ICCE 2014

The 22nd International Conference on Computers in Education

Workshop Proceedings

Nov 30 - Dec 4, 2014 Nara, Japan

Workshop Proceedings of the 22nd International Conference on Computers in Education ICCE 2014

November 30, 2014 - December 4, 2014
Nara, Japan

ISBN 978-4-9908014-2-7

Publisher
ICCE 2014 Organizing Committee, Japan
ICT Unit, Center for Graduate Education Initiative, Japan Advanced Institute of Science and Technology
1-1, Asahidai, Nomi, Ishikawa, 923-1292, Japan

Editors

Ying-Tien WU
Thepchai SUPNITHI
Tomoko KOJIRI
Chen-Chung LIU
Hiroaki OGATA
Siu Cheung KONG
Akihiro KASHIHARA

PREFACE

Welcome to the Workshop Proceedings of the 22nd International Conference on Computers in Education (ICCE 2014), held from November 29th to December 4th in Nara, Japan.

Established in 1989, ICCE is now an annual international conference organized by the Asia-Pacific Society for Computers in Education (APSCE), and it has become a major event for scholars and researchers in the Asia-Pacific region to share ideas and to discuss their works in the use of technologies in education.

This year we accepted 19 proposals—16 workshops, two interactive events, one tutorial and we aim to explore focused issues in various themes related to the use of technologies in education. Each proposal was peer-reviewed by international reviewers with relevant expertise to ensure high-quality work. All the workshops, organized by the International Program Committees, are in the min-conference format and this proceeding includes all the papers submitted to our workshop. We believe that the workshop provides a valuable opportunity for researchers to share their works and to seek further collaboration. Papers of our workshop will certainly stimulate more interesting research works in these relative areas in Asia-Pacific countries and beyond. We hope that readers will find the newly ideas relevant to their research works in this proceeding.

Finally, we would like to thank the Executive Committee of the Asia-Pacific Society of Computers in Education and the ICCE 2014 International Program Committee Coordination Chair and Co-Chair for entrusting us with this important task of chairing the workshop program and giving us an opportunity to work with many outstanding researchers. We would also like to thank the Local Organizing Committee for helping with the logistic of the workshop program.

Workshop, Tutorial and Interactive Event Coordination Co-Chairs:

Ying-Tien WU, National Central University, Taiwan
Thepchai SUPNITHI, National Electronics and Computer Technology Center, LST, Thailand
Tomoko KOJIRI, Kansai University, Japan

TABLE OF CONTENTS

Workshop 1: Natural Language Processing Techniques for Educational Applications (NLP-TEA), with a Shared Task on Grammatical Error Diagnosis for Learning Chinese as a Foreign Language (CFL)

Workshop 8: Computer-Supported Personalized and Collaborative Learning

Workshop 11: International Workshop on "Technology Enhanced Language Learning"

Workshop 13: Technology Enhanced Learning of Thinking Skills (TELoTS)

Workshop 14: The 1st Workshop on Application of Advanced Technologies in Environmental Education

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

A Sentence-Pattern Learning Support System for Japanese

Takahiro Ohno*, Ayato Inoue* & Dongli Han**
**Graduate School of Integrated Basic Sciences, Nihon University, Tokyo, Japan*
**Graduate School of Integrated Basic Sciences, Nihon University, Tokyo, Japan*
***Department of Information Science College of Humanities and Sciences, Nihon University, Tokyo, Japan*
**han@chs.nihon-u.ac.jp

Abstract: Computers have played more and more important roles in Japanese education. We in this paper focus on sentence-patterns in Japanese and have developed a Web system to support sentence –pattern learning. Specifically, our system is able to indicate a sentence pattern from a freely composed Japanese sentence by using a digital sentence-pattern dictionary generated in advance. After that, our system determines whether the sentence pattern has been used properly and gives some instructions on how to use the sentence pattern correctly. Experiments show the effectiveness of our method.

Keywords: Japanese education, learning materials, Web application, sentence pattern, free composition oriented system

1. Introduction

The number of Japanese learners all over the world is increasing every year. According to a report by Japan Foundation (2013), the total number has reached approximate 4 millions spreading in 136 countries, which is 9.2% larger than the total number counted four years before. On the other hand, a parallel investigation towards all the Japanese education institution around the world shows that the lack of learning material and facility, the monotony in learning method, and the learners' lukewarm attitude still remain as the most serious problems in Japanese learning. Computers and learning support tools based on Web will undoubtedly provide a good solution to the above problems more or less. They might be helpful not only to the learners but also to the teachers as well.

Many e-learning systems have been developed to help users learn Japanese in a simple manner with true-false or fill-in-the-blank questions only (e.g., Samidori[1], WebCM[2], and Nihongo-Dekimasu[3]). In contrast, free-composition oriented systems are considered more helpful but more difficult to implement. Here are some existing web systems based on free composition. Obi2 analyzes an input sentence and classifies it into 13 difficulty-levels (Sato etc., 2008). Asunaro conducts dependency-relation analysis on input sentences (Abekawa etc., 2002). Reading Tutor explains each morpheme contained in an input Japanese sentence in six different languages, and discovers the sentence-patterns used (Kawamura, 2002). These systems support free composition and hence make themselves more practical and feasible. However, none of the above systems is capable of handling wrong texts, which is more realistic and hence has more actual significance. Here, a wrong text indicates the text containing grammatically wrong usages.

In this paper, we focus on sentence-patterns in Japanese and aim to develop a free-composition oriented Web system for Japanese learners in all levels. Specifically, our system is able to recognize a sentence pattern from an input Japanese sentence by using a digital sentence-pattern dictionary

[1] http://www.ku-japanese.jp/
[2] http://opal.ecis.nagoya-u.ac.jp/webcmj/
[3] http://www.erin.ne.jp/

generated in advance. After that, our system determines whether the sentence pattern has been used properly and gives some instructions on how to use the sentence pattern correctly. The difference between our system and Reading Tutor is that our system does not expect users to input correct sentences at all, whereas Reading Tutor only accepts sentences containing correctly used sentence patterns.

In the rest of this paper, Section 2 describes the pre-created sentence-pattern dictionary, Section 3 details the discovering and correcting process of sentence-pattern, and Section 4 shows the experimental results for our approach. Finally, we end this paper with a conclusion in Section 5.

2. Sentence-pattern Dictionary

A Japanese sentence pattern is composed of a set of words in a fixed order to express some particular meaning (Han and Song, 2011). A simple example is "～しだい" meaning *as soon as*. The symbol "～" is a placeholder where only expressions satisfying some certain conditions could be inserted. For the sentence pattern "～しだい" here, only two kinds of expressions could be used to replace "～" in front of "しだい":

(1)predicative forms of verbs
(2)Sahen-verbs which are formed by adding "する" to action nouns

Sentence patterns are supposed to be one of the most difficult issues during the process of learning Japanese. And we consider it is necessary and important to make users aware of their own usage of sentence patters when composing Japanese documents.

The first step we take to build a learning support system with the capability of discovering and correcting sentence-patterns is generate a sentence-pattern dictionary. Totally, seven types of structures exist in Japanese sentence patterns as shown in Figure 1. As stated above, "～" in Figure 1 is a placeholder, and each symbol other than "～" stands for a fixed expression contained in a particular sentence pattern. For example, the sentence pattern "～しだい" is classified into Type 1, "～○".

Type. 1	～○
Type. 2	～○～
Type. 3	～～○
Type. 4	○～△
Type. 5	～○～△
Type. 6	～○～△～□
Type. 7	～○～△～□～

Figure 1. Structure Types of Sentence Patterns.

Each entry included in the Japanese sentence-pattern dictionary (Ask Shuppan, 2008) is first classified into one of the above seven structure-types, then segmented into multiple elements according to the number of "～" and other symbols, and finally stored as a whole record into our digital dictionary. In the second step, the matching conditions to be examined for each "～" have been rewritten to conform to Cabocha[4], a free morphological analyzer, which is used to analyze input sentences later in Section 3. Our digital dictionary is composed of 371 records, each containing the Structure-type, the matching conditions as described above, and other necessary information as well. One of the authors has spent 1.5 months in editing this digital dictionary manually.

[4] https://code.google.com/p/cabocha/

3. Sentence-pattern Examination

This section describes the flow of our system. After a Japanese sentence is composed, it is first segmented into multiple morphemes using Cabocha at the back. Then the pre-created digital dictionary is employed to examine whether the sentence is likely to contain a sentence pattern. If the answer is positive, another examination will be conducted to see whether the sentence pattern has been properly used. Finally, a feedback will be prompted to the user telling the examining result and a guide as well in case the usage is not correct. Figure 2 shows the algorithm for the specific case "～○".

Step1. Seek "○" in the input sentence.
Step2. Get the part-of-speech (POS) and conjugation information of "～" using Cabocha if Step1 returns a success.
Step3. Compare the POS and the conjugation information of "～" with those in the sentence-pattern dictionary for the corresponding "～○".
Step4. Prompt the user with a guide on how to use the sentence pattern properly in case a mismatch occurs in Step3.

Figure 2. Examining Steps for Sentence Pattern "～○"

Algorithms for examining other structure types in Figure 1 are similar to "～○", but a little more complicated. A screen shot of our system is shown as Figure 3.

Figure 3. A Screen Shot of the Interface of Our Web System

4. Experiments for Sentence-pattern Examination

We have conducted some experiments to examine the effectiveness of our approach. We employ two kinds of test dataset: D1 and D2. D1 is composed of 200 correct illustrative sentences extracted from a book on Japanese sentence patterns (ALC Shuppan 2007). Each illustrative sentence in D1 contains at least one sentence pattern. D2 contains 200 sentences which have been extracted randomly from a corpus (Tomoya Mizumoto etc., 2011) generated by using Lang-8, a language learning platform where native speakers correct what learners have composed[5]. Compared with D1, D2 is much closer to the practical input sentences. We hope to observe the difference arising from correctly used sentence patterns and the opposite by this means.

Table 1: Experimental results for D1

Number of sentence patterns discovered			287
Number of true sentence patterns contained			272
Number of true sentence patterns discovered			267
Number of sentence patters with correct feedbacks			253
Precision	Recall	F-value	Feedback Precision
0.93	0.98	0.96	0.95

Table 2: Experimental results for D2

Number of sentence patterns discovered			311
Number of true sentence patterns contained			180
Number of true sentence patterns discovered			179
Number of sentence patters with correct feedbacks			167
Precision	Recall	F-value	Feedback Precision
0.58	0.99	0.73	0.93

Table 1 and Table 2 show the experimental results on D1 and D2. Precision, Recall, and F-value are calculated to measure how effective our approach is in discovering sentence patterns. Feedback Precision indicates the success rate by which proper feedbacks have been given for correctly discovered sentence patterns. A significant performance degradation could be observed when the test dataset varies from correct sentences to real-world texts, whereas the latter is what we have to consider and important when evaluating learning support systems developed especially for less capable users.

Table 3: New experimental results for D1

Number of sentence patterns discovered	266
Number of true sentence patterns contained	272
Number of true sentence patterns discovered	252
Number of sentence patters with correct feedbacks	241

[5] http://lang-8.com/

Precision	Recall	F-value	Feedback Precision
0.94	0.93	0.94	0.96

Table 4: New experimental results for D2

Number of sentence patterns discovered	230
Number of true sentence patterns contained	180
Number of true sentence patterns discovered	164
Number of sentence patters with correct feedbacks	154

Precision	Recall	F-value	Feedback Precision
0.71	0.91	0.80	0.94

Based on the initial experimental results, we have modified Step1 in Figure 2 by incorporating morphological analysis. After locating "〇" in the input sentence, we conduct a morphological analysis on the whole sentence to see whether "〇" is segmented as a separate morpheme or partially attached to other morphemes around. Only sentences containing separate "〇" are left over for further process. Table 3 and Table 4 show the new experimental results.

It is obvious that tightening the conditions for discovering sentence-patterns improved the system's performance, especially for real-world texts. The Precision has increased from 0.58 to 0.71, indicating the effectiveness of the modification in algorithm. Totally, we have obtained a reasonable F-value for discovering sentence patterns and a satisfying Feedback Precision. We believe the system could be helpful for less capable users who are trying to use sentence patterns during the process of free composition. Among all the sentence patterns discovered by the system, about 30% might not really be sentence patterns according to the experimental results in Table 4. However, we don't consider this as a major issue. Users could easily ignore the over-discovered sentence patterns in texts composed by themselves.

However, there are some drawbacks with this approach.

1. false analytical results from the morphological analysis
2. incompleteness of the sentence-pattern dictionary.
3. confusions between normal expressions and sentence patterns.

The first problem arises from the accuracy of the morphological analyzer we have been using. The second one is attributable to the exhaustiveness of the sentence-pattern dictionary we adopt. Some conditions to be examined for each "～" in a sentence pattern are not exhaustive. And this leads to some discovering failures as a result. We might be able to remove the these problems partially from our system by trying other analytical tools or sentence-pattern books. The last problem indicates another case where a normal expression is over-discovered as a sentence pattern. Here is an example.

Input:

ともだちのミシャールさんのうちにあそびに行きました。

(I hang out with Mishaal to her house.)

Feedback:

「～うちに」の文型を使っているようです。 - OK

(Correct usage with the sentence pattern "～うちに")

The system has recognized “うちに” as a sentence pattern meaning *before you know it / while*, and prompted a message showing that the user has used this sentence pattern correctly, whereas “うちに” could also be used to express *to someone’s home* which is correct here according to the context of this input sentence. This issue is hard to handle and we might need some extra statistical information to make better decisions.

5. Conclusion

We focus on sentence-patterns in Japanese and have developed a Web system to support sentence –pattern learning. Specifically, our system is able to discover a sentence pattern from an input Japanese sentence by using a digital sentence-pattern dictionary generated in advance. After that, our system determines whether the sentence pattern has been used properly and gives some instructions on how to use the sentence pattern correctly if the answer is negative. The difference between our system and Reading Tutor, an existing system also capable of discovering sentence patterns, is that our system does not expect correct input sentences at all, whereas Reading Tutor only accepts sentences containing sentence patterns that have been used correctly.

Experiments have been conducted to examine the effectiveness of our approach. On the whole, we have obtained a reasonable F-value for discovering sentence patterns and a satisfying Feedback Precision. We believe the system could be helpful for less capable users who are trying to use sentence patterns during the process of free composition.

Our future tasks include increasing the precision further in discovering sentence patterns, and a questionnaire by less-capable Japanese learners to examine the effectiveness of our Web System through practical application.

References

Japan Foundation, (2013). http://www.jpf.go.jp/j/japanese/survey/result/survey12.html

ASK Shuppan. (2008). “Ikita Reibun De Manabu Nihongo Hyougen Bunkei Jiten”. Japan. (in Japanese)

ALC Shuppan. (2007). “Donna Toki Dou Tsukawu Nihongo Hyougen Bunkei Jiten”. Japan. (in Japanese)

Satoshi Sato, Suguru Matsuyoshi, and Yohsuke Kondoh (2008). Automatic Assessment of Japanese Text Readability Based on a Textbook Corpus. *Proceedings of LREC-08.*

Takeshi Abekawa, Nishina Kikuko, Manabu Okumura, Yutaka Yagi, Norihisa Totsugi, Sigeki Sugimoto and Liang Fu (2002). Development of Japanese Reading System of Multi-Lingual Environment "Asunaro." *Proceedings of 3rd International Conference on Computer Assisted Systems for Teaching & Learning/Japanese*, 71-74.

Yoshiko Kawamura, (2002). Implementation and Evaluation of a Computer-Based Reading Comprehension Program for Students from Non-Kanji Backgrounds, *Proceedings of CASTEL/J'2002,* 147-150.

Tomoya Mizumoto, Mamoru Komachi, Masaaki Nagata and Yuji Matsumoto. (2011). Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners, *Proceedings of the 5th International Joint Conference on Natural Language Processing (oral)*, 147-155.

Han, D., and Song, X. (2011). “Japanese Sentence Pattern Learning with the Use of Illustrative Examples Extracted from the Web”, *IEEJ Transactions on Electrical and Electronic Engineering*, Vol.6, No.5, 490-496.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

CYCCDC: A Chiayi Chinese Conversation Dialogue Corpus

Jui-Feng Yeh, Yun-Yun Lu, Yi-Syun Tan
Department of Computer Science and Information Engineering, National Chiayi University
No.300 Syuefu Rd., Chiayi City 60004, Taiwan (R.O.C.)
ralph@mail.ncyu.edu.tw, s1020447@mail.ncyu.edu.tw, s1000441@mail.ncyu.edu.tw

Abstract: Speech is one of the most natural ways of communication between human. In recent years, the spoken dialogue systems on human machine interaction (HMI) is more and more popular. In order to develop effective and natural human machine interaction, the corpus collected is relatively more important. Due to various types of corpus, classifying the corpus is a needed process. In this paper, we collected, transcribed, and classified the corpus, we named ChiaYi Chinese Conversation Dialogue Corpus (CYCCDC). We collected this corpus with multiple ways, and then we arranged and classified this corpus. The corpus includes multiple useful information to research spoken dialogue system and human communication field. The CYCCDC includes tourism information, food information, clothing, housing information, traffic information and part of Orange Technology, such as elderly health care and accident handling. This corpus can be extensively applied, such as tourism plan in spoken dialogue system.

Keywords: CYCCDC, Chinese corpus, tourism, orange technology, spontaneous speech.

1. Introduction

With the advance of the internet and technology, people chat with each other not only through voice. People can also make a communication by text with various well-developed online instant messaging software, such as Skype, Line, Aim, and so on. Since text is easier to record than voice, it is helpful for searching history record or finding out particular conversation information with proper classification in the future. There is some corpus analysis as (Agrawal, 2011) which is classified the emotions of the Hindi corpus. In (Jia et al., 2011), the present study systematically states the construction of the corpus on the English learners in Asia. There is also some conversation corpus, which has been collected in (Bechet et al., 2012), the goal of this paper is to reduce the development cost of speech analytics systems by reducing the need for manual annotation. In (Takezawa et al., 2002), they collected the travel conversation corpus and a broad-coverage bilingual basic expression corpus, and they compared the characteristics of vocabulary and expressions between these two corpus. In this paper, we collect the conversation transcript documents and analyze the topic classification. Finally, we gave each conversation script a topic classification. In particular, the corpus of care and accident handling are classified into at the category of Orange Technology (Wang & Chen, 2011). Orange technology is the idea mentioned by the National Cheng Kung University professor Jhing-Fa Wang. The main idea of orange technology is to bring health, happiness and care for human. It also include elderly health care and child care. Besides, people can enhance safety and quality of humanism to foreign culture with orange technology. We also collected other corpus on several topics: tourism, food, sports, solicitude and others. Each topics can be classified into more categories. Ohtake et al. proposed a new corpus of consulting dialogues which is designed for training a dialogue manager (Ohtake et al., 2010). They also collected more than 150 hours of tourist guidance dialogue. In section 2, we describe the method how we collect corpus. Then we explain how we arrange and classify in section 3. In section 4, we show that our corpus can be applied in various fields. Eventually, we discuss the future work and conclusions in section 5 and 6.

2. Method

In this paper, We collected the Chiayi Chinese conversation dialogue corpus to increase the variability of response sentence for the spoken dialogue systems. Our approach is to ask at least two people for a chat and record their conversation. There are several rules as following:

- Sentence composed by the conversation dialogue (unlimited number of the sentences)
- Period of each conversation occurred at least 8 minutes
- Each dialogue included at least two topics
- Each conversation took by native Chinese speakers

The rule 1 is established to make the system spoken dialogue more humane. The purpose of the rule 2 is to let speakers converse in a spontaneous way. The rule 3 is helpful to increase the variability of contents that can enrich response sentence. In order to reduce Chinese grammar errors and obtain a native Chinese spoken dialogue, we augmented the rule 4. We obtained 392 audio files which length is at least 8 minutes, then turned the audio files into the transcribed text files. There are at least 20 sentences in each transcribed file. The sentences are related to the fields of tourism information, elderly care and accident disposal, which included health, unexpected events, food, traffic, housing, clothing and others. An example of the conversation is shown in Figure 1.

Speaker A：這個周末我也要去嘉義玩
Speaker A：希望也能有好天氣。
Speaker B：你要去嘉義哪裡玩？
Speaker A：先去看射日塔之後再去阿里山看日出吧。
Speaker B：二二八紀念碑也可以去看看。
Speaker A：要怎麼去呢?
Speaker B：你可以坐公車去。
Speaker A：附近有加油站嗎?
Speaker A：我想開車去嘉義。
Speaker B：附近過橋有一家中國石油。
Speaker A：嘉義有甚麼好吃的在地小吃嗎?
Speaker B：有一間簡單火雞肉飯不錯吃。
……

Figure 1. An example of the conversation.

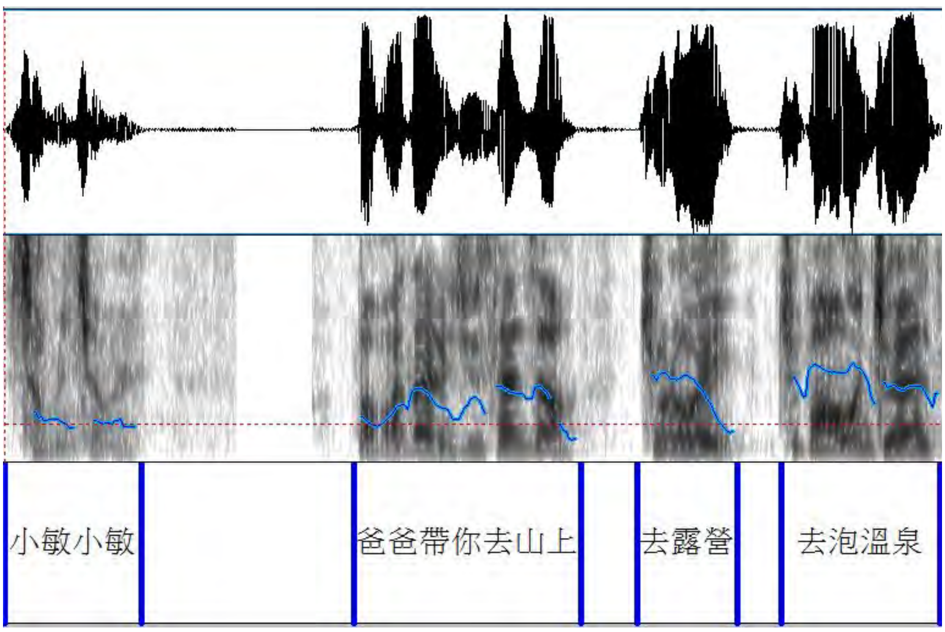


Figure 2. Spectrogram of the recording sentence.

First, the conversations must be recorded in a quiet room. Speakers need a computer, a microphone, and the recording software Praat before starting a conversation.. Audio file format is designated as the WAV format, with a bit rate of 256 kbps and a sampling rate of 16 kHz. The sampling resolution in 16 Bit is recorded in a mono mode with the PCM. The speakers can think about what they want to talk about before recording. Each conversation should be took at least 8 minutes and recorded at least 24 minutes in total. An example of Spectrogram is shown the show in figure 2.

The Chiayi Chinese conversation dialogue corpus has been recorded by 199 speakers, which are the students of National Chiayi University, Taiwan. There are 88 native Taiwan females and 111 native Taiwan males. The age range of these speakers is from 18 to 21. In total, there are 399 audio files and 27.5 hours of conversation speech. Consequentially, the Chiayi Chinese conversation dialogue corpus comprised 9,734 sentences of 399 transcribed text files.

3. Corpus Analysis

Table 1: Instruction of the topic.

Topics	Instruction
Tourism	Dining Location Lodging Location Modes of Transport (Living Travel) Navigation Information (Modes of Transport) Traveling Spot Shopping Information (Living Travel) Cost of Time Weather/Climate Care Agency (Location)
Food	Dinging Place (Restaurant、Snack Bar) Transportation (The Modes of Transport Go for Dining) Cost of Time (Meal Time)
Healthy Care	Transportation (Healthy Care) Navigation Information (Healthy Care) Shopping Information (Healthy Care) Weather/Climate (Health Care) Health Status
Accident Handle	Emergency Incident Detection (Fall Down, Faint, Asthma Attack)
Recreation	Leisure Entertainments (Knowledge/Reading, Recreational, Artistry, and so on)
Solicitude	Health State Greeting Consolation Family/Neighbor Trivia Working And so on
Sports	Ball Running Swimming Hiking And so on
Others	Not classified in the above topics

In this section, we describe how we design and classify these collected corpus. Due to a wide variety of corpus, we sorted all the corpus before the classification. It will make developers more convenient while using our corpus. And we will introduce main topics in section 3.1, then spoken speech phenomenon in section 3.2. In section 3.3 is describe about what is our method to distinguish topic in the sentences.

A “topic” is an abstract or a representative summary of the contents in dialogues. Developers can comprehend key points with the information of the topic instead of reading whole contents of dialogues. We can also classify conversations into appropriate categories. For example, a speaker says: “這個週末連假我們去嘉義的阿里山走走吧”, this sentence not only can be regarded as an effect of the topic "tourism" from a conversation, but also speculate from the keyword that the topic of this diaglogue might be “tourism”. If there exists a topic in one interpersonal conversation this conversation would not be interrupted easily. Besides, this uninterrupted conversation can be continued to develop new contents or changed into different topic timely can motivate speakers for keeping the interpersonal conversation. Therefore, we designed 8 topics to classify the corpus. There are tourism, food, healthy care, accident handle, recreation, solicitude, sports and others. Each topic can be subdivided into a lot of classes as shown in table 1. According to transcription which is tagged each topic by speakers. There are 2,064 sentences in tourism, 1,042 sentences in food, 177 sentences in healthy care, 268 sentences in accident handle, 1,801 sentences in recreation, 2,128 sentences in solicitude, 513 sentences in sports and 1,741 sentences in others. The distributing graph is shown in figure 3. There are several issues need to be discussed.

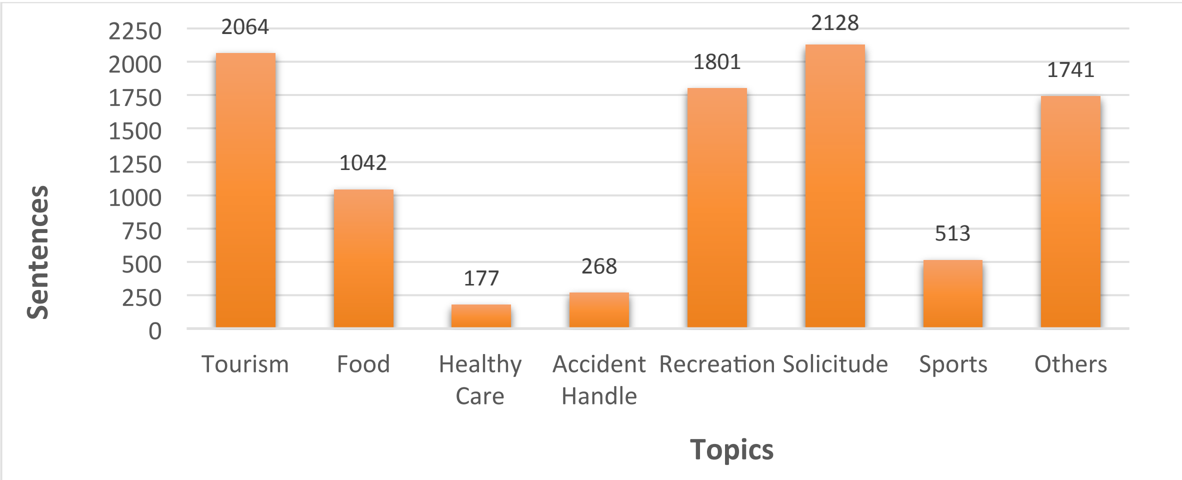


Figure 3. The distributing graph of each topics sentence.

3.1 *Main Topics*

According to human thought, the dialogue is nothing more than their food, clothing, housing, traffic and entertainment for main topic. For example, an old sick people want to travel, then he goes to some fun places, eat goodies and experience what he could not engage in leisure activities during illness. In this example, we designed the 7 topics, including the examples mentioned in the travel and food, while healthy Care, accident handle and solicitude are especially designed for the elderly.

3.2 *Spoken speech Phenomenon*

Through listening to the recorded speech, we found that corpus contains many voices speaking tone, because people thinking and emotion factors. Due to the recording of the conversation dialogue is not like reading speech which has a transcript. There is some spoken speech phenomenon. The most common of spoken speech phenomenon is particles, and the other is paralinguistic phenomena, the situation of pronunciation is not correct that is pronunciation error, and non-native language (Chang et al., 2005). There are an example of paralinguistic phenomena shown in figure 4. Because of the speakers conversed spontaneously, we classify those corpus to other topic.

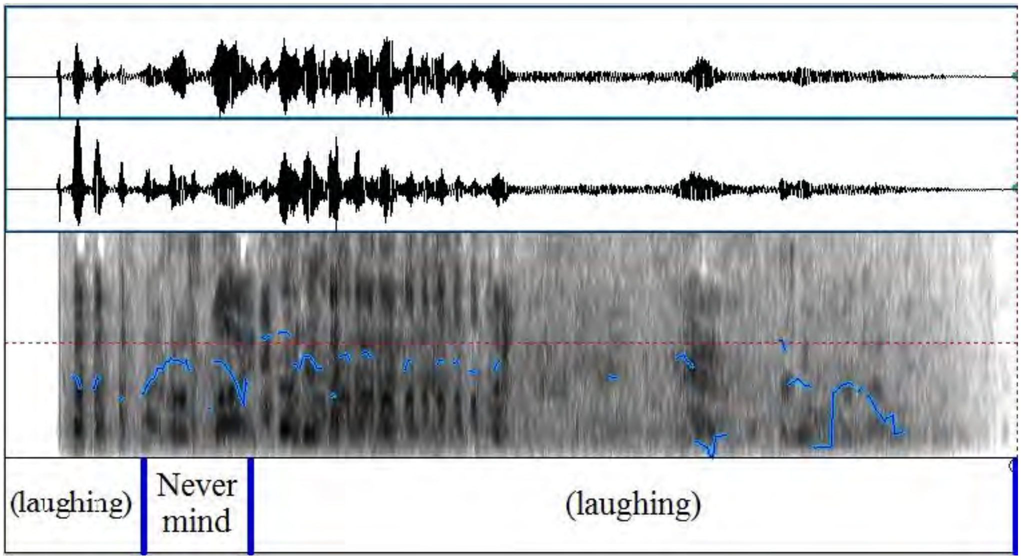


Figure 4. Spectrogram of the laughter in conversation (paralinguistic phenomena).

3.3 *Keywords of topic*

For example, a speaker A says: “最近有聽說什麼好吃的美食嗎?”, a speaker B says: “有阿!聽說在嘉義奮起湖有好吃的鐵路便當”. There has different keyword of topics in this example. We will judged according to the context of the sentence belongs to what topics.

4. Applications

The most application of corpus is automatic speech recognition (Wu et al., 2014). There is also another application of corpus, such as generating responses sentences. Since the CYCCDC contains many different topics of sentences and rich vocabularies, the CYCCDC can applied in many aspects. For instance, developers can utilize the CYCCDC to generate responses sentences with some tour information for a tourism planning system, or refine an existing dialogue system with corpus of solicitude. The developers also can exploit the CYCCDC to analyze what do people talk about or when they want to have a travel. The CYCCDC is a conversational corpus so that the generated responses sentences of a dialogue system can make users feel like having a conversation with an actual human. This type of dialogue system is called as chat oriented dialogue system. Banchs et al. have also proposed the informal response interactive system (IRIS) (Banchs et al., 2012), which is a chat-oriented dialogue system based on the vector space model framework.

5. Future Work

In the future work, we will continue to improve the consistency of the database. With more corpus from the refined CYCCDC, all the related applications can be expected to be improved greatly, and can be applied in more studies. The spoken speech phenomenon also is still a research issue.

6. Conclusion

In this paper, our corpus was recorded and transcribed by the speakers. Different from other common corpus, the Chiayi Chinese conversation dialogue corpus is based on conversation for academic research but also substantial contribution. We designed 8 topics of the CYCCDC, which included the vast majority of tourism and solicitude. The native regional range of CYCCDC involved Yunlin, Chiayi and Tainan in Taiwan. Thus, the corpus also can apply in tourism planning system which focused on these regions. Although all the sentences are still in the testing phase, but the quality of the sentences is adequate enough for doing researches. The other parts of corpus in CYCCDC are mainly collected for the Orange Technology, which focused on health care and accident handling for children and elders. The CYCCDC also involoved some dialogue about disadvantaged groups or people who need social care. This part of the corpus can assists with the academic research of society. Certainly, the CYCCDC also contained some of common corpus, such as recreation, sports and other categories that are irrelevant to the main collection. Because the CYCCDC is a real conversation between humans, the CYCCDC is also helpful for a chat-oriented dialogue system.

Acknowledgments

This work is supported in part by the National Science Council, Taiwan, R.O.C., under the project grant numbers NSC 102-2221-E-415-006-MY3.

References

Agrawal, S. S. (2011, October). Emotions in Hindi speech-analysis, perception and recognition. In Speech Database and Assessments (Oriental COCOSDA), 2011 International Conference on pp. 7-13. IEEE.

Jia, Y., Wang, M., Zhai, H., & Li, A. (2011, October). Construction of speech corpus of AESOP-SD. In Speech Database and Assessments (Oriental COCOSDA), 2011 International Conference on pp. 42-46. IEEE.

Bechet, F., Maza, B., Bigouroux, N., Bazillon, T., El-Beze, M., De Mori, R., & Arbillot, E. (2012). DECODA: a call-centre human-human spoken conversation corpus. In LREC on pp. 1343-1347.

Takezawa, T., Sumita, E., Sugaya, F., Yamamoto, H., & Yamamoto, S. (2002, May). Toward a Broad-coverage Bilingual Corpus for Speech Translation of Travel Conversations in the Real World. In LREC on pp. 147-152.

Wang, J. F., & Chen, B. W. (2011). Orange computing: challenges and opportunities for awareness science and technology. In 2011 3rd International Conference on Awareness Science and Technology (iCAST) pp. 533-535.

Ohtake, K., Misu, T., Hori, C., Kashioka, H., & Nakamura, S. (2010). ”Dialogue acts annotation for NICT Kyoto tour dialogue corpus to construct statistical dialogue systems” Seventh International Conference on Language Resources and Evaluation, LREC 2014, pp. 2123-2130.

Wu, C., Shen, H., & Yang, Y. (2014) “Chinese-English Phone Set Construction for Code-Switching ASR Using Acoustic and DNN-Extracted Articulatory Features” IEEE/ACM Transactions on Audio, Speech & Language Processing 22(4), IEEE, 2014, pp. 858-862.

Banchs, R. E., & Li, H. (2012). “IRIS: a Chat-oriented Dialogue System based on the Vector Space Model” Proceedings of the Association for Computational Linguistics 2012 System Demonstrations, pp. 37-42

Chang, L. H., Wang, Y. R., & Chen, S. F. (2005) “Evaluation of Mandarin Broadcast News Transcription System” ROCLING 2005

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Partial and Synchronized Caption Generation to Develop Second Language Listening Skill

Maryam Sadat MIRZAEI [a*], Yuya AKITA [a], Tatsuya KAWAHARA [a]
[a] *Graduate School of Informatics, Kyoto University, Japan*
*maryam@ar.media.kyoto-u.ac.jp

Abstract: Captioning is widely used by second language learners as an assistive tool for listening. However, the use of captions often leads to word-by-word decoding and over-reliance on reading skill rather than improving listening skill. With the purpose of encouraging the learners to listen to the audio instead of merely reading the text, the study introduces a novel technique of captioning, partial and synchronized, as an alternative listening tool for language learners. Using TED talks as a medium for training listening skill, the system employs the ASR technology to synchronize the text to the speech. Then, the system uses the learner's proficiency level to generate partial captions based on three features that impair comprehension: speech rate, word frequency and specificity. To evaluate the system, the performance of Kyoto University students in two CALL classes was assessed by a listening comprehension test on TED talks under three conditions: no caption, full caption and the partial-and-synchronized caption. Results revealed that while reducing the textual density of captions to less than 30%, the proposed method realizes comprehension performance as well as full caption condition. Besides, it performs better than other conditions on new segments of the video without captions.

Keywords: Computer-Assisted Language Learning, Automatic Speech Recognition, Listening Comprehension, Word Frequency, Speech Rate

1. Introduction

The process of learning a foreign language involves mastering different skills such as listening, speaking, reading and writing. Of these, acquiring listening often entails a complex cognitive process and demands the use of different strategies which in turn make a phase of frustration for many language learners (Leveridge and Yang, 2013). In order to improve listening, one must be exposed to authentic and comprehensible input. Authentic input, however, makes listening more challenging especially when the phonological systems of the first and the second language are distant (e.g. Japanese vs. English).

Listeners can overcome this problem by benefiting from assistive tools such as "captioning" that textualizes the verbatim speech and makes it more recognizable through neatly dividing the word boundaries. Nevertheless, when it comes to using captions, both language learners and teachers face a dilemma. In fact, when reading captions is part of watching a video, learners often rely on their reading skill to compensate for their listening skill deficiencies, whereas in a real-world communication, learners should solely use their listening skill as no assistive tools are available.

To address these issues, this study proposes a new method of captioning, "partial and synchronized" as an alternative tool for enhancing second language (L2) learners' listening comprehension skills. The term "synchronized" captioning is to present caption text word by word aligned in precise timing with the speech signal of the respective words, which effectively shows the correspondence between words and the audio channel. This method is realized by the automatic word-level alignment feature of automatic speech recognition (ASR) technology, which precisely maps each word to its corresponding speech signal. In the "partial" captioning method we select a subset of words from the transcript and present them in the caption while hiding the rest of the words. Although seems similar to keyword captioning, in this method "important" words are not the selection criteria. Instead, words that impair comprehension or the ones beyond the learner's current level of competence form the basis of this selection. Moreover, the selection of keywords is content-specific and does not consider the proficiency level of the learners, whereas the features of the proposed method are tuned to the learner's knowledge to meet the requirements of each individual.

Unlike conventional captions, in Partial and Synchronized Captioning, comprehension cannot be gained by solely reading the captions, but by listening to the audio and reading only for difficult or unrecognizable words. Thus the method is effective for reducing learners' dependence on captions.

Following this introduction, this paper reviews previous studies and describes the proposed technique of captioning. Then, the experimental procedure together with the results is demonstrated and a discussion over the findings is addressed. This paper ends with conclusion and future directions.

2. Literature Review

2.1 Captioning and L2 Listening Comprehension

To overcome the listening problems, assistive materials, such as captions, are used to help L2 listeners. Captioning is defined as "visual text delivered via multimedia that matches the target language auditory signal verbatim" (Leveridge and Yang, 2013, p.1). Captions neatly demonstrate the word boundaries without being affected by accent, pronunciation and audio deficiencies (Vanderplank, 1993) and allow the learners to parse the speech stream into meaningful chunks, an essential process for learning (Ellis, 2003). A considerable amount of literature has been published on various beneficial effects of captions. Some of these studies have investigated the effect of captioning on vocabulary acquisition (Bird and Williams, 2002; Griffin and Dumestre, 1992), reading development (Bean and Wilson, 1989), word recognition (Bird and Williams, 2002; Markham, 1999) and listening comprehension (Danan, 2004; Garza, 1991; Markham, 1999; Montero Perez et al., 2014; Vanderplank, 1993; Winke et al., 2010).

For instance, Garza (1991) conducted an experiment with 70 high-intermediate learners of English and 40 three to four year learners of Russian, and assessed their comprehension of videos with/without captions. His results indicated significant improvement on the captioning condition in both groups. Studies in Japan such as Suzuki (1996) reported the positive effect of English caption on Japanese listening comprehension development.

The type and manner of captioning may influence the effect of this assistive tool on language learning. Garza (1991) suggests using various types of open captioning, such as verbatim, paraphrase and keywords as means of training listening skill.

2.2 Aligned and Synchronized Captioning

Correspondence between caption and speech may also affect the learning process. Advancement of speech technology has enabled precise text-to-speech alignment. Munteanu et al. (2007) used ASR to generate transcripts of webcast lectures for examining native speakers' comprehension on the videos. They found out that ASR generated transcripts are useful when word error rate (WER) is lower than 20%. This finding was generalized to L2 learners in a study by Shimogori et al. (2010) who suggest that captions with 80% accuracy improve the understanding of Japanese learners of English.

Accordingly, "karaoke-style" display, where the text is highlighted in colors as the audio moves by, has gained some instructional value. Bailly and Barbour (2011) developed a system that exploits the alignment of text with audio at various levels (letters, phones, syllables, etc.). This system uses a data driven phonetizer trained on an aligned lexicon of 200,000 French entries to display a time-aligned text with speech at phoneme level. The results showed that the multimodality of synchronous reading systems is beneficial for overcoming the problem of word decoding in a text/audio-only environment.

It should be noted that this method may lead to over-reliance on the caption and needs to be refined. This can be accomplished through highlighting only particular words or sentences in the caption, as in keyword captioning.

2.3 Keyword Captioning

Guillory (1998, p.95) examined the use of keyword captioning for learners of French. The results demonstrated that students who received keyword captions performed as well as those who received full captions. Guillory discussed that “learners no longer need to be subjected to a volume of text to read; they can in fact comprehend authentic video with considerably less pedagogical support”.

In a recent study by Montero Perez et al. (2014), the perceived effectiveness of keyword captioning is criticized. The study investigated the effect of full text captions and keyword captions versus no captioned condition. The results demonstrated that full captioning group outperformed the other two groups on the global comprehension questions while both the keyword captioning and the no-captioning group had equal performance on this test. Analysis of the responses received from the keyword-captioning group revealed that this type of captioning is distractive. According to the researchers, a plausible explanation may be the salient and irregular appearance of the keywords on the screen, which causes distraction. However, not every learner can benefit from presenting the keywords in captions since the selection of keywords is content-specific and may not provide each learner with his/her required amount of support. In line with this assumption, Guillory (1998) noted that the keyword captions used for her study contained a tiny portion of the total script, which may not have provided enough information for the beginners.

2.4 Limitations on Captioning

In spite of the beneficial aspects of captioning, there are some criticisms on the use of this assistive tool. It is skeptical whether learners provided with captions are training their listening or their reading skills. Kikuchi (1995) examining subtitles in Japanese and captions in English reported that students who watched the movie with Japanese subtitles merely read the text without listening to the movie. Using an eye tracker, Winke et al. (2013) investigated learners’ use of captions and reported that learners read the captions on average 68% of the time it is on the screen.

On one hand, the learner needs to be able to deal with real-world situation where there is no access to any supportive tool, and on the other hand we cannot expect a non-native listener to follow the authentic input without any support. Hence, the listening instruction should focus first and foremost on assisting the language learners to cope with aural input difficulties while maintaining a tendency to develop compensatory strategies for listening in real-time. Thus, further research should be conducted to investigate an effective method for assisting learners to gain adequate comprehension, without becoming too much dependent on captions.

3. Proposed Method: Partial and Synchronized Captioning

We propose a new type of captioning called Partial and Synchronized Captioning (hereinafter, PSC). In this method the text is synchronized to the speech in word-level and only a subset of words are shown in the caption while the rest are masked to keep the learner listening to the speech. Thus, this method consists of two components; synchronization and partialization where the two are complementary and counteract the demerits of one another.

First, synchronized caption is automatically generated; word-level synchronization of text with speech is realized by ASR. The word-level alignment, which synchs each word with the speaker’s utterance, presents the phonological visualization of the words and thus leads to improvement in aural word recognition skills through mapping between the speech stream and the verbatim text.

Moreover, this method neatly presents word boundaries, which often cannot be easily recognized in authentic speech input. Synchronized captions, although in favor of many language learners, may bring too much assistance for the learner and makes them more and more dependent on the caption (Vandergrift, 2004; Garza, 1991). In order to solve the disadvantages of this method, we propose partial captioning which builds on synchronized captions to provide the students with reduced transcription of the videos in order to better train them for real-world situations.

This method can act as an intermediary stage before the learner is totally independent of captions. In this method, the filtering process of words to be presented takes into account not only the hindering factors of comprehension, but also the assessed knowledge of the learner. Hence, adjusted to a particular learner's need, the method selects words which are beyond the proficiency level of the learner. However, if using partial captions alone, as in keyword captioning, the students are often distracted by the sudden and irregular appearance of a word on the screen (Montero Perez et al., 2014). Nevertheless, this problem is mitigated by synchronization in PSC.

To conclude, this new tool, PSC, is anticipated to make the learner less dependent on caption and more prepared to handle listening in real-world situations. Table 1 summarizes the advantages of PSC compared to other captioning methods and Figure 1 shows the screenshot of a generated PSC.

Table 1: Comparison of different caption methods.

Caption Type / Advantage	Full Caption	Keyword Caption	Proposed Partial Caption	Synchronized Caption	PSC
Aid word boundary detection	✓			✓	✓
Speech-to-text mapping				✓	✓
Avoid over-reliance on reading		✓	✓		✓
Avoid being distractive	✓			✓	✓
Automatic	✓		✓	✓	✓
Adjustable to learners' knowledge			✓		✓
Adjustable to the content		✓	✓		✓

Figure 1. Screenshot of PSC on a TED talk. From the original transcript: *"how we motivate people how we apply our human resources"*.

3.1 Feature Selection

In order to decide which words to show in the caption and which ones to hide, the following features were picked as the selection criteria. These features were chosen for being identified as major contributing factors in listening comprehension impair. Besides, these factors can be quantified automatically and are easy to be implemented.

3.1.1 Speech Rate

Previous studies showed that high speech rate can negatively affect L2 listeners' comprehension (Dunkel, 1994) and this is even true for native speakers (Wingfield et al., 1985). For Japanese learners of English, particularly, fast rates of speech and inability to perceive the sounds in English are the major factors to impair comprehension (Osuka, 2007). Some studies suggested modification of speech rate as a solution, however, this is not close to real-world situation. Instead, we provide the learner with PSC that presents words/phrases uttered faster than normal rate of speech, or that of tolerable for the learner.

3.1.2 Word Frequency

When the vocabulary chosen by the speaker exceeds the vocabulary size of the listener, comprehension will be impeded. In such cases the unknown words confine the learner's attention, and as the speech proceeds the learner cannot pursue the subsequent parts. In other words, the listener invests a lot of time trying to understand what s/he missed (Goh, 2000).

The frequency of word usage in a language is a measure to assess word difficulty. For instance, learners are less likely to be familiar with low-frequency words (Nissan, 1996). Word frequency is calculated based on its occurrence in spoken or written corpora. A well-cited paper by Nation (2006) categorizes English vocabulary into High-frequency (the most frequent 2000-3000 word families), Mid-frequency (anything between 3000-9000 word families), and Low-frequency (beyond the 9000 frequency band). The term word family here refers to a base word and all its derived and inflected forms that can be recognized by a learner without having to learn each form separately.

To assist L2 listeners, PSC presents words or phrases, which are less frequent and hence make comprehension difficult.

3.1.3 *Word Specificity*

The occurrence of specific words in a video would make comprehension difficult since limited knowledge of academic words is often seen as a reason for L2 listening comprehension deficiency (Goh, 2000). Thus, when considering word frequency, it is important to consider word specificity as well. Using academic talks as the material for this study, this feature is also taken into account in PSC.

4. System Architecture

Figure 2 depicts the data flow and main components of the system. The procedure of generating a PSC starts with an alignment phase where the ASR system outputs the transcript with estimated word timing, which is aligned and adjusted with the given transcript of the caption. Next, word frequency, word specificity and speech rate are used to serve as the selection criteria for making PSC. The feature extraction module further processes the transcript and converts it into a feature vector for the decision making module.

A rule engine in the decision making module decides whether a word should be shown or not. This decision not only depends on the features, but also relies on the user input (i.e. quiz results).

In the formatting and display module, the captions are altered as the desired output of the system. Being synchronized with the utterance of the word, the corresponding dictation of the word (or character mask) should appear on the screen. Eventually this module plays back the media with the generated caption, and offers a pre-made comprehension test afterwards.

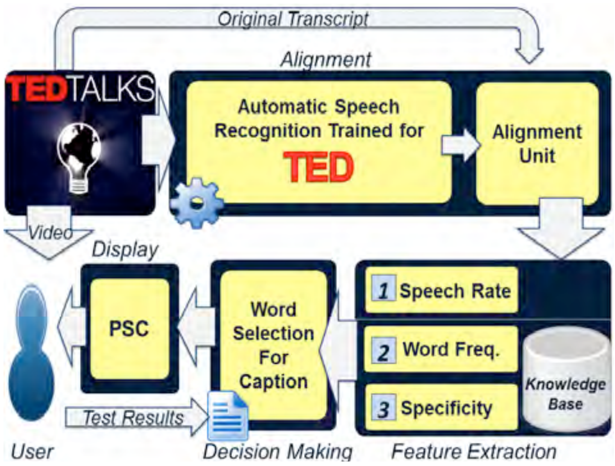


Figure 2. Data flow and the main components of the system.

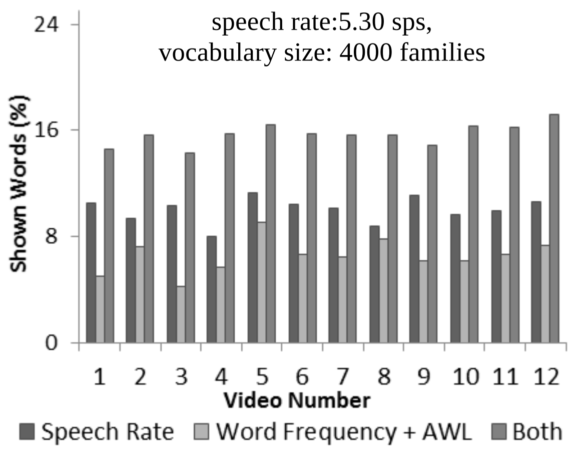


Figure 3. Percentage of words shown in PSC for intermediate learners

4.1 *Alignment Module*

The input data is composed of a video and its transcript text. To obtain the time tag of the tokenized words automatically, the audio should be ripped from the video to be passed to our ASR system, Julius v4.3.1 (Lee and Kawahara, 2009). Since Julius itself is a language-independent decoding program, it is possible to make a recognizer of a language, given an appropriate language model and acoustic model for the target language. The performance of ASR largely depends on these models. In this study TED talks were selected as the material. Thus, for precise alignment to take place, it is necessary to train the ASR models using a matched corpus, in this case TED talks. This model training was done in our laboratory, based on the lightly-supervised training approach using 780 TED talks (Naptali and Kawahara, 2012). The transcript and ASR output then got aligned using the force-alignment procedure.

4.2 Feature Extraction Module

This module extracts the main features and calculates them. The following elaborates on these features.

4.2.1 Speech Rate

The speech rate is often measured in Words per Minute (WPM) or Syllables per Second (SPS). The former may be affected by pauses and change of speech rate within a minute which causes inaccurate measurement while the latter is more suitable to measure short speeches and thus is used in this study.

The first step to calculate this feature is to estimate the speech rate where we need to count the number of syllables in each word, and then calculate the duration of its utterance. Calculation of the syllables is based on the structural syllabification of the corresponding text, which was realized using Natural Language Toolkit (NLTK). The full calculation of speech rate requires the duration of a word, which is calculated by the time tags obtained in the alignment phase after excluding the long pauses.

4.2.2 Word Frequency

Word frequency is defined by referring to corpus-based studies. Nation (2006) has designed 25 word family lists each including 1000 word families, plus four additional lists: (i) an ever-growing list of proper names; (ii) a list of marginal words including swear words and exclamations; (iii) a list of transparent compounds; and (iv) a list of abbreviations. The first two lists are carefully hand-selected while the rest are based on the following two famous corpora.

- The British National Corpus (BNC) which involves 100 million word collections of samples of written and spoken language from British English.
- Corpus of Contemporary American English (COCA), gathered by Mark Davies (from 1990 to 2012), includes 450+ million words. The corpus is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts.

This study is based on aforementioned word family lists and COCA. Every word is lemmatized first, and the result is looked up for the word family, created offline from the COCA and BNC corpus. The family of the lemmitized word serves as the difficulty index. The word is also cross-checked with the spoken genre section of COCA.

4.2.3 Word Specificity

In this method specific words are determined using a popular catalogue called Academic Word List (AWL) by Coxhead (2000) which includes 570 headwords and about 3000 of academic words altogether. Besides, these words are cross-referenced with COCA’s academic words (Gardner and Davies, 2013) for more accuracy. The system is also capable of handling other features such as abbreviations, proper names, numbers, transparent compounds, and repeated appearance of words.

4.3 Decision Making Module

Based on the features, the system decides whether a word should be included in the final partial caption or not. This decision not only relies on the value of the features, but also considers general features.

In the first stage, the main features - word frequency, speech rate, and specificity - are accounted. If only one of them require a word to be shown, the word is marked to appear in caption. To decide on the word frequency feature, a vocabulary size test (Nation and Beglar, 2007) is employed to assess the vocabulary size of the learner and to determine the appropriate frequency threshold for him/her. Similarly, a decision about whether a word should be a candidate for being shown in partial caption is taken by comparing the calculated speech rate of the word to that of preferable for the learner. Thus, if the utterance of the word (measured by speech rate feature) is faster than the tolerable threshold of the learner, the word will be shown in caption as a textual clue. This threshold can be adjusted by the user.

In the second stage, the general features act on each word. The features are either excitatory or inhibitory. The decision on general features is made on top of the first stage. For instance, abbreviations and proper names are being marked to be displayed while interjections are marked to be discarded.

The third stage of decision-making is about the sequence of the words that should be readable for the learners. The rules also handle words after numbers and words after “apostrophe s”.

4.4 Formatting and Display Module

This module generates the final partial and synchronized caption using the user display parameters. If the word is decided to be shown, it will be copied intact in the partial caption; otherwise a character mask (here we use “dots”) replaces every letter of the word. This will emulate the speech flow, by showing each and every word in the given speech in synch with their utterance. (e.g. “express” will be replaced by “.......” and “don’t” will be replaced by “....”).

5. Experiment

Given the novelty of partial and synchronized captioning method, experiments were needed to evaluate the effectiveness of this technique. Thus, the study investigates the following questions:

- Do captioned videos result in better comprehension of video compared to non-captioned ones?
- Can the proposed captioning method substitute the conventional full-text captioning?
- Do proficiency differences affect the benefits obtained from the proposed captioning method?
- Does the proposed method help the learner comprehend the video better without any captions?

5.1 Participants

The participants were 28 and 30 Japanese students at Kyoto University ranging from 19 to 22 years old. These students were undergraduates of different majors who enrolled at a CALL course. The experiments were carried out over this course, in two different classes, for three consecutive sessions.

5.2 Material

Videos: The video materials of this research were selected from TED website which provides us with authentic videos plus almost accurate captions without the copyright issue (www.TED.com). The selection criteria were bound to “popularity” and “recentness” of the videos. The selection was carefully done to include only videos of native American speakers, to avoid the influence of other accents. All videos were trimmed to 5-mintue meaningful segments.

Pre-study Vocabulary Size Test: A vocabulary size test created by Nation (2007) was used to evaluate the vocabulary reservoir of each student. The results of this test were used both as a placement criteria of dividing students into groups of proficiency and as a value to determine the frequency threshold for our caption generator. This test consists of 140 multiple-choice questions, with 10 items from each 1000 word family level. Since the caption generator uses the same word family lists as its references, the result of the test is appropriate to be set as our threshold.

Partial and Synchronized Caption Statistics: Taking into account the result of the vocabulary size test and the tolerable rate of speech, the system generates appropriate captions for learners with different proficiency levels. The percentage of words to be shown in the final caption does not exceed 30% for any of the videos as illustrated in Figure 3. This figure presents how the generated captions show fairly equal amount of words per video for a particular intermediate learner.

Comprehension Tests: After watching each video with assigned caption, the students were asked to take a listening comprehension test in the form of multiple choice and cloze test on summary.

5.3 *Procedure*

The study was conducted in CALL classes where students were provided with a 20 inch-wide screen and a headphone. Although the experiment was held in two different classes, the same procedure was adopted for both. Same videos were captioned with a different method (PSC↔FC) for each class.

We considered learner's proficiency as a blocking factor, with three levels: "beginner", "pre-intermediate" and "intermediate". For the purpose of dividing the students into these three groups, the assessed vocabulary size together with the students' TOEIC/CASEC scores were considered.

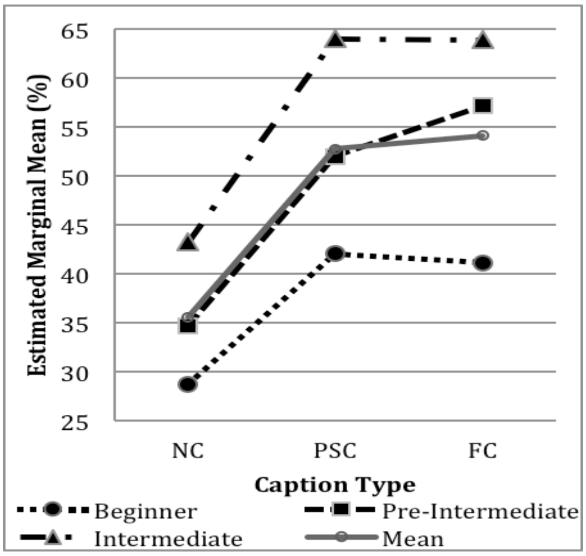
Each video, regardless of the caption type assigned to that was divided into two segments; 70% from the beginning and the rest of 30%. The students watched the first part of the video (70%) under one of these three conditions: no-caption (NC), full-caption (FC) and partial and synchronized caption (PSC). This was followed by a listening comprehension test. Next, the subjects were asked to watch the rest of the same video (30%) "without any caption" (regardless of the type of caption in the previous phase), and took another test. The procedure remained the same for all videos, while the type of caption was changed. To be more specific, the second part of each video is dedicated to evaluate students' performance on a non-captioned video as in real-world condition.

6. Results and Discussion

The scoring system was easily constructed because of the objective format of multiple-choice and cloze-on summary items. One point was awarded for each correct answer to multiple-choice questions while partial credit (0.25) was given to each item in cloze test. The total score was finally calculated in percentage for all participants in each group. One-way ANOVA test was used to analyze the result of the tests and to investigate whether any statistically significant difference is found between different conditions under which the learners watched the videos.

Table 2: Comprehension performance of both classes on the first part of video with (NC, PSC, FC)

Proficiency Level		Mean	SD	N
NC	Beg.	28.7	13.6	19
	Pre. Int.	34.7	11.8	19
	Int.	43.3	15.1	20
	Total	35.7	14.7	58
PSC	Beg.	42.0	16.7	19
	Pre. Int.	52.0	17.5	19
	Int.	64.0	18.0	20
	Total	52.9	19.4	58
FC	Beg.	41.1	12.3	19
	Pre. Int.	57.2	14.8	19
	Int.	63.9	16.4	20
	Total	54.2	17.3	58



* **NC:** No Caption **PSC:** Partial and Synchronized Caption **FC:** Full Caption

As shown in Table 2, analysis of the first part of the experiment (watching 70% of the videos) revealed a significant difference between NC (M =35.7, SD =14.7) condition and PSC (M =52.9, SD =19.4) and also FC condition (M =54.2, SD = 17.3) at the p < .05. The results provide a positive answer to our first research questions, which concerns the effectiveness of PSC as compared to NC. However, no significant difference was found between the scores gained under PSC and FC conditions in this part of the experiment [F (1, 57) = 25, p =.62]. The findings suggest that PSC leads to the same level of comprehension as FC while providing less than 30% of the transcript.

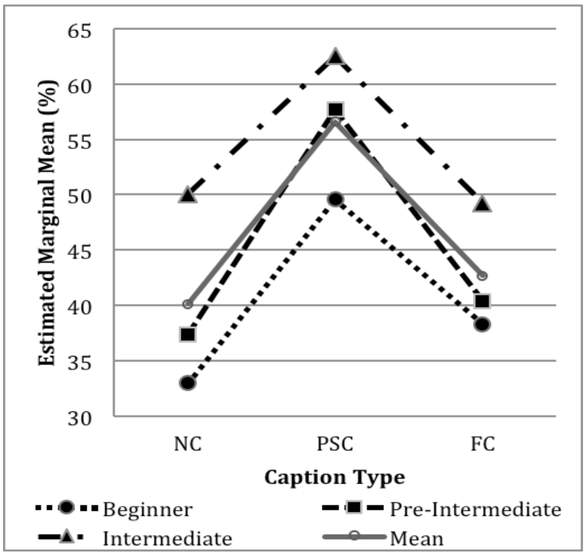
Consequently, the assumption of our second research question is plausible and hence PSC can be used as an alternative to full captioning method for training L2 listening. Furthermore, the results reveal that students with different proficiency levels gained almost equal scores under PSC and FC conditions and could benefit from our method. A possible explanation for deriving such results may lie in the adaptability of PSC that considers the proficiency level of the learners for generating appropriate amount of caption for them and provides adequate assistance for any learner.

Table 3 presents the results of comprehension tests on the second part of the experiment where students watched the rest of videos without any captions immediately after having watched the first parts under different conditions (NC, FC, PSC).

In the second part of the experiment (30% without caption), the best performance is associated with the condition in which the learners first watched the video with PSC [F (2,118) = 20.5, p < .05] as compared to FC and NC. The findings highlight the effectiveness of PSC on preparing the learner for real-world situation where captioning is not available. While this result indicate a short-term enhancement partly because of adaptation to the video, this finding is still of value.

Table 3: Comprehension performance of both classes on the second part of video without caption

Proficiency Level		Mean	SD	N
NC	Beg.	33.0	16.0	19
	Pre. Int.	37.4	16.6	19
	Int.	50.0	15.6	20
	Total	40.1	17.4	58
PSC	Beg.	49.6	15.8	19
	Pre. Int.	57.7	17.2	19
	Int.	62.5	17.4	20
	Total	56.6	17.3	58
FC	Beg.	38.3	13.5	19
	Pre. Int.	40.4	11.9	19
	Int.	49.3	12.7	20
	Total	42.7	13.4	58



*** NC:** No Caption **PSC:** Partial and Synchronized Caption **FC:** Full Caption

7. Conclusion and Future Work

The study introduced a novel technique of captioning, partial and synchronized, which is based upon speech rate, word frequency and specificity, to generate a smart type of caption that deals with limitation of previous methods. This method is based on the premise that the presence of infrequent or specific words and fast delivery of speech by the speaker hinder learner's listening comprehension. Additionally, by synchronization, the system emulates the speech flow which facilitates text-to-speech mapping and avoids the salient appearance of the words on the screen. Besides, to generate a suitable caption for a particular learner, the system assesses the tolerable rate of speech and vocabulary size of the leaner and prepares the captions in accordance to his/her level of competence.

Evaluated in two CALL classes, the results of the experiment showed that students' scores using PSC overtook that of the no-caption condition while resulted in almost equal comprehension as the full-caption condition. Furthermore, learner's scores on a new segment of the video without caption was significantly higher than other conditions when they watched the video with PSC first. The finding highlights the positive effect of PSC in preparing learners for listening in real-world situations.

The results also indicate that our method can assist learners to obtain adequate comprehension of the video by presenting less than 30% of the transcript to them. Such a method is assumed to be effective particularly for Japanese students who heavily rely on caption text in order to comprehend the content of the video. The findings further suggest that this form of captioning can be effectively incorporated into CALL systems as an alternative method to enhance L2 listening comprehension.

Long-term study requires both time and dedicated resources such as CALL classes that in this stage of the study was infeasible. Instead, we considered the immediate effect of the proposed method presuming a real-world situation by checking the learner's comprehension of a new segment of the video without any caption after being exposed to our proposed method. Although the findings has shown comprehension improvement on a short-time adaptation experiment, given the nature of listening skill, overall improvement could not be realized unless the participants undertake long-term experiments, hence such an experiment is suggested.

References

Bailly, G., & Barbour, W. S. (2011). Synchronous reading: learning French orthography by audiovisual training. *Proceedings of Interspeech 2011*, 1153-1156.

Bean, R. M., & Wilson, R. M. (1989). Using closed captioned television to teach reading to adults. *Literacy Research and Instruction*, *28*(4), 27-37.

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, *23*(04), 509-533.

Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, *34*(2), 213-238.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Journal des traducteursMeta:/Translators' Journal*, *49*(1), 67-77.

Dunkel, P. A., & Davis, J. N. (1994). The effects of rhetorical signaling cues on the recall of English lecture information by speakers of English as a native or second language. *Academic listening: Research perspectives*, 55-74.

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. The handbook of second language acquisition, 63-103.

Gardner, D., & Davies, M. (2013). A new academic vocabulary list. *Applied Linguistics*, amt015.

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals*, *24*(3), 239-258.

Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55-75.

Griffin, R., & Dumestre, J. (1992). An initial evaluation of the use of captioned television to improve the vocabulary and reading comprehension of navy sailors. *Journal of Ed. Tech. Systems*, *21*(3), 193-206.

Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal*, *15*(1-3), 89-108.

Lee, A., & Kawahara, T. (2009). Recent development of open-source speech recognition engine julius. In *Proceedings: APSIPA ASC* 2009, 131-137.

Leveridge, A. N., & Yang, J. C. (2013). Testing learner reliance on caption supports in second language listening comprehension multimedia environments. *ReCALL*, *25*(02), 199-214.

Markham, P. (1999). Captioned Videotapes and Second-Language Listening Word Recognition. *Foreign Language Annals*, *32*(3), 321-328.

Montero Perez, M., Peters, E., & Desmet, P. (2014). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL*, *26*(01), 21-43.

Munteanu, C., Penn, G., & Baecker, R. (2007). Web-based language modelling for automatic lecture transcription. In *Interspeech*, 2353-2356.

Naptali, W. and Kawahara, T. (2012): Automatic Transcription of TED Talk, *Proceedings IWSLT 2012*.

Nation, I. S. (2006). How large a vocabulary is needed for reading and listening?. *Canadian Modern Language Review/La revue canadienne des langues vivantes*, 63(1), 59-82.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.

Nissan, S. (1996). An Analysis of Factors Affecting the Difficulty of Dialogue Items in TOEFL Listening Comprehension. TOEFL Research Reports, 51.

Nitta, H., Okazaki, H., & Klinger, W. (2011). Speech Rates and a Word Recognition Ratio for Listening Comprehension of Movies. *Bulletin of English Movie Education Society*, (16), 5-16.

Osuka, N. (2007). What factors affect Japanese EFL learners' listening comprehension, *JALT2007 Challenging Assumptions*, *40*(5), 337-344.

Shimogori, N., Ikeda, T., & Tsuboi, S. (2010). Automatically generated captions: will they help non-native speakers communicate in english?. In *Proc. of the 3rd Int'l conf. on Intercultural collaboration*, 79-86.

Suzuki, H. (1996). The effects of closed-captioned videos on the listening and viewing processes of EFL learners. *Studies in Language*, 19, 19-34.
Vandergrift, L. (2004). Listening to Learn or Learning to Listen? *Annual Rev. of Applied Linguistics*, *24*(1), 3-25.
Vanderplank, R. (1993). A very verbal medium: Language learning through closed captions. *TESOL*, *3*(1), 10-14.
Wingfield, A., Poon, L. W., Lombardi, L., & Lowe, D. (1985). Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time. *Journal of Gerontology*, *40*(5), 579-585.
Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*(1), 65-86.
Winke, P., Gass, S., & Sydorenko, T. (2013). Factors Influencing the Use of Captions by Foreign Language Learners: An Eye-Tracking Study. *The Modern Language Journal*, *97*(1), 254-275.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Challenges in the annotation of article errors in Spanish learner texts

María del Pilar VALVERDE IBAÑEZ
[a]*Department of European Studies, Aichi Prefectural University, Japan*
valverde.mp@gmail.com

Abstract: Annotating learner texts with article error information is a difficult task. To identify which are the main difficulties for annotators, we carry out an annotation experiment in Spanish texts written by Japanese learners. Two expert and two non-experts raters annotate 300 noun phrases containing a definite, indefinite or zero article. We calculate inter-annotator agreement and analyse the sources of disagreement. We find article usage governed by pragmatic factors causes disagreement the most, while lexico-semantic factors are the most reliable. Finally, a learner corpus sample of 30,000 words is annotated with the revised annotation scheme.

Keywords: articles, error annotation, learner Spanish, inter-annotator agreement, reliability, corpus annotation

1. Introduction

The annotation of learner texts with error information is necessary for linguistic research as well as for the development of educational applications for language learning.While research has focused on the development of learner corpora and tools for English as a foreign language, the field of Grammatical Error Detection (GEC) is expanding and there is a need to develop resources for other languages. However, annotation and evaluation best practices are still an open issue.

First, inter-rater reliability for error annotation can vary widely: while for some errors, like number and gender agreement, rules are clearly defined, and using one rater may be acceptable, other kind of errors like article or preposition presence and choice are harder to annotate (Tetreault et al. 2010), so using only one annotator is not enough reliable. For article and noun number selection, for example, in Lee et al. (2009) raters found more than one valid construction for more than 18% of noun phrases. For prepositions, Tetreault and Chodorow (2008) found that even in native texts, “native raters can disagree with each other by 25% in the task of preposition selection”. In spite of this, learner corpora are typically annotated only once because double annotation would be too expensive, and few annotation projects provide measures of inter-annotator agreement (Rozovskaya and Roth, 2010; Lee et al., 2012). The need to improve annotation quality has been put forward by the NLP community, that has found difficulties for evaluating error detection systems in the last GEC shared tasks: in the first Helping Our Own task (2011), systems were penalized for valid corrections not annotated in the data while in last three tasks (H00 2012, CoNLL 2013 and 2014) teams could request the organizers to make changes in the annotation (Tetreault et al. 2014).

Second, as noted by Reidsma and Carletta (2008), there are different types of disagreement: chance disagreement, caused by random slips or lack of knowledge of the annotators, and systematic disagreement, due to different intuitions of the annotators or to a misinterpretation of the annotation guidelines. This distinction is crucial for the development of gold standards, since systematic disagreement can have a worse effect on machine learning than noise-like disagreement.

In this scenario, how can we improve the quality of annotations? Multiple annotations by more than one trained annotator is unrealistic for large projects and while crowdsourcing has shown good results for English preposition error annotation in a pilot study (Tetreault et al. 2010), more research is needed to deal with other languages and error types. To date, there has not been found a good method to improve the quality and number of annotations for learner texts

The goal of this paper is to investigate the main difficulties faced in the annotation of article errors in Spanish learner texts, so that measures can be taken to improve the quality of future annotation efforts. To do that, we carry out an experiment on article error annotation with a preliminary annotation scheme. In section 2 we describe the annotation principles, in 3 we briefly describe the data collection and annotation procedure, and in section 4 we investigate the sources of disagreement and main difficulties. In section 5 we apply the revised annotation scheme to a learner corpus sample and in 5 we present the conclusions.

2. Annotation principles

2.1. Article errors

In Spanish, articles can be definite (as in English *the*) or indefinite (in English *a/an*), and their form changes according to the gender and number of the noun they complement, as shown in Table 1 (base form in bold face).

Table 1: Spanish articles (gender and number)

	Definite		Indefinite	
	Masculine	Feminine	Masculine	Feminine
Singular	***el***	*la*	***un***	*una*
Plural	*los*	*los*	*unos*	*unas*

The definite article *el* 'the' is the most frequent word in Spanish and article usage is also one of the most frequent grammatical errors[1] among learners (specially for speakers of languages that do not have articles like Chinese, Japanese, Korean or Russian.), because article choice it is the result of interacting pragmatic, semantic, syntactic and lexical constraints.

We consider the following type of errors: missing article, extraneous article and article confusion. We are only concerned with article presence and choice, so we did not tag malformation (e.g. spelling or agreement errors) or order errors. A missing article occurs when the learner does not use any article but the noun phrase should contain one. An extraneous article occurs when the article used by the learner is not necessary (zero article is correct). A confusion error occurs when the learner used a definite instead of an indefinite, or vice versa, or when the learner uses a different type of determiner instead of an article.

2.2. Level of confidence in the judgments

It was expected that the annotators would sometimes be unsure about the acceptability of article usage in a given sentence, or unable to determine the most likely correction. With regard to the level of confidence in the annotators' judgments, annotated corpora do not explicitly provide confidence levels for every annotated item. Only in some annotation experiments the annotators are asked to indicate their level of confidence (as "low" or "high") (Tetreault and Chodorow, 2008).

We did not want to force the annotators to make a best guess in "difficult" sentences because that would lower inter-annotator agreement. Instead, we gave the possibility of marking such sentences as "difficult to judge" (as in Han et al. 2006), so later we could look at the sentences marked as problematic, and analyse what they have in common.

[1] . Fernández (1997) found 2.2 article errors per 100 words in a 4,433 words sample.

2.3. Number of tags

With regard to the number of possible analysis a sentence can receive, error-annotated learner corpora typically contain only one tag per error. However, the ”single correct construction” approach has been questioned and in recent annotation efforts there is a tendency to allow the inclusion of several alternative codes for the same item (Lüdeling et al., 2005; Boyd, 2010; Lee et al., 2012; Rozovskaya and Roth, 2010). However, it is unattainable to list all possible interpretations for every error, so this is done only when the error analysis is doubtful. In our experiment, we decided to allow only one tag per item to detect the sources of disagreement. After the revision of the annotation scheme, double tagging would be allowed in some specific cases (4.3.2).

3. Experiment

We carried out an experiment on article error annotation with the following objectives. First, calculate inter-annotator agreement for this task, which can be considered as the limit for an automatic article error detection system. Second, analyse the types and sources of disagreement, to find out which are the main difficulties the annotators face when annotating article errors in learner texts, so that measures can be taken to refine the annotation scheme and future annotation can be improved.

3.1. Data collection

A teacher of Spanish as a Foreign Language extracted sentences containing at least one article error from students' written assignments,[2] 50 sentences for each kind of article (definite, indefinite and zero article). The same number of sentences, but with at least one correct article usage, was then collected from the same texts. The distribution of the data is as Table 2 shows. In every sentence only one highlighted noun phrase had to be annotated, so the number of sentences and the number of annotated noun phrases is the same.

Table 2: Number of noun phrases and article they contain.

	Definite	Indefinite	0 article	Total
Correct	50	50	50	150
Incorrect	50	50	50	150
Total	100	100	100	300

3.2. Annotation procedure

The 300 noun phrases were tagged by 4 annotators. The annotators were two experts (teachers of Span- is as a Foreign Language, who correct learners' texts on a regular basis), which we will call E1 and E2, and two non-experts (native speakers of Spanish with higher education but without experience in corpus annotation), which we will call NE1 and NE2.

They all annotated the same noun phrase in the same sentences, but presented in different orders, using a Microsoft Excel spreadsheet. Annotators were provided with the target sentence plus the preceding and the following sentence, which they could resort to if they needed more context. They were asked to classify article usage for every noun phrase using one of the following tags: missing definite (AD), missing indefinite (AI), extraneous article (E), confusion error (C), difficult to judge (NC), article is correct (OK). They were not given any more guideline or training about the expected level of intervention in the texts: they were only asked to classify the noun phrases in one of the categories.

[2] The texts were written by 4th grade Japanese students of Spanish with an intermediate level of proficiency, at Aichi Prefectural University.

4. Inter-annotator agreement

Tables 3 and 4 show the confusion matrices for expert and non-expert annotations. Observed agreement, defined as the proportion of items on which annotators agree, is 0.79 for expert annotators and 0.76 for non- experts.

Table 3: Confusion matrix for Expert 1 and Expert 2 annotators

E1↓ E2→	AD	AI	C	E	NC	OK	Total
AD	**37**	0	0	0	2	2	41
AI	0	**5**	0	0	2	0	7
C	0	0	**30**	3	2	1	36
E	0	0	3	**39**	**7**	1	50
NC	1	0	1	4	5	**8**	19
OK	4	0	4	7	10	122	**147**
Total	42	5	38	53	28	134	300

Table 4: Confusion matrix for Non-expert 1 and Non-expert 2 annotators

NE1↓ NE2→	AD	AI	C	E	NC	OK	Total
AD	**31**	2	0	1	0	10	44
AI	2	**5**	0	0	0	2	9
C	1	0	**23**	2	2	6	34
E	0	0	4	**57**	2	10	73
NC	0	0	0	1	**0**	0	1
OK	5	1	5	7	2	**119**	139
Total	39	8	32	68	6	147	300

However, using observed agreement to measure reliability does not take into account agreement that is due to chance and hence is not a good measure of reliability. Therefore, an analysis using Cohen's Kappa statistic (Cohen, 1960) was performed. Perfect agreement would equate to a kappa of 1, and chance agreement would equate to 0. For the whole set of noun phrases (300, correct or incorrect), inter- annotator agreement for experts was found to be Kappa = 0.71 ($p < 0.001$), 95% CI (0.65, 0.77), and for non-experts it was 0.68 ($p < 0.001$), 95% CI (0.62, 0.75). If we exclude 45 sentences marked as "difficult to judge" by at least one annotator, kappa is 0.85 and 0.73 respectively. If we exclude 97 sentences tagged as correct by the four of them kappa is 0.62 and 0.58. If we exclude both sentences marked as NC by at least one annotator and sentences marked as OK by four annotators (remaining only 159 noun phrases all of them containing "safe" article errors) kappa is 0.79 and 0.61. Although kappa values vary depending on the set of sentences used to calculate it, agreement is over 0.60, which indicates "substantial agreement".

These figures are slightly lower than those for English. In Han et al. (2006) annotators classify noun phrases in the same categories as our experiment with a kappa of 0.86, excluding correct noun phrases and sentences where they are unable to determine correct usage, which for us was 0.79 and 0.61 for experts and non-experts. The difference in the kappa values can in part be explained by the different proportion of article types in the data: while in our experiment article types are balanced (one third of noun phrases for every article type), in real texts like those used in Han et al. (2006) the zero category is the most common (followed by the definite and indefinite article) and this category also has the highest inter-annotator agreement[3], which may raise the total kappa value.

In the following sections we examine different types of disagreement: disagreement due to the annotators' individual biases (4.1), due to the annotation scheme (4.2) and genuine disagreement (4.3).

[3] 3Full agreement (that is, by the four annotators) in sentences with an indefinite article is lower (45%) than in sentences with the zero article (71.0%) $\chi^2(4, N = 299) = 16.7$, $p = 0.02$.

4.1. Disagreement due to the annotators' individual biases

As expected, non-expert annotators are less reliable than experts. First, non-expert annotators make more mistakes (they add tags which are incompatible with certain noun phrases, e.g. a missing article tag in a noun phrase already containing an article). To avoid this kind of mistakes, we should constrain the available tags depending on the input (e.g. if there is already an article in the noun phrase, do not allow the "missing" tag). Second, even though non-experts are supposed to be less confident on their annotation because pointing out errors in a text is a task for which they have no previous training, in fact they are less cautious than experts when they correct texts. This bias explains why, for example, NE1 uses the tag "difficult to judge" only one time (0.3%), while E2 uses it almost once every 10 sentences (9.3%), and non-experts use the tag "extraneous article" (specially for definite articles) more frequently than experts (23.5% vs 12.2% of times).

Part of the variability in annotators' attitude could be reduced by giving clear guidelines about the optimum level of intervention in the texts. In this regard, we advocate for following a principle of minimal change: so we should not mark as errors the sentences where the learner choice is acceptable, even if the learner choice is not the best choice, that is, the goal of the annotator should be to produce an acceptable rather than a perfect result.

In relation to that, annotators should be instructed about the halo effect, by which the judgment of a sentence as acceptable or unacceptable is influenced by our overall impression of previous sentences. In other words, one is more likely to find errors in a text if this text already contains other errors. While expert annotators (teachers of a foreign language) are trained in evaluation methods and therefore they are aware of the importance of reliability in students' evaluation, know how external factors (e.g. the halo effect and contrast effect) can have a negative impact and what can be done to reduce it, non-experts lack this training and do not know how to perform a fair evaluation -annotation. Therefore, non-expert annotators should receive training in evaluation methods to be able to reliably correct learner texts.

4.2 Disagreement due to the annotation scheme

With regard to the reliability of the 6 tags used for annotation, "difficult to judge" is the one that causes more disagreement: most of the times (67.7%) it is used by only one of the four annotators, and it is never used by three or four annotators in the same sentence. On the contrary, the rest of tags have a much higher agreement: on average, they are used by the four annotators 63.2% of the times, by three 19.9%, by two 9.2% and by one 7.7% of times. Therefore, this tag should at most be used to filter out problematic sentences, which annotators cannot comprehend, and not for proper annotation of sentences.

We advocate for not using this tag and instead set clear principles in the annotation guidelines specifying what the annotators should do when they are not confident about the error analysis of a sentence: exclude the sentence if a reasoned annotation is considered impossible (e.g. in incomprehensible fragments of text), or use more than one error tag if both are equally possible (as we will see in 4.3.2).

4.3 Genuine disagreement

Article presence and choice can be determined by different types of factors: it mainly depends on pragmatic factors (in our data, 69.0% of noun phrases), lexico- semantic factors (20.7%) and syntactic factors (10.3%).

As for pragmatic factors, for example the definite article is used to generalize, that is, to refer to a whole class of things or people, as in (1) (we underline the noun phrase and indicate the type of article in brackets) and to refer to something that is identifiable to the listener, as in (2).[4]The indefinite is used to refer to any object of a particular class, as in (3), and no article is used when we are talking about an indefinite amount of something, as in (4) (examples from Alonso et al. (2013)).

(1) Los hijos dan muchos disgustos. [DEFINITE]
'Children cause a great deal of trouble.'

(2) El hijo de María tiene dos años. [DEFINITE]
'Maria's son is two years old.'

(3) Tener un hijo es lo mejor que te puede pasar en esta vida. [INDEFINITE]
'Having a child is the best thing that can happen in life.'

(4) No tengo hijos pero tengo sobrinos. [NO ARTICLE]
'I do not have children but I have nephews.'

As for lexico-semantic factors, for example, place names usually have no article (*México*), while the definite is obligatory for rivers, mountains, seas and oceans (*el Mediterráneo*), and there exist many set phrases and idioms which require definite (e.g. *con el objetivo de* 'with the objective of'), indefinite (*por una parte*, 'on the one hand') or zero article (e.g. *a corto plazo*, 'in the short run'). As for syntactic factors, for example two or more nouns should have their own article if they refer to different things: *un gato y un perro*, "a cat and dog" (*un gato y perro* suggests a cross between a cat and a dog) (Butt and Benjamin, 2014).

Leaving aside sentences tagged as "correct" by 4 annotators, agreement is higher when the article choice depends on lexico-semantic factors (k = 0.835 for experts and 0.780 for non-experts) and lower with pragmatic factors ((k = 0.514 for experts and 0.496 for non-experts). Syntactic factors seem to be in between (k = 0.750 for experts and 0.523 for non-experts), although their low frequency makes the figures less reliable. Therefore, more care should be paid to pragmatic distinctions.

Specifically, disagreement is more likely in noun phrases where two pragmatic interpretations (and article choices) are possible, and annotators choose one of the alternatives in an inconsistent manner (§ 4.3.1 and § 4.3.2). Disagreement can also be due to a lack of the world knowledge that is needed to be able to determine the correct article usage (§ 4.3.3). As for syntactic and lexico-semantic factors (§ 4.3.4), disagreement occurs because annotators do not have a good knowledge about the existing prescriptive rules about article usage.

4.3.1. Vacillation between definite article and zero article

Frequently both the definite and zero article are acceptable for the same noun phrase. This happens when the noun phrase can refer to a whole class of things or people in general (definite article) as in (1) or to an indefinite amount of something (zero article) as in (4). This distinction frequently does not change the meaning of the sentence significantly and in fact some languages with articles like English usually use the zero article to express both situations.

In our experiment, when the two pragmatic interpretations are possible for a given sentence, annotators inconsistently choose one of them: some annotators tag the noun phrase for a missing

[4] In (2) Maria's son must be identifiable for the listener because a) Maria has only one son, or b) we have talked about him before.

article in (5) (OK|AD|AD|OK)[5] while they tag it for an extraneous article in (6) (E |NC|OK|E), although both noun phrases can have the same pragmatic interpretations.

(5) Los políticos hablan en público y manifiestan sus opiniones con el objeto de conseguir votos de ciudadanos [...] [NO ARTICLE]

'Politicians talk in public and show their opinion with a view to get votes from the citizens [...].'

(6) Concretamente los cursos que consiguieron participantes japoneses y que ofrecen los certificados oficiales como IMEC(Instituto de Medicina China) continuarán existiendo [...]. [DEFINITE]

'Specifically the courses which obtained Japanese participants and offer official certificates like IMEC (Chinese Medicine Institute) will continue existing [...].'

In these cases, when both the definite and the zero article are acceptable, according to the principle of minimal change, we opt in favor of leaving the learners' choice unchanged if it is acceptable.

4.3.2 Vacillation between indefinite article and zero article

Sometimes annotators agree in considering a noun phrase as unacceptable but they do not agree in the type of correction. This can happen when the learner wrongly uses a definite article, as in (8) (E|C|C|E), and the annotators propose different corrections: it can be an extraneous article if the noun phrase refers to an indefinite amount of something (zero article), or a confusion error if the noun phrase refers to any object of a particular class (indefinite).

(8) En cambio, la cocaína tiene el efecto tóxico. [DEFINITE]
'On the contrary, cocaine has a toxic effect.'

When the two are equally acceptable and the annotator considers she cannot make a reasoned choice, we consider the best solution is to allow two error tags (E/C). In our experiment, this only happens with the pair of tags E and C.

4.3.3 Lack of world knowledge

In some sentences, annotators have insufficient extra-linguistic knowledge to be able to determine the right article usage. For example, in (9) (OK|E|E|E) the annotator needs to know whether in Nagoya there are only nine interesting and touristy places (definite article) or there are more than nine (no article).

(9) Sale cada treinta minutos aproximadamente desde la estación de Nagoya y paran en los nueve sitios muy interesantes y turísticos, por ejemplo El castillo de Nagoya. [DEFINITE]

'It runs approximately every thirty minutes from Nagoya station and stops in nine very interesting and touristy places, for example Nagoya Castle.'

For future annotation, if the learner's choice is acceptable in some context, as in (9), we do not mark it as wrong. If the learner's choice is not acceptable, we tag the noun phrase as usual.

4.3.4 Lack of knowledge about syntactic and lexico-semantic rules

[5] For very example from the learner data, in parenthesis we indicate the tags chosen by the 4 annotators, in the following order: Expert 1, Expert 2, Non-expert 1, Non-expert 2. We also indicate in brackets the article choice of the learner.

Unlike article usage governed by pragmatic factors, which is subject to interpretation by the annotator, for article usage determined by syntactic and lexico-semantic constraints there exist some clear rules about what is considered correct and incorrect by the linguistic norm. These rules are part of language planning efforts by the Spanish language academy, but native speakers -even experts- do not have sufficient knowledge about them and as a result sometimes do not follow them when they annotate learner texts. For example, in (10) (AD|AD|OK|OK) experts marked as error an article usage that is actually accepted (RAE, 2006): the zero article between the preposition ***a*** ('to') and the relative pronoun *que* ('which').[6]

> (10) [...] el capítulo 2 dice sobre el proceso del portuñol y los problemas a que el portuñol se enfrenta actualmente. [NO ARTICLE]
> '[...] chapter 2 is about the "portuñol" process and the problems that the "portuñol" confronts nowadays.'

Therefore, to determine the acceptability of article usage, annotators should not rely only on their intuition as native speakers but they should also consult existing rules and recommendations published in reference dictionaries and grammars as RAE (2006) and RAE (2009) to avoid contradictions between their corrections and what the linguistic norm actually says.

4. Corpus annotation

After analysing the main sources of disagreement, we have revised the annotation scheme as explained in Valverde & Ohtani (2014). Then, we have applied the revised annotation scheme to the annotation of an approximately 30,000 words sample of the CEDEL2 learner corpus (Lozano & Mendikoetxea 2013), as shown in table 5. The texts in the sample were written without preparation, by learners whose first language is English.

Table 5. 30,000 words sample from the CEDEL2 learner corpus.

Level	Words	Texts/Learners
Beginner	10390	40
Intermediate	9960	22
Advanced	10293	20
Total	30643	82

The following categories were used: 1) Missing definite article, 2) Missing indefinite article, 3) Extraneous definite article, 4) Extraneous indefinite article, 5) Confusion error: indefinite instead of definite, 6) Confusion error: definite instead of indefinite, 7) Confusion error: another determiner instead of definite, 8) Confusion error: indefinite instead of another determiner

Annotation has been carried out by one trained annotator with the software UAM Corpus Tool (O'Donnell 2010).We have found 196 errors in 30643 words, that is 0.64/100 words. Results are shown in Table 6. As expected, the most frequent error type involves the presence/absence of article: 92 missing articles (as in 10) and 95 extraneous articles (as in 11) give account of 95.41% of errors, and confusion errors represent only 4.59% (as in 12). This proportion is very close to that found in English learner texts: Han et al. (2006) found 21.5% of extraneous articles, 58.6% of missing articles and only 6.2% of ***a-the*** confusion in English texts written by Japanese learners. However, the frequency of extraneous articles in our texts –very close to missing articles- is higher than in the English texts.

[6] The definite article is also acceptable but not *obligatory*. The definite article would be obligatory if the antecedent referred to a person, or if the subordinate clause was negative.

Table 6. Frequency of error tags by language level

Error tag	Beginner		Intermediate		Advanced	
	N	%	N	%	N	%
Missing article	51	40.80	25	64.10	16	50.0
Extraneous article	68	54.40	12	30.77	15	46.88
Confusion error	6	4.80	2	5.13	1	3.12
Total	125	100	39	100	32	100

(10) Jude Law tiene pelo rubio y es Ingles. [0 → DEFINITE]
'Jude Law has blond hair and is English'

(11) Fui en el junio y no llovó allí. [DEFINITE →0]
'I went in June and it did not rain there'

(12) Me encanta ir a la Universidad porque es la experiencia Buena.
[DEFINITE→INDEFINITE]
'I love going to the University because it is a good experience'

Among missing articles, the most frequent is the omission of the definite (88/92), as shown in Table 7. Among extraneous articles, the proportion of definite and indefinites is more balanced (37 vs 58).

Table 7. Frequency of error types by language level

Error type	Beginner		Intermediate		Advanced	
	N	%	N	%	N	%
Missing type	51	100	25	100	16	100
Missing definite	50	98.04	25	100.00	13	81.25
Missing indefinite	1	1.96	0	0.00	3	18.75
Extraneous type	68	100	12	100	15	100
Extraneous indefinite	30	44.12	2	16.67	5	33.33
Extraneous definite	38	55.88	10	83.33	10	66.67
Confusion type	6	100	2	100	1	100
Definite instead of indefinite	5	83.33	0	0.00	0	0.00
Indefinite instead of definite	0	0.00	0	0.00	0	0.00
Indef. instead of another det.	0	0.00	1	50.00	1	100.00
Another det. instead of definite	1	16.67	1	50.00	0	0.00

From this data we can extract some statistically significant differences among learners. For example, for beginners learners extraneous articles (54.40%) are more frequent than missing articles (40.80%), while for intermediate learners the opposite is true: missing articles (64.10%) are more frequent than extraneous ones (30.77%). As for the type of missing article, among beginner learners the indefinite is very rare (1.96%), but not so rare for advanced learners (18.75%).

5. Conclusions

Although article errors have been annotated in a number of small-scale studies, to date there has not been any study about article error annotation and inter-annotator agreement in Spanish learner texts. In this paper we have tested the results of an annotation scheme for article errors in a sample of learner texts written by Japanese learners. We have calculated agreement among 4 annotators (2 experts and 2 non-experts) and have found kappa values between 0.85 and 0.62 for expert annotators and from 0.73 to 0.58 for non-experts, depending on the collection of sentences considered.

Non-experts are less reliable than experts, and the annotation scheme (the tag “difficult to judge”) is also responsible for part of the disagreement.
As for genuine disagreement among annotators, some pragmatic distinctions are specially problematic: the distinction between a) a whole class of things or people in general (definite article) and b) an indefinite amount of something (zero article), and the distinction between a) an indefinite amount of something (zero article) and any object of a particular class (indefinite article). In addition to that, some times more world knowledge is needed to determine whether article presence and choice is acceptable or not. As for article usage governed by syntactic and lexico-semantic factors, annotators sometimes disagree in determining the right article usage because they lack knowledge about the existing prescriptive rules published by the Spanish language academy.

To improve annotation reliability, annotators need to be trained in language evaluation methods and have to consult published prescriptive rules about article usage. After annotating a 30,000 words sample from the CEDEL2 learner corpus with a revised annotation scheme, we have found that the most frequent error types are missing and extraneous articles.

Acknowledgements

This work was supported by kakenhi (25770207), Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science.

References

Alonso, R. Castañeda,A. Martínez, P. Miguel, L. Ortega, J. & Ruiz, J. (2013). Students’ Basic Grammar of Spanish. Difusión.

Boyd, A. (2010). Eagle: an error-annotated corpus of beginning learner german. *Proceedings of LREC-10*, Malta.

Butt, J. & Benjamin, C. (2014). A New Reference Grammar of Modern Spanish. Routledge.

Chodorow, M. Dickinson, M. Israel, R. & Tetreault, J. (2012). Problems in evaluating grammatical error detection systems. *Proceedings of COLING 2012*, pages 611–628, Mumbai, Desember.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.

Davies, M. (2005). A Frequency Dictionary of Spanish: Core Vocabulary for Learners (CD). Routledge.

Fernández, S. (1997). Interlanguage and Error Analysis in the Learning of Spanish as a Foreign Language. Edelsa, (In Spanish: Interlengua y análisis de errores en el aprendizaje del español como lengua extranjera).

Han, N., Chodorow, M. and Leacock, C. (2006), Detecting errors in English article usage by non-native speakers, Natural Language Engineering 12 (2): 115–129.

Lee, S. Dickinson,M. & Israel, R. (2012). Developing learner corpus annotation for korean particle errors. Association for Computational Linguistics (editor), *Proceedings of the Sixth Linguistic Annotation Workshop, LAW VI*, pages 129–133, Stroudsburg.

Lozano, C., & Mendikoetxea, A. (in press 2013). Learner corpora and second language acquisition: the design and collection of CEDEL2. In N. Ballier, A. Díaz-Negrillo, & P. Thompson (Eds.), *Automatic Treatment and Analysis of Learner Corpus Data*. Amsterdam: John Benjamins.

Lüdeling,A., Walter, M., Kroymann,E. & Adolphs, P. (2005). Multi-level error annotation in learner corpora. *Proceedings of the Corpus Linguistics 2005 Conference*, Birmingham, United Kingdom, July.

O’Donnell, M. (2010) UAM Corpus Tool. Available at http://www.wagsoft.com/CorpusTool/

Real Academia de la Lengua Española (RAE). (2006). Diccionario panhispánico de dudas. Real Santillana.

Real Academia de la Lengua Española (RAE). (2009). New Grammar of Spanish Language (In Spanish: Nueva gramática de la lengua española). Espasa Calpe.

Reidsma, D & Carletta, J. (2008). Reliability measurement without limits. *Computational Linguistics*, 34(3):319–326.

Rozovskaya, A. & Roth, D. (2010). Annotating ESL errors: Challenges and rewards. *Proceedings of NAACL’10 Workshop on Innovative Use of NLP for Building Educational Applications*. University of Illinois at Urbana–Champ.

Tetreault, J. & Chodorow, M. (2008). Native judgments of non-native usage: Experiments in preposition error detection. *Proceedings of the Workshop on Human Judgments in Computational Linguistics at the COLING 2008*, pages 24–32.

Tetreault, J, Filatova, E. & Chodorow, M. (2010), Rethinking Grammatical Error Annotation and Evaluation with the Amazon Mechanical Turk, Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications, pages 45–48, Los Angeles, California, June 2010.

Tetreault, J. & Leacock (2014), Automated Grammatical Error Correction for Language Learners, Turorial at *25th International Conference on Computational Linguistics (COLING 2014)*, Dublin, August 23-29.

Valverde, M.P & A. Ohtani (2014, forthcoming), Annotating article errors in Spanish learner texts: design and evaluation of an annotation scheme. *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing (PACLIC 2014).*

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Tools for Supporting Language Acquisition via Extensive Reading

Alexandra UITDENBOGERD*
[a]*School of Computer Science and IT, RMIT University, Australia*
*alexandra.uitdenbogerd@rmit.edu.au

Abstract: Extensive reading, that is, the reading of large quantities of text at a comfortable level of difficulty, has been shown to be of great benefit to second and foreign language learners skills. We define seven types of reading support as management, text generation, text selection, text simplification, preparation, translation, and revision. We describe and propose a range of tools that assist learners to locate reading material of an appropriate level of difficulty and manage their progress. Our prototype implementation of two of these tools, the Bilingual eReader and the Readable Extract Search Engine, demonstrate their feasibility.

Keywords: CALL, extensive reading

1. Introduction

Imagine attempting to make sense of the following extract while reading:

> She was the youngest of the two daughters of a most XXX, XXX father; and had, in XXX of her sister's XXX, been XXX of his house from a very early XXX.

The “XXX”s above represent unknown words. For this extract, you are reading a text in which you know 81.25% of the words. If you are a native or highly fluent speaker of English, you have the advantage of excellent knowledge of the grammatical flow of English sentences, which will greatly assist you in your ability to predict the meaning of the missing words. Research has shown that for most foreign language learners to predict the meaning of words in text requires knowing at least 95% of the words (Liu and Nation, 1985). However, for most native texts, the learner would need a vocabulary of 5,000 word families in order to achieve that coverage (Hirsh and Nation, 1992). A typical 1,000 hour course of English as a second language will achieve a 2,000 word vocabulary (Laufer, 2000). This leaves a substantial shortfall in vocabulary knowledge.

It is very challenging to become proficient in a foreign language to the extent that it can be used for discourse at a high level. One useful activity for improving language skills that is readily available is extensive reading. However, the reading material needs to be at an appropriate level of difficulty as well as interesting to be the most effective. For English and French there is an extensive collection of reading material available, such as the Oxford Bookworms series. Other languages are less resource rich. However, there are large quantities of text produced in a variety of languages on the Web and elsewhere. A proportion of these are potentially suitable for reading practice. Uitdenbogerd (2006) examined a corpus of English web sites and found that the readability range of Web-sites adequately covered that of typical stories written for English language learners. However, many such web-sites provided uninteresting reading, such as navigation pages. Heilman et al. (2010) reached a similar conclusion in their work, and developed methods to eliminate web pages that don’t have sufficient prose, as well as providing categories for students to choose from to increase the chance that the material retrieved would be interesting to them.

The contributions of this paper are:

1. a taxonomy of extensive reading support that can be provided by computer (and other) systems (See Section 3)
2. tools and techniques for providing appropriate text from a corpus of literature (See Section 4)

2. Background

The need for language skills is becoming more relevant in globally connected societies. For example, millions of Chinese students are studying English (Zheng and Cheng, 2008), and 27 million people have taken the TOEFL English language test (TOEFL website). It is generally agreed that more exposure to language improves skills in the language. For example, various studies have shown vocabulary gain from reading (Waring and Takaki, 2003).

In a previous paper we observed that there are short extracts and sentences to be found in the classical French literature that meet strict vocabulary criteria such as consisting only of the 20 most frequently occurring words in news text, French-English cognates, and proper nouns (Uitdenbogerd, 2010). We provided an estimate of exact cognates in native text (10%) and a frequency distribution of sentence structures.
Clearly the more strict the constraints on the text, the fewer suitable extracts will be found. However, even at the strictest constraints such as 1-word sentences, or sentences consisting only of the most frequent 20 words, proper nouns and French-English cognates, extracts could be found. Relaxing constraints to allow 95% coverage provides larger quantities of extracts.

3. Extensive Reading Support Taxonomy

Support for extensive reading falls into seven main categories: management of an extensive reading-based programme, generation of readable text, selection of readable text, simplification of text, learner preparation, glossary or translation support, and revision.

- ***Management*** via an extensive reading system frees the user from tracking their reading and their progress. It can work in conjunction with text selection and other categories of support by overseeing the different types of activity available to the user.
- ***Text Generation*** involves generating stories or other content based on strict readability constraints. For early stages that require much reading practice with a small set of vocabulary and grammar this can be a useful addition to the choices availble for reading.
- ***Text Selection*** via readability-based search enables the learner to find easier texts to read from large corpora. When combined with topic search, texts can be both relevant and readable. Text selection can also be based on recommendations in the manner of typical ratings-based recommender systems. Text selection support can help ensure that text is both readable and interesting, increasing the motivation of the learner to read and benefit from reading.
- ***Text Simplification*** has traditionally been done manually, typically for classic works of literature to make them accessible to both children and foreign language learners. Typically the works are not only simplified in language, but made significantly shorter.
- ***Preparation*** consists of materials that assist the learner with vocabulary and background knowledge before they commence reading. (See Section 4.1)
- ***Translation*** of difficult vocabulary occurring in text helps the learner to read fluently, and when a gloss is used to look up a word, the word is more remembered than if it is merely read over (Lomicka, 1998). Translations of difficult passages can also help a learner tackle more difficult texts.
- ***Revision*** typically consists of a set of questions that the learner answers once they have completed reading a text. The questions test the learner's comprehension of the text, as well as their knowledge of vocabulary and grammar seen in the text. The more involved the learner is with the language, the more they will remember. Therefore, exercises based on the reading will improve language retention (Laufer 2000).

Applications vary in their support across these categories. In Section 5 we present a range of applications and discuss them in relation to the categories of support and their applicability to language acquisition via extensive reading.

4. Applications

We discuss various existing and proposed applications that provide support of the types described in our taxonomy.

4.1 The Readable Document Search Engine

An idea that was independently developed by several research groups, and is now a part of major search engines is to make readability a criterion for search for documents. The most developed and used system that retrieves documents for foreign language reading practice is the REAP system (Heilman et al. 2010). It allows instructors to choose target vocabulary to be studied by students, permits students to select broad interest areas, and presents the students with recommended texts of an appropriate reading level that provide practice in the target vocabulary. REAP determines the student's existing vocabulary, so that it can better estimate the reading level required. While it doesn't appear to provide preparation activities, it does provide comprehensive translation support by making any word searchable in an on-line dictionary, and target words for study are hyperlinked to a definition. Revision consists of a practice exercises on target words. The system has been shown to improve language knowledge.

Where systems recommend native texts and also track a student's vocabulary knowledge, a personalised set of preparation material could be provided. For example, based on the word frequencies in the document and the known vocabulary, a set of 5-10 words that would improve the student's ability to understand the document the most, could be presented in a small pre-reading lesson. A simple but not necessarily optimal approach would be to select the highest frequency document words that are not in the student's current known vocabulary. These are likely to be "topic" words. For example, in a story about pirates, the words "pirate", "sail", "mast", "anchor" and "cabin" may occur frequently, but be generally unknown by a student studying the English Language. Where personalised vocabulary tracking doesn't exist, the preparation material can be selected purely by the frequency of words in the document versus their frequency in background text. Some published reading books for foreign language learners do this to some extent.

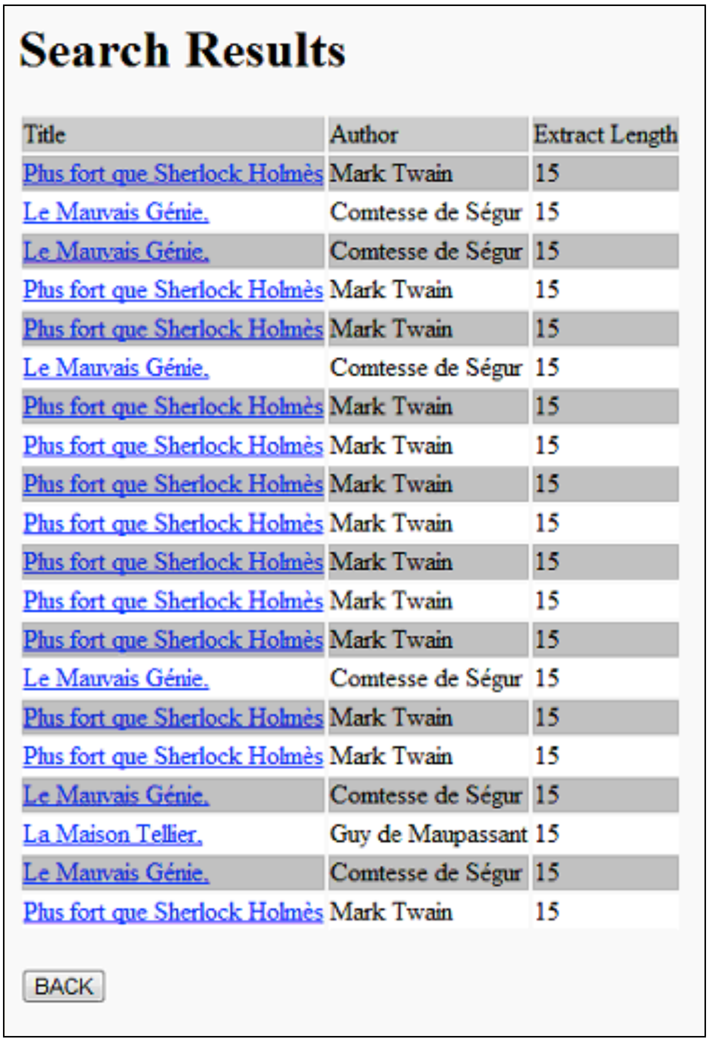
4.2 The Readable Extract Search Engine

The idea behind the readable extract search engine is that a collection of native texts that are of high quality and interest, such as a collection of literature, may provide short extracts for reading. In an initial exploration we found it was possible to locate short readable extracts based on both vocabulary and grammar constraints (Uitdenbogerd, 2010). In our prototype we used sentence length as the readability measure, which has been shown to work very well for the case of French as a foreign language for English speakers (Uitdenbogerd, 2006). However, a more sophisticated readability measure, or one that is appropriate for a different language can be substituted.

Users set the sentence length criteria and the ideal extract size (Figure 1), and then retrieve a list of results (Figure 2). Users can then choose which extract is of interest from a list of titles and extract lengths.

In our prototype implementation, the extract is shown highlighted within the full text, allowing users to continue reading if they so choose (Figure 3). While the usefulness of short extracts hasn't been demonstrated as yet, extracts that provide multiple examples of a particular word or phrase has increased vocabulary knowledge (Webb, 2007).

Figure 1. The Readable Extract Search Engine splash screen.



Search Results

Title	Author	Extract Length
Plus fort que Sherlock Holmès	Mark Twain	15
Le Mauvais Génie.	Comtesse de Ségur	15
Le Mauvais Génie.	Comtesse de Ségur	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Le Mauvais Génie.	Comtesse de Ségur	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Le Mauvais Génie.	Comtesse de Ségur	15
Plus fort que Sherlock Holmès	Mark Twain	15
Plus fort que Sherlock Holmès	Mark Twain	15
Le Mauvais Génie.	Comtesse de Ségur	15
La Maison Tellier.	Guy de Maupassant	15
Le Mauvais Génie.	Comtesse de Ségur	15
Plus fort que Sherlock Holmès	Mark Twain	15

BACK

Figure 2. Retrieved list of extracts.

```
Les armées françaises, maîtresses de toute la rive gauche du Rhin, et
prêtes à déboucher sur la rive droite, menaçaient la Hollande et
l'Allemagne: fallait-il les porter en avant ou les faire entrer dans
leurs cantonnemens? telle était la question qui s'offrait.

Malgré leurs triomphes, malgré leur séjour dans la riche Belgique, elles
étaient dans le plus grand dénuement.

Le pays qu'elles occupaient, foulé
pendant trois ans par d'innombrables légions, était entièrement épuisé.


Aux maux de la guerre s'étaient joints ceux de l'administration
française, qui avait introduit à sa suite les assignats, le _maximum_ et
les réquisitions. Des municipalités provisoires, huit administrations
intermédiaires, et une administration centrale établie à Bruxelles,
gouvernaient la contrée en attendant son sort définitif. Quatre-vingts
millions avaient été frappés sur le clergé, les abbayes, les nobles, les
corporations. Les assignats avaient été mis en circulation forcée; les
prix de Lille avaient servi à déterminer le _maximum_ dans toute la
Belgique. Les denrées, les marchandises utiles aux armées étaient
soumises à la réquisition. Ces règlemens n'avaient pas fait cesser la
disette. Les marchands, les fermiers cachaient tout ce qu'ils
possédaient; et tout manquait à l'officier comme au soldat.
```

Figure 3. Highlighted extract in main body of text.

4.3 The Bilingual e-Reader

The bilingual e-Reader combines simplification with translation, in that the more difficult sentences are presented in the learner's native language. The idea of mixed language reading material is not new. The approach has been used for gradually increasing the amount of target language in a story for English-German with good results compared to a normal German lesson (Weible, 1980), and has also been used to introduce Japanese kanji within an English story (Watanabe, 2002).

We developed a simple prototype using sentence length as the readability criterion, with a collection of movie subtitles (downloaded from http://opus.lingfil.uu.se/OpenSubtitles.php) as the corpus.

Figure 4 shows the main screen, including parameters for source and target language, a list of movie subtitles available, a readability parameter (maximum sentence length), and the resulting mixed language subtitles for the movie Rocky with the maximum sentence length set to 7.

The idea is a little controversial in the sense that it prevents total immersion in the target language, but it can be useful for earlier stages of language learning. Later stages would benefit from sentences that translate when clicked on. This way translations only are shown when requested.

Figure 4. The Bilingual e-Reader Search Interface.

4.4 The Book Bootstrapper

This application re-orders the text of a book into readability order. This is a similar idea to the Textladder application that sorts a collection of documents into readability order (Ghadirian, 2002), but is done with a single large text. Fiction books with significant amounts of dialogue have many simple sentences and considerable variation in readability across the text. For example, the novel *Emma* by Jane Austen has 31 occurrences of the sentence “Oh!”, 6 of the sentence “Ah!”, and many sentences consisting of a single-word name. Its average sentence length is approximately 21.
If it is true that “meeting” a word 10 times during reading allows one to learn its meaning, then reading a novel of about 85,000 words should provide a learner with a 1,000 word reading vocabulary (Waring, 2009). However, the learner must be able to comprehend enough of the text to gain vocabulary from it. Ordering the text on readability criteria may allow vocabulary to be acquired more comfortably than reading it in the normal order. The learner can then re-read the book in the normal order with greater ease. We have developed a prototype book bootstrapper, and are currently determining its range of usefulness.

4.5 The Text Simplifier

An alternative to the careful selection of text, or the efforts of writing simple text, is the automatic text simplifier. Using similar techniques to document summarisation in addition to word substitution, the text simplifier both shortens a given text and reduces the vocabulary and grammar difficulty. The usefulness of simplification over translation was demonstrated by Eskenazi, Lin and Saz (2013).

4.6 The Text Generator

For extensive reading practice, the text generator should generate interesting stories on a small vocabulary and grammar repertoire. Shim and Kim (2002) developed a story generator that uses autonomous agents. To our knowledge, this idea hasn’t been applied to foreign language learning yet. While not exclusively reading-based, chat-bots can also provide engaging reading practice.

5. Conclusion and Future Work

Reading extensively in the target language improves language skills, and this happens most efficiently when the reading material is both interesting and at a suitable difficulty level. Therefore systems that can provide appropriate reading material in sufficient quantities will be of great benefit to the language learner. We identified seven areas of support that applications can provide for extensive reading: management, generation of readable text, selection of readable text, simplification of text, learner preparation, glossary or translation support, and revision.

We described simple prototype systems that allow the user to locate suitable reading material, either as extracts, or as mixed language texts, exploiting the variability in language difficulty across a typical native text. Other systems were described that are either already in existence or may be worthwhile additions to the range of tools for the language learner. We are currently determining the applicability of one of these: the book bootstrapper. Future work will include determining the effectiveness of the proposed applications for language learning.

Acknowledgements

Software development for the: Readable Extract prototype was by Christopher Dore, Len Wang and XinYang Yao; and the Bilingual eReader was by Wei Yan, Yueyao Zhuang and Ziteng Zhang. The Book Bootstrapper is currently being developed by Daniel Collins.

The Bilingual eReader concept was originally discussed with Laurianne Sitbon, who suggested the use of movie subtitles.

Support for this project has been provided by the Australian Government Office for Learning and Teaching. The views in this project do not necessarily reflect the views of the Australian Government Office for Learning and Teaching.

References

Eskenazi, M., Lin, Y., Saz, O., 2013, Tools for non-native readers: the case for translation and simplification, *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*, NAACL2013, Atlanta, p.20-28

Ghadirian, S. (2002). Providing controlled exposure to target vocabulary through the screening and arranging of texts. *Language Learning and Technology, 6*(1), 147-164.

Heilman, M., Collins-Thompson, K., Callan, J., & Eskenzi, M. (2010) Personalization of Reading passages Improves Vocabulary Acquisition. *International Journal of Artificial Intelligence in Education*, 20, 73-98.

Hirsh, D. & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2):689–696.

B. Laufer. (2000). Task effect on instructed vocabulary learning: the hypothesis of involvement. In *Selected Papers from AILA '99* Tokyo, pages 47–62. Waseda University Press, Tokyo.

Liu, N. & Nation, I.S.P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16(1), 33-42.

Lomicka, L.L. (1998). To gloss or not to gloss: an investigation of reading comprehension online, *Language Learning and Technology*, 1(2), 41-50.

Shim, Y. & Kim, M. (2002). Automatic Short Story Generator Based on Autonomous Agents. *Intelligent Agents and Multi-Agent Systems*. Lecture Notes in Computer Science, 2413, pp 151-162

Uitdenbogerd, A. L. (2006). Web readability and computer-assisted language learning. In Cavedon, L. and Zukerman, I., editors, *Australasian Language Technology Workshop*, pages 99–106.

Uitdenbogerd, A.L. (2010). Fun with Filtering French. *Australasian Language Technology Association Workshop*.

Waring, R. (2009). The inescapable case for extensive reading. *Extensive Reading in English Language Teaching*, Andrzej Cirocki (ed.).

Waring, R. & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2).

Watanabe, H. (2002). Coco & the Gold Flute. *Kanji Dojo*, North Balwyn, Vic. Australia.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.

Weible, D.M. (1980). Teaching reading skills through linguistic redundancy. *Foreign Language Annals*, 6, 487-493.

Zheng, Y., Cheng, L. (2008), Test review: College English Test (CET) in China. *Language Testing*, 25(3), 408-417.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Overview of Grammatical Error Diagnosis for Learning Chinese as a Foreign Language

Liang-Chih YU[a,b], Lung-Hao LEE [c,d*] & Li-Ping CHANG [e]
[a]*Department of Information Management, Yuen-Ze University, Taiwan*
[b]*Innovation Center for Big Data and Digital Convergence, Yuen-Ze University, Taiwan*
[c]*Information Technology Center, National Taiwan Normal University, Taiwan*
[d]*Department of Computer Science and Information Engineering, National Taiwan University, Taiwan*
[e]*Mandarin Training Center, National Taiwan Normal University, Taiwan*
*lhlee@nlg.csie.ntu.edu.tw

Abstract: We organize a shared task on grammatical error diagnosis for learning Chinese as a Foreign Language (CFL) in the ICCE-2014 workshop on Natural Language Processing Techniques for Educational Applications (NLPTEA). In this paper, we describe all aspects of this shared task, including task description, data preparation, evaluation metrics, and testing results. The aim is, through such evaluation campaigns, more advanced computer-assisted Chinese learning techniques will be emerged.

Keywords: Computer-assisted language learning, shared task, Mandarin Chinese

1. Introduction

China's growing global influence has prompted a surge of interest in learning Chinese as a foreign language (CFL), and this trend is expected to continue. However, whereas many computer-assisted learning tools have been developed for use by students of English as a Foreign Language (EFL), support for CFL learners is relatively sparse, especially in terms of tools designed to automatically detect and correct Chinese grammatical errors. For example, while Microsoft Word has integrated robust English spelling and grammar checking functions for years, such tools for Chinese are still quite primitive.

In contrast to the plethora of research related to EFL learning, relatively few studies have focused on computer-assisted language learning for CFL learners. Relative position and parse template language models have been adopted to detect Chinese errors written by US learners (Wu et al. 2010). Machine learning models have been applied to detect word-ordering errors in Chinese sentences from the HSK dynamic composition corpus (Yu and Chen, 2012). Ranking SVM based model has been further explored to rank the candidates and suggest the proper corrections of word ordering errors (Cheng et al. 2014). A penalized probabilistic First-Order Inductive Learning (pFOIL) algorithm has been proposed for grammatical error diagnosis (Chang et al. 2012). Linguistic rule based approach has been presented to detect grammatical errors written by CFL learners (Lee et al. 2013). A sentence judgment system has been implemented to integrate rule-based linguistic analysis and n-gram statistical learning for detecting grammatical errors (Lee et al. 2014). SIGHAN 2013 bakeoff on Chinese spelling check evaluation focus on developing automatic checker to detect and correct spelling errors (Wu et al. 2013). In summary, human language technologies for Chinese learning have attracted more attentions in recent years.

In the ICCE-2014 workshop on Natural Language Processing Techniques for Educational Applications (NLPTEA), we organize a shared task on Chinese grammatical error diagnosis that provides an evaluation platform for developing and implementing computer-assisted learning tools. The data sets in our task are collected from the Chinese as the Foreign Language (CFL) learners' written essays. Given a sentence with/without one of grammatical errors, *i.e.*, redundant word, missing word, word disorder, and word selection, the developed system should indicate whether contains grammatical errors and further points out which one of defined error types. The hope is that, through such evaluation campaigns, more advanced Chinese grammatical error detecting techniques will be emerged.

We give an overview of the shared task on grammatical error diagnosis for learning Chinese as a foreign language. The rest of this article is organized as follows. Section 2 details the designed task. Section 3 introduces the data sets provided in this evaluation. Section 4 proposes the evaluation metrics. Section 5 presents the results of participants' approaches for performance comparison. Finally, we conclude this paper with the findings and future research direction in the Section 6.

2. Shared Task Description

The goal of this shared task is developing the computer-assisted tools to detect several kinds of grammatical errors, that is, redundant word, missing word, word disorder, and word selection. The input sentence contains one of defined error types. The developed tool should indicate which kind of error type is embedded in the given sentence. If the input sentence, which is given a unique sentence number SID, contains no grammatical errors, the tools should return "SID, Correct". If an input sentence contains a defined grammatical error, the output format should be "SID, error_type". We simplify the task that there are only one error type may be in the given sentence. Examples are shown as follows. In example 1, the character "被" is a redundant word. There is a missing word "有" in the example 2 and its correct usage is shown in example 3. The sentence in the example 4 has word disorder error, *i.e.*, the word "很早" should be preceded the word "起床". The word "一個" in the example 5 is an incorrect word selection, the correct word should be "一件".

- Example 1
  Input: (sid=B2-1447-6)　　希望沒有人再被食物中毒
  Output: B2-1447-6, Redundant
- Example 2
  Input: (sid=C1-1876-2)　　對社會國家不同的影響
  Output: C1-1876-2, Missing
- Example 3
  Input: (sid=C1-1876-2)　　對社會國家有不同的影響
  Output: C1-1876-2, Correct
- Example 4
  Input: (sid=A2-0775-2)　　我起床很早
  Output: A2-0775-2, Disorder
- Example 5
  Input: (sid=B1-0110-2)　　我會穿著一個黃色的襯衫
  Output: B1-0110-2, Selection

3. Data Sets

Mandarin Training Center (MTC) of National Taiwan Normal University (NTNU) was founded in 1956 for teaching Chinese as a foreign language. Currently, MTC is the most renowned Chinese language center in Taiwan, in which around 1700 CFL learners from more than 70 countries enrolled each academic quarter. The learner corpus used in our task is collected from the computer-based writing Test of Chinese as a Foreign Language (TOCFL). The writing test is designed according to the six proficiency levels of the Common European Framework of Reference (CEFR). Test takers have to complete two different tasks for each level. For example, for the A2 (Waystage level) candidates, they will be asked to write a note and describe a story after looking at four pictures. All candidates are asked to complete the writings on line.

We further ask the annotators to label the grammatical errors in CFL learners' written sentences and provide their correct usage. Our prepared data is further divided into three distinct sets. (1) **Training set**: 1,506 CFLs' writings are collected in which 5,607 grammatical errors are annotated. Each CFL learners' writing is represented in SGML format shown in Figure 1. The title attribute is used to describe the topic of the writing test. There is only one grammatical error in an annotated sentence. The error types are also indicated along with their corresponding correct usages. All sentences in this set can be used to train the developed grammatical error detection tool. (2) **Dryrun set**: Total 33 sentences are given for participants to familiarize themselves with the final testing process. Each participant can submit several runs generated using different models with different parameter settings. In addition to make sure the submitted results can be correctly evaluated, participants can fine-tune their developed models in the dryrun phase. The purpose of dryrun is for output format validation only. No matter which performance can be achieved that will not be included in our official evaluation. (3) **Test set**: In total, there are 1,750 testing sentences. A half of these instances contain no grammatical errors. Another half of testing cases includes one grammatical error per sentence. The number of error type redundant, missing, disorder, and selection is 279, 350, 120, and 126, respectively. The distribution is the same with our given training set. The policy of our evaluation is an open test. In addition to our provided data sets, registered research teams can employ any linguistic and computational resources to detect grammatical errors in the sentences.

```
<ESSAY title="寫給即將初次見面的筆友的一封信">
<TEXT>
<SENTENCE id="B1-0112-1">我的計畫是十點早上在古亭捷運站</SENTENCE>
<SENTENCE id="B1-0112-2">頭會戴著藍色的帽子</SENTENCE>
</TEXT>
<MISTAKE id="B1-0112-1">
<TYPE>Disorder</TYPE>
<CORRECTION>我的計畫是早上十點在古亭捷運站</CORRECTION>
</MISTAKE>
<MISTAKE id="B1-0112-2">
<TYPE>Missing</TYPE>
<CORRECTION>頭上會戴著藍色的帽子</CORRECTION>
</MISTAKE>
</ESSAY>
```

Figure 1. An essay represented in SGML format.

4. Performance Metrics

Table 1 shows the confusion matrix used for performance evaluation. In the matrix, True Positive (TP) is the number of sentences with grammatical errors that are correctly proposed by the developed tool; False Positive (FP) is the number of sentences without grammatical errors that are incorrectly proposed; True Negative (TN) is the number of sentences without grammatical errors that are identified correctly; False Negative (FN) is the number of sentences with grammatical errors that are incorrectly regarded as correct sentences.

Table 1: The confusion matrix for evaluation.

Confusion Matrix		**System Result**	
		Positive (With grammatical errors)	Negative (Without grammatical errors)
Gold Standard	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

The criteria for judging correctness are distinguished into two levels. (1) **Detection level**: all error types are regarded as incorrect. Binary classification of a testing instance, *i.e.*, correct or incorrect, should be completely identical with the gold standard. (2) **Identification level**: this level could be considered as a multi-class categorization problem. In addition to correct instances, all error types should be clearly identified, *i.e.*, Redundant, Missing, Disorder, and Selection. The following metrics are measured in both levels with the help of the confusion matrix.

- **False Positive Rate** (FPR) = FP / (FP+TN)
- **Accuracy** = (TP+TN) / (TP+FP+TN+FN)
- **Precision** = TP / (TP+FP)
- **Recall** = TP / (TP+FN)
- **F1** = 2 * Precision *Recall / (Precision + Recall)

For example, give 8 testing inputs with gold standards shown as "A2-0802-4, correct", "A2-3344-1, Selection", "B1-3419-8, Missing", "B1-3520-3, correct", "B2-1918-7, correct", "B2-2231-4, Disorder", "C1-1744-1, Redundant", and "C1-1873-7, correct". The system may output the result shown as "A2-0802-4, Disorder", "A2-3344-1, Redundant", "B1-3419-8, Selection", "B1-3520-3, correct", "B2-1918-7, correct", "B2-2231-4, Disorder", "C1-1744-1, Redundant", and "C1-1873-7, Missing". The evaluation tool will yield the following performance.

- **False Positive Rate (FPR)** = 0.5 (= 2 / 4).
  Notes: {"A2-0802-4, Disorder", "C1-1873-7, Missing"} /{"A2-0802-4, correct", "B1-3520-3, correct", "B2-1918-7, correct", "C1-1873-7, correct"}
- **Detection Accuracy** = 0.75 (=6/8).
  Notes: {"A2-3344-1, Redundant", "B1-3419-8, Selection", "B1-3520-3, correct", "B2-1918-7, correct", "B2-2231-4, Disorder", "C1-1744-1, Redundant"} /{"A2-0802-4, correct", "A2-3344-1, Selection", "B1-3419-8, Missing", "B1-3520-3, correct", "B2-1918-7, correct", "B2-2231-4, Disorder", "C1-1744-1, Redundant", "C1-1873-7, correct"}
- **Detection Precision** = 0.67 (=4/6).
  Notes: {"A2-3344-1, Redundant", "B1-3419-8, Selection", "B2-2231-4, Disorder", "C1-1744-1, Redundant"} /{"A2-0802-4, Disorder", "A2-3344-1, Redundant", "B1-3419-8, Selection", "B2-2231-4, Disorder", "C1-1744-1, Redundant", "C1-1873-7, Missing"}
- **Detection Recall** = 1 (=4/ 4).
  Notes: {"A2-3344-1, Redundant", "B1-3419-8, Selection", "B2-2231-4, Disorder", "C1-1744-1, Redundant"} /{"A2-3344-1, Selection", "B1-3419-8, Missing", "B2-2231-4, Disorder", "C1-1744-1, Redundant"}
- **Detection F1**= 0.8024 (=2*0.67*1/(0.67+1))
- **Identification Accuracy** = 0.5 (=4/8).
  Notes: {"B1-3520-3, correct", "B2-1918-7, correct", "B2-2231-4, Disorder", "C1-1744-1, Redundant"} /{"A2-0802-4, correct", "A2-3344-1, Selection", "B1-3419-8, Missing", "B1-3520-3, correct", "B2-1918-7, correct", "B2-2231-4, Disorder", "C1-1744-1, Redundant", "C1-1873-7, correct"}
- **Identification Precision** = 0.33 (=2/6).
  Notes: {"B2-2231-4, Disorder", "C1-1744-1, Redundant"} /{"A2-0802-4, Disorder", "A2-3344-1, Redundant", "B1-3419-8, Selection", "B2-2231-4, Disorder", "C1-1744-1, Redundant", "C1-1873-7, Missing"}
- **Identification Recall** = 0.5 (=2/ 4).
  Notes: {"B2-2231-4, Disorder", "C1-1744-1, Redundant"} /{"A2-3344-1, Selection", "B1-3419-8, Missing", "B2-2231-4, Disorder", "C1-1744-1, Redundant"}
- **Identification F1**= 0.3976 (=2*0.33*0.5/(0.33+0.5))

5. Evaluation Results

Table 2 shows the participant teams and their testing submission statistics. Our shared task attracted 13 research teams. There are 4 teams that come from Taiwan, *i.e.*, AS, KUAS & NTNU, NCYU, and NTOU. 3 teams originate from China, *i.e.*, HITSZ, PKU, and PolyU. The remaining 6 teams are CIRU from United States of America, MU from New Zealand, SPBU from Russia, TMU from Japan, UL from United Kingdom, and UDS from Germany. Among 13 registered teams, 6 teams submitted their testing results. In total, we had received 13 runs in the formal testing phase.

Table 2: Result submission statistics of all participants.

Participants (Ordered by abbreviations of names)	**#Submissions**
Academia Sinica (**AS**)	0
Confucius Institute of Rutgers University (**CIRU**)	1
Harbin Institute of Technology Shenzhen Graduate School (**HITSZ**)	0
National Kaohsiung University of Applied Sciences & National Taiwan Normal University (**KUAS & NTNU**)	3
Massey University (**MU**)	0
National Chiayi University (**NCYU**)	1
National Taiwan Ocean University (**NTOU**)	2
Peking University (**PKU**)	0
The Hong Kong Polytechnic University (**PolyU**)	0
Saint Petersburg State University (**SPBU**)	0
Tokyo Metropolitan University (**TMU**)	2
Saarland University (**UDS**)	4
University of Leeds (**UL**)	0
Total	**13**

Table 3: Testing results of our shared task.

Submission	**FPR**	**Detection Level**				**Identification Level**			
		Acc.	**Pre.**	**Rec.**	**F1**	**Acc.**	**Pre.**	**Rec.**	**F1**
CIRU-Run1	0.496	**0.6446**	**0.6128**	0.7851	**0.6884**	**0.4589**	**0.4548**	0.4137	**0.4333**
KUAS & NTNU-Run1	0.904	0.5006	0.5003	0.9051	0.6444	0.2149	0.2696	0.3337	0.2983
KUAS & NTNU-Run2	0.2686	0.5217	0.5374	0.312	0.3948	0.4109	0.2516	0.0903	0.1329
KUAS & NTNU-Run3	0.904	0.5006	0.5003	0.9051	0.6444	0.2074	0.2607	0.3189	0.2869
NCYU-Run1	**0.1189**	0.4983	0.4927	0.1154	0.187	0.4594	0.2409	0.0377	0.0652
NTOU-Run1	1	0.5	0.5	**1**	0.6667	0.16	0.2424	0.32	0.2759
NTOU-Run2	1	0.5	0.5	**1**	0.6667	0.2074	0.2932	**0.4149**	0.3436
TMU-Run1	0.1977	0.5171	0.5399	0.232	0.3245	0.4554	0.3545	0.1086	0.1662
TMU-Run2	0.1691	0.5103	0.5287	0.1897	0.2792	0.4531	0.3084	0.0754	0.1212
UDS-Run1	0.792	0.4914	0.4945	0.7749	0.6037	0.2337	0.2467	0.2594	0.2529
UDS-Run2	0.6286	0.4949	0.4959	0.6183	0.5504	0.2869	0.2435	0.2023	0.221
UDS-Run3	0.5783	0.4949	0.4955	0.568	0.5293	0.3057	0.247	0.1897	0.2146
UDS-Run4	0.2491	0.5046	0.509	0.2583	0.3427	0.428	0.2968	0.1051	0.1553

Table 3 shows the testing results of our shared task. In addition to achieving promising detection effects of grammatical errors, reducing the false positive rate, which is percentage of the correct sentences that are incorrectly reported containing grammatical errors, is also important. The research teams, NCYU and TMU, achieved relatively low false positive rates.

Detection level evaluations are designed to detect whether a sentence contains grammatical errors or not. Accuracy is usually adopted to evaluate the performance, but it is affected by the distribution of testing instance. The baseline can be achieved easily by always guessing without errors. That is accuracy of 0.5 in this evaluation. Some systems achieved slightly better than the baseline, *i.e.*, CIRU, KUAS&NTNU, TMU and UDS. Registered teams may send different runs that aimed at optimizing the recall or precision rates. These phenomena guide us to adopt F1 score to reflect the tradeoff between precision and recall. In the testing results, CIRU accomplished the best detection effects of indicating grammatical errors, which resulted the best F1 score 0.6884. For identification level evaluations, the systems need to identify the error types in the given sentences. The research team came from CIRU accomplished the best correction accuracy 0.4589. Most systems cannot effectively identify the input sentences to point out possible grammatical errors. Our testing results indicate that the system developed by CIRU accomplished the best identification F1 0.4333.

In summary, it is a really difficult task to develop the computer-assisted learning tool for grammatical error diagnosis, especially learning Chinese as a foreign language, since there are only target sentences without the help of their context. We cannot find a relatively promising system according to our testing results. In general, this research problem still has long way to go.

6. Conclusions and Future Work

This paper describes the overview of NLPTEA 2014 shared task on grammatical error diagnosis for learning Chinese as a foreign language. We introduce the task designing ideas, data preparation details, evaluation metrics, and the results of performance evaluation. This task also encourages researchers to bravely propose various ideas and implementations for possible breakthrough. No matter how well their implementations would perform, they contribute to the community by enriching the experience that some ideas or approaches are promising or impractical, as verified in this shared task. Their reports in the proceeding will reveal the details of these various approaches and contribute to our knowledge about computer-assisted Chinese learning.

All data sets and their accompanying gold standards and evaluation tool are publicly available for research purposes at http://ir.itc.ntnu.edu.tw/lre/nlptea14cfl.htm. We hope our provided data can serve as a benchmark to help developing better Chinese learning tools. This shared task also motivates us to build more language resources in the future to possibly improve the state-of-the-art techniques.

Acknowledgements

This research was supported by the Ministry of Science and Technology, under the grant MOST 102-2221-E-155-029-MY3, 103-2221-E-003-013-MY3, 103-2911-I-003-301 and the “Aim for the Top University Project” and “Center of Learning Technology for Chinese” of National Taiwan Normal University, sponsored by the Ministry of Education, Taiwan.

References

Chang, R.-Y., Wu, C.-H., & Prasetyo, P. K. (2012). Error diagnosis of Chinese sentences using inductive learning algorithm and decomposition-based testing mechanism. *ACM Transactions on Asian Language Information Processing*, 11(1), article 3.

Cheng, S.-M., Yu, C.-H., & Chen, H.-H. (2014). Chinese word ordering errors detection and correction for non-native Chinese language learners. *Proceedings of COLING’14* (pp. 279-289), Dublin, Ireland: ACL Anthology.

Lee, L.-H., Chang, L.-P., Lee, K.-C., Tseng, Y.-H., & Chen, H.-H. (2013). Linguistic rules based Chinese error detection for second language learning. *Proceedings of ICCE’13* (pp. 27-29). Bali, Indonesia: Asia-Pacific Society for Computers in Education.

Lee, L.-H., Yu, L.-C., Lee, K.-C., Tseng, Y.-H., Chang, L.-P., & Chen, H.-H. (2014). A sentence judgment system for grammatical error detection. *Proceedings of COLING’14* (pp. 23-29), Dublin, Ireland: ACL Anthology.

Wu, C.-H., Liu, C.-H., Harris, M., & Yu, L.-C. (2010). Sentence correction incorporating relative position and parse template language models. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(6), 170-1181.

Wu, S.-H., Liu, C.-L., & Lee, L.-H. (2013). Chinese spelling check evaluation at SIGHAN bake-off 2013. *Proceedings of SIGHAN’13* (pp. 35-42), Nagoya, Japan: ACL Anthology.

Yu, C.-H. & Chen, H.-H. (2012). Detecting word ordering errors in Chinese sentences for learning Chinese as a foreign language. *Proceedings of COLING’12* (pp. 3003-3017), Mumbai, India: ACL Anthology.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

KNGED: a Tool for Grammatical Error Diagnosis of Chinese Sentences

Tao-Hsing CHANG[a*], Yao-Ting SUNG[b], Jia-Fei HONG[c] & Jen-I CHANG[b]

[a]*Department of Computer Science and Information Engineering, National Kaohsiung University of Applied Sciences, Taiwan*
[b]*Department of Educational Psychology and Counseling, National Taiwan Normal University, Taiwan*
[c]*Department of Applied Chinese Language and Culture, National Taiwan Normal University, Taiwan*
*changth@gm.kuas.edu.tw

Abstract: The main purpose of this paper is to propose a method that can automatically detect whether there are any grammatical errors as well as identify their error types. The framework of this method is based on a rule base to identify common grammatical errors. This rule base contains manually constructed rules and rules that are automatically machine generated. This paper further proposes algorithms which can apply these rules to determine whether a sentence is incorrect as well as what types of errors it belongs to. Experimental results show that the F1-measure of the proposed method is 0.64 and 0.30 on detection and identification, respectively.

Keywords: KNGED, Chinese grammar, grammatical error, CFL, automatic diagnosis, rule-based method.

1. Introduction

Automatically detecting grammatically incorrect sentences is fundamental and important for numerous NLP studies and related applications. For instance, in language teaching, automatic detection of grammatically incorrect sentences produced by learners can help a teacher teach grammar more effectively. However, the detection of grammatically incorrect sentences in Chinese is challenging. The main reason is that it is difficult to detect sentence boundaries in Chinese. In English, the period provides a clear signal of the end of a sentence, allowing the segment of text between two periods to be taken as a sentence and analyzed grammatically. However, two periods between Chinese sentences represent a complete semantic expression. An excessively long sentence may contain several commas as delimiters. Moreover, a sentence segment formed by a comma may comprise the complete sentence, a clause and even a phrase. This phenomenon makes the detection of sentence boundaries in Chinese difficult.

The above difficulty makes the method of detecting grammatical errors in Chinese sentences via using a parser to completely deconstruct a parsing tree infeasible. To detect errors in an English sentence, a parsing tree constructed using a parser can provide criteria for sentence judgment. However, to deal with grammatical errors generated by learners of Chinese as a second language, the parsing tree method may not have been thoroughly examined. This is because the main cause of common errors by second language learners is the language transfer phenomenon in language learning. For example, Korean students often write the following incorrect sentence.

我 來 台灣 四年 工作 了

The error pattern here is that the time noun ‘四年’ appears before the verb ‘工作’. This common error occurs because Korean is a subject-object-verb (SOV) language. An important characteristic of an SOV language is that all other elements such as nouns, adverbs, and numbers come before the verb. However, Chinese is a subject-verb-object (SVO) language, so the correct sentence would be '我 來 台灣 四年 工作 了 (I came to work in Taiwan for four years)' , where the verb '工作(work) ' must appear before the time noun'四年 (four years)'.

Consequently, by generalizing common errors made by learners of Chinese as a second language, it is possible to further analyze which specific syntactic structures those common errors belong to. These specific syntactic structures are considered error detection rules. Moreover, the error detection rules possess a syntactic error pattern and its corresponding syntactically correct pattern. Sufficient patterns and rules enable the generation of a rule base. Sentences can be compared using the rule base to identify grammatical errors. If a sentence contains multiple segments that conform to error detection rules, this segment is most likely the syntactic structure of that error. Therefore, sufficient error detection rules are collected, grammatical errors can be identified by comparing the rule base. The method of error detection rule collection can be obtained through analysis of sentences in a learner error corpus. The bigger the corpus, the more error detection rules can be generated and the more grammatical errors detected.

This study primarily proposes a method capable of automatically detecting and identifying grammatical errors. The framework of this method is based on a rule base to identify common grammatical errors. This rule base contains manually constructed rules and rules that are automatically machine generated. This study further proposes algorithms that can apply error detection rules to determine whether a sentence is incorrect and what types of errors it belongs to and classify any errors.

This paper is organized as follows. Section 2 reviews some related research and illustrates the impact of these studies on the research motivations. Section 3 then lists the corpus used in this paper and illustrates a learner corpus that is designed to automatically detect grammatically incorrect sentences produced by Chinese as second language writers. Next, section 4 introduces the method to manually construct error detection rules and the method for the program to automatically generate error detection rules. Subsequently, section 5 describes the algorithm for automatically identifying incorrect sentences. Section 6 demonstrates the performance of the proposed approach, while Section 7 draws conclusions.

2. Related Works

Recently, Chinese learning has become a growing trend, making Chinese one of the most popular foreign languages globally, besides English. To learners, learning a new language frequently involves grammar difficulties, and grammatically incorrect sentences are a common error. Previous research on second language acquisition indicated that effective provision of corrective feedback can contribute to the development of grammatical competence in second language learners (Fathman and Whalley, 1990; Ashwell, 2000; Ferris and Robers, 2001; Chandler, 2003). Currently, in the field of natural language processing, the development of tools and technologies to automatically detect grammatical errors is an important research trend.

On the one hand, regarding common error types made by learners of English as a second language] and the development of the related automatic detection research, Donahue (2001) used the error taxonomy of native English learners proposed by Connors and Lunsfor (1998) to analyze two hundred writing tests taken by learners of English as a second language. The most common error types committed by learners of English as a second language were found to differ from those of native English learners. The three most common error types of learners of English as a second language were incorrect usage of commas, incorrect word usage, and missing words. However, this corpus was insufficient to understand common error types committed by most learners of English as a second language. Cambridge University Press collaborated with the University of Cambridge to create the Cambridge Learner Corpus (CLC), which tags approximately 16 million words. Among these words, the three most common error types are incorrect word selection, preposition errors, and determiner errors (Nicholls, 2003).

During the past ten years, natural language processing specialists have designed automated grammatical error detection techniques and tools focused on common error types in the corpus. Examples include preposition error detection by Eeg-Olofsson and Knuttson (2003), Tetreault and Chodorow (2008), DeFelice and Pulman (2009) and Tetreault and Chodorow (2009), article and preposition error detection by Gamon et el. (2009) and Dale and Kilgarriff (2011), determiner and proposition error detection by Dale *et al.* (2012), and determiner, article, and proposition error detection by Ng *et al.* (2013).

On the other hand, regarding common error types in learners of Chinese as a second language, Wang (2011) observed that the most common grammatical error types among Chinese learners whose mother tongue is English are missing language components, incorrect word order, and incorrect sentence structure. Additionally, analysis of the HSK corpus of 35,884 erroneous sentences has demonstrated that the three most common error types are incorrect word order, missing adverb components, and missing predicate components (Cheng *et al.* 2014). With the development of related automatic detection research, Cheng *et al.* (2014) and Yu and Chen (2012) designed word order error detection technology focused on the Chinese sentences in the HSK Dynamic Composition Corpus. In developing a sentence grammatical error detection system, Lee *et al.* (2014) further used the HSK Dynamic Composition Corpus and additional manually constructed rules of common Chinese sentence errors.

The above literature indicates that, in English learning, there exists widespread use of learning assistance tools developed from natural language processing technology. These tools can automatically detect and correct the grammatical errors of learners. This is valuable as a means to help learners learn correct grammar and improve their compositional skills (Chodorow *et al.*, 2012; Leacock, Chodorow, Gamon, &Tetreault, 2010). However, little research has examined automatic detection of Chinese grammatical errors. This study proposes the integration of rule-based and machine learning methods to identify reliable rules from the corpora to detect the grammatical errors of learners of Chinese as a second language.

3. Corpora

This study seeks to obtain reliable rules to detect grammatical errors committed by learners of Chinese as a second language. The three corpora used in this study include (1) the dry run data provided by the convention; (2) the formal run data provided by the convention; (3) Chinese Written Corpus developed herein. The following focuses on introducing Chinese Written Corpus.

This study has continually developed a Chinese Written Corpus primarily comprising a single topic at different levels. This corpus was developed using Chinese composition scoring guidelines based on the ACTFL (2012) language proficiency criteria. The research samples are compositions written by foreign students who have learned enough Chinese to have basic competence. Samples were collected from September 2010 to June 2013. The source of the corpus is foreign Chinese learners studying at the National Taiwan Normal University Mandarin Training Center and 11 other Taiwanese Chinese educational institutions. The corpus currently includes foreign learners representing 37 different mother languages. During composition collection, complete information was collected on each composition. This information included the title of the composition, the Chinese and English names, nationality, and mother tongue of the learner, and the Chinese education institution in Taiwan. This information was saved as text and image documents. Currently, the texts of this corpus deal with two topics, and there are 1,147 compositions in total, comprising approximately 750 thousand words.

Following the creation of the corpus, each composition text was assessed by two experts or personnel trained in evaluation. To ensure reliability, the texts were cross-evaluated using the Chinese Composition Scoring Standard. This standard assigns compositions to different rankings of Distinguished, Superior, Advanced, Intermediate, and Novice. The Advanced, Intermediate, and Novice categories are each divided into three subcategories, including High, Medium, and Low, amounting for a total of 11 categories. These 11 categories account for Chinese users of all levels, from learners unable to construct a full sentence to native level writers. This study manually scored the compositions in each complete topic based on the above scoring standards and procedures. The composition scores were collected, and learner and corpus information were inputted into an error tagging system developed herein for compositions by learners of Chinese as a second language.

This error tagging system compiles learner and corpus information, and also includes word segmentation, part-of-speech tagging, and error tagging functions. In the main error tagging system, methods and standards for the analysis of learner language errors can be roughly divided into “linguistic form taxonomy” and “surface structure taxonomy”. Linguistic form taxonomy classifies error types – word class, sentence, and specific sentence errors – using language components as a structure. Meanwhile, surface structure taxonomy classifies error types using their structure. That is, it compares the correct and incorrect forms. Typical surface structures comprise four categories: omission, addition, selection, and disorder (Dulay, Burt & Krashen, 1982; James, 1998). The function of error tagging in this study integrates the two taxonomies, first classifying errors based on the surface structure, then carefully analyzing them based on the language form.

4. Rule Generation and Extraction

Based on the above three data types used in this study, this section explains the optimal method of generating error detection rules to identify ungrammatical sentences.

4.1 Manually Constructed Rules

The study uses five steps to generate manually constructed rules. First, based on the training data provided in this shared task, this study handcrafted syntactic patterns of grammatically incorrect sentences and corrected sentences. Second, to ensure the reliability of manually constructed rules for detecting incorrect sentences, this study also devised a program in which the Chinese Written Corpus developed in this study is embedded. Thirdly, on program completion, we enter syntactic patterns of grammatically incorrect sentences into the interface, and the program can then show the number of sentences contained in the Chinese Written Corpus. Moreover, those sentences conform to syntactic patterns of grammatically incorrect sentences.

Meanwhile, this study entered syntactic patterns of corrected sentences into the program, and then recorded the number of sentences contained in the Chinese Written Corpus, as well as those sentences that conform to the syntactic patterns of corrected sentences. Finally, this study retains the number of syntactic patterns of corrected sentences such that it exceeds that of incorrect sentences. These rules are considered the reliable error detection rules for identifying grammatically incorrect sentences in formal run data. This study contains 840 manually constructed rules, which contain 90 rules for identifying sentences with Missing words, 73 for identifying sentences with Redundant words, 51 for identifying sentences with Selection words, and 626 for identifying sentences with Word disorder.

4.2 Machine Generated Rules

The advantage of manually constructed rules is that complex rules can be detected with high accuracy. However, using manually constructed rules to identify grammatical errors suffers from a disadvantage. Specifically, the number of manually constructed rules is limited, and errors may exist. This study thus employs a program to retrieve syntactic rules of ungrammatical sentences from the learner corpus.

Unlike manually constructed rules, the rules generated by the program are fixed in length. For example, the learner corpus contains the following sentence.

這些	地方	是	在	巴西
Neqa	Na	SHI	P	Nc

In this sentence, each part of speech is labeled. This sentence in the learner corpus is tagged as the Redundancy error, and ‘SHI’ is a redundant word. We hypothesize that every word in this sentence can be collocated with its beginning and end, and their parts-of-speech to generate rules. Therefore, we combine “是” and its part-of-speech “SHI” with the first and last parts of the word “是” and their associated parts of speech, which yields 32 possible Redundant rules, as shown in Fig. 1.

In Figure 1, the symbol "+" represents two adjacent words or parts-of-speech, while the symbol ">" indicates that both the front and the back of a word or its associated part-of-speech should not be adjacent to that symbol. For a rule *pr* included in these 32 possible rules, if it meets the following criteria, it will be recognized as an error detection rule:

$$positive(pr) > p \quad \text{and} \quad r > k,$$
$$r = positive(pr) \,/\, negative(pr)$$

where *positive(pr)* indicates the number of *pr* that occurred in the corpus with erroneous sentences; and *negative(pr)* indicates the number of *pr* that occurred in the corpus with correct sentence. In this study, the value of *positive(pr)* divided by *negative(pr)* is denoted as the *r*-value. The *r*-value of rules used by the grammatical error diagnosis algorithm described in Section 5.

Parameters *p* and *k* are thresholds obtained via experiment. Larger *p* is associated with more occurrence of rule *pr* in the incorrect sentences. That is, the rule of *pr* does not appear randomly. Meanwhile, larger *k* represents the possibility of a high degree of precision when using *pr* to identify a sentence as erroneous. Take 32 rules in Fig. 1 for example; if *p* and *k* are set to 2, then just 11 rules with borders in Fig. 1 are collected in the rule base for detection. This study uses the above method to automatically generate 13,890 Redundant rules and 2,497 Missing rules.

(1)	這些>是+在	(17)	Neqa>是+在
(2)	這些>是+P	(18)	Neqa>是+P
(3)	這些>是>巴西	(19)	Neqa>是>巴西
(4)	這些>是>Nc	(20)	Neqa>是>Nc
(5)	地方+是+在	(21)	Na+是+在
(6)	地方+是+P	(22)	Na+是+P
(7)	地方+是>巴西	(23)	Na+是>巴西
(8)	地方+是>Nc	(24)	Na+是>Nc
(9)	這些>SHI+在	(25)	Neqa>SHI+在
(10)	這些>SHI+P	(26)	Neqa>SHI+P
(11)	這些>SHI>巴西	(27)	Neqa>SHI>巴西
(12)	這些>SHI>Nc	(28)	Neqa>SHI>Nc
(13)	地方+SHI+在	(29)	Na+SHI+在
(14)	地方+SHI+P	(30)	Na+SHI+P
(15)	地方+SHI>巴西	(31)	Na+SHI>巴西
(16)	地方+SHI>Nc	(32)	Na+SHI>Nc

Figure 1. Examples of Rules Generated by Machine.

5. Grammatical Error Diagnosis Algorithm

For each sentence, the following steps are performed to determine whether it is incorrect.

Step 1. Check for rules that conform to the error detection rule of Word selection. If such rules exist, the sentence is considered to contain Word selection error and so the error identification is concluded].

Step 2. Check whether rules exist that conforms to the error detection rule of Word disorder. If so, the sentence is considered to contain Word order error and so the error identification is concluded.

Step 3. Check for rules that conform to the error detection rule for Redundant and Missing words.

Step 3.1. If the rule only conforms to one of the error detection rules related to redundant or missing words, then it is considered a sentence that contains that type of error and so the error identification is concluded.

Step 3.2. If the rule simultaneously conforms to more than one error detection rule of redundant or missing words, then among the rules that conforms to both types of error, that with the highest *r* value is selected.

Step 3.3. It is assumed that among the Missing word rules, the highest value of *r* is *mr*, and among the Redundant word rules, the highest value of *r* is *rr*. If the *r* value of *rr* exceeds *y* times that of *mr*, the sentence is considered to suffer from Redundant word error; otherwise, it is considered to suffer from Missing word error. The identification is concluded following sentence judgment.

Step 4. If the sentence is not recognized as erroneous via the last three steps, then it is considered correct.

Because different types of error detection rules exert different effects, based on analysis of error detection rules from the dry run corpus, their effectiveness reveals that the Selection has higher accuracy than other types of rule. Consequently, when a sentence is identified as containing segments of the rule of Selection, it is recognized that the sentence contains that type of error. Similarly, although the accuracy of the Word disorder rule is lower than that of the Selection rule, it is far higher than that of the Redundant word and Missing word rules. Therefore, when a sentence is identified as containing the Word disorder rule, it is first recognized that the sentence contains that type of error.

Compared to the Missing word rule, the redundant word rule can more easily obtain a higher *r* value. Thus, if the *r*-value of the Redundant word rule must exceed the missing word rule by y times, then the result of the detection of the rule of Redundant word can be reliable; otherwise, the sentence should be recognized as containing a Missing word error. The next section illustrates the value of each parameter used in the proposed method.

6. Experimental Results

In the NLPTEA 2014 CFL shared task, three parameters are established and combined with three runs to evaluate the effectiveness of the proposed method. In Run 1, the p-value is 3, the k-value is 2, and the y-value is 50. In Run 2, the p-value is 10, the k-value is 1000 and the y-value is 50. In Run 3, the p-value is 3, the k-value is 2 and the y-value is 1. Table 1 lists the experimental results.

Table 1: An example of a table for the ICCE proceedings.

Submission		Run1	Run2	Run3
False Positive Rate		0.9040	0.2686	0.9040
Detection Level	Accuracy	0.5006	0.5217	0.5006
	Precision	0.5003	0.5374	0.5003
	Recall	0.9051	0.3120	0.9051
	F1	0.6444	0.3948	0.6444
Identification Level	Accuracy	0.2149	0.4109	0.2074
	Precision	0.2696	0.2516	0.2607
	Recall	0.3337	0.0903	0.3189
	F1	0.2983	0.1329	0.2869

7. Discussion

We have made a few discoveries regarding the process of this experiment and the results obtained. First, manually constructed rules are more complicated than machine-generated rules. However, the accuracy of manually constructed rules does not necessarily exceed that of machine generated rules. Fairly reliable error detection rules can be obtained by establishing parameters based on automatically generated rules. Second, many automatically generated rules are not listed in manually constructed rules. This means the method of using machines to identify error detection rules is feasible. Considering these two perspectives, if the program has an enhanced ability to search for rules, then it is feasible to fully automatically identify grammatical errors made by Chinese as second language learners.

Several aspects of our proposed method can be further improved. First, rules in this study are primarily based on Chinese written error corpus. However, the corpus currently remains in the expansion phase. The increasingly rich content of the corpus can enhance the system performance. Second, only Redundant word and Missing word errors can be automatically generated by the current program. Also, the error detection rules contains only three terms. If more types of rules that are automatically generated by the program can be added in the program and the program can identify more complex rules, the system performance will be further improved.

Acknowledgements

This research is partially supported by the 'Aim for the Top University Project' and 'Center of Learning Technology for Chinese' of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the "International Research-Intensive Center of Excellence Program" of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301

References

ACTFL Proficiency Guidelines 2012 - Writing. (2012). Retrieved August 25, 2014, from http://actflproficiencyguidelines2012.org/writing

Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by form feedback the best method? Journal of Second Language Writing, 9, 227–57.

Cheng, S. M., Yu, C. H., & Chen, H. H. (2014). Chinese Word Ordering Errors Detection and Correction for Non-Native Chinese Learners. *Proceedings of COLING 2014*, (pp. 279-289), Dublin, Ireland.

Chang, T. H., Sung, Y. T., & Lee, Y. T. (2012). A Chinese word segmentation and POS tagging system for readability research. *Proceedings of SCiP 2012*, Minneapolis, MN.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. Journal of Second Language Writing, 12, 267–96.

Chodorow, M., Dickinson, M., Israel, R., & Tetreault, J. R. (2012). Problems in Evaluating Grammatical Error Detection Systems. *Proceedings of COLING 2012* (pp. 611-628), Mumbai, India.

Connors, R. J., & Lunsford, A. A. (1988). Frequency of formal errors in current college writing, or ma and pa kettle do research. *College Composition and Communication, 39*(4), 395-409.

Dale, R., Anisimoff, I., & Narroway, G. (2012). HOO 2012: A report on the preposition and determiner errorcorrection shared task. *Proceedings of the 7th Workshop on Building Educational Applications Using NLP* (pp. 54-62), Montréal, Canada.

Dale, R., & Kilgarriff, A. (2011). Helping our own: The HOO 2011 Pilot Shared Task. *Proceedings of the 13th European Workshop on Natural Language Generation,* Nancy, France.

De Felice, R., & Pulman, S. 2009. Automatic detection of preposition errors in learner writing. *CALICO Journal, 26*(3), 512-528.

Donahue, S. (2001). Formal errors: Mainstream and ESL students. Presented at the 2001 Conference of the Two-Year College Association (TYCA); cited by Leacock et al. 2010.

Dulay, H. C., Burt, M. K., & Krashen, S. D. (1982). Language Two. New York: Oxford University Press.

Eeg-Olofsson, J., & Knuttson, O. (2003). Automatic grammar checking for second language learners: the use of prepositions. *Proceedings of the 14th Nordic Conference in Computational Linguistics*, Reykjavik, Iceland.

Fathman, A. K., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), Second language writing: Research insights for the classroom (pp. 178-190). Cambridge, UK: Cambridge University Press.

Ferris, D. & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? Journal of Second Language Writing, 10, 161–184.

James, C. (1998). Errors in Language Learning and Use: Exploring Error Analysis. London: Addison Wesley Longman.

Lee, L. H., Yu, L. C., Lee, K. C., Tseng, Y. H., Chang, L. P., & Chen, H. H. (2014). A Sentence Judgment System for Grammatical Error Detection. *Proceedings of COLING 2014* (pp. 67-70), Dublin, Ireland.

Leacock, C., Chodorow, M., Gamon, M., & Tetreault, J. (2010). Automated Grammatical Error Detection for Language Learners. Morgan and Claypool Publishers.

Gamon, M., Leacock, C., Brockett, C., Dolan,W. B., Gao, J. F., Belenko, D., & Klementiev, A. (2009). Using Statistical Techniques and Web Search to Correct ESL Errors. *CALICO Journal, 26*(3), 491-511.

Ng, H. T., Wu, S. M., Wu, Y., Hadiwinoto, C., & Tetreault. J. (2013). The conll-2013 shared task on grammatical error correction. *Proceedings of the 17th Conference on Computational Natural Language Learning*.

Nicholls, D. (2003) The Cambridge learner corpus: error coding and analysis for lexicography and ELT. *Proceedings of the Corpus Linguistics 2003 Conference* (pp. 572-581), Lancaster, UK..

Tetreault, J., & Chodorow, M. (2009). Examining the use of region web counts for ESL error detection. *Proceedings of the Web as CorpusWorkshop (WAC-5)*, San Sebastian, Spain.

Tetreault, J., & Chodorow, M. (2008). The ups and downs of preposition error detection in ESL writing. *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 865-872), Manchester, UK.

Wang, Z. (2011). A Study on the Teaching of Unique Syntactic Pattern in Modern Chinese for Native English-Speaking Students. Master Thesis. Northeast Normal University.

Yu, C. H., & Chen, H. H. (2012). Detecting word ordering errors in Chinese sentences for learning Chinese as a foreign language. *Proceedings of COLING 2012* (pp. 3003-3018), Bombay, India.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Extracting a Chinese Learner Corpus from the Web: Grammatical Error Correction for Learning Chinese as a Foreign Language with Statistical Machine Translation

Yinchen ZHAO[a], Mamoru KOMACHI[a*] & Hiroshi ISHIKAWA[a]
[a]*Graduate School of System Design, Tokyo Metropolitan University, Japan*
*komachi@tmu.ac.jp

Abstract: In this paper, we describe the TMU system for the shared task of Grammatical Error Diagnosis for Learning Chinese as a Foreign Language (CFL) at NLP-TEA1. One of the main obstacles in grammatical error correction for CFL is a data bottleneck problem. The Chinese learner corpus at hand (NTNU learner corpus) contains only 1,208 sentences in total, which is obviously insufficient for supervised learning-based techniques. To overcome this problem, we extract a large-scale Chinese learner corpus from a language exchange site called Lang-8, which results in 95,706 sentences (two million words). We use it as a parallel corpus for a phrase-based statistical machine translation (SMT) system, which translates learner sentences into correct sentences.

Keywords: Chinese learner corpus, web mining, grammatical error correction, statistical machine translation

1. Introduction

Recently, educational applications of natural language processing techniques are actively studied. For example, grammatical error correction for English as a Second Language (ESL) learners has gained large attention in the past few years. Specifically, there were a number of shared tasks of grammatical error correction for ESL learners such as Helping Our Own (HOO) and Conference on Natural Language Learning (CoNLL). However, little attention has been paid to Chinese as a foreign language (CFL). One of the reasons why it is difficult to develop a grammatical error correction system for CFL is the lack of learner corpora. In this paper, we present a method to extract a learner corpus of Chinese from the web, and use it to build a grammatical error correction system for CFL. The main contributions of this paper is as follows:

1. To best of our knowledge, this is the first work that constructs a large-scale learner corpus of Chinese from the web. It contains 100,000 sentences (2M words) annotated with corrections.
2. It is the first work that adopts statistical machine translation (SMT) to grammatical error correction task for CFL. The experimental result shows that our proposed approach is effective to build a precise error correction system.
3. Unlike previous using phrase-based SMT for Chinese spelling correction task, we propose to use character-wise tokenization and prove that character-wise tokenization is more robust than word-wise tokenization.

2. Extracting a Chinese Learner Corpus from the Web

To alleviate the problem of shortage of training data, we resort to extract a Chinese learner corpus from the web. We focus on a language exchange social networking service (SNS) called Lang-8[1]. Lang-8 offers a wide variety of languages that you can use to write a blog entry. Other users correct your blog entry written in your learning language, and you in turn correct other users' blog entry written in your mother tongue. Lang-8 facilitates the process of mutual "language exchange". Up to date (August 2014), Lang-8 has about one million users where 50,000 of them are Chinese learners.

3. Grammatical Error Correction with Statistical Machine Translation

We decompose the task of grammatical error correction into two parts. First, we identify the location of errors using statistical machine translation trained on a Chinese learner corpus. Second, we classify the type of errors using a simple heuristic rule using dynamic programming.

3.1 Error Identification with Statistical Machine Translation

We follow (Brockett, Dolan, & Gamon, 2006) to make a grammatical error correction system with phrase-based statistical machine translation. One of the advantages of the approach is that we can use an off-the-shelf machine translation toolkit to build a grammatical error correction system if we have a learner corpus with sufficient size.

In their paper, the grammatical error correction process is modeled using a noisy-channel model as follows:

$$\hat{e} = \arg\max_e P(e \mid f) = \arg\max_e P(f \mid e) P(e)$$

where $P(e)$ is a language model and $P(f \mid e)$ is a translation model. In this paper, f corresponds to a learner sentence and e corresponds to a corrected sentence, respectively. The phrase-based SMT toolkit we use in this paper actually uses a log linear model which contains the noisy-channel model as follows:

$$\hat{e} = \arg\max_e \mathbf{w}^{\mathrm{T}}\mathbf{h}$$

where **w** is a weight vector and **h** is a feature function, respectively.

We propose two types SMT systems: word-based system and character-based system, depending on the pre-processing step of a learner corpus. The intuition behind using a character-wise segmentation is that learners of Chinese tend to write incorrect sentences, which may hurt the accuracy of the word segmentation. Character-based SMT is free from tokenization errors, while it is able to learn word-to-word or phrase-to-phrase correction patterns thanks to the phrase extraction heuristics.

3.2 Error Classification with Dynamic Programming

Once we identify the location of errors, we classify the type of errors using a simple heuristic rule. We use a dynamic programming algorithm to calculate the number of insertion, deletion and replacement operations for each sentence pair. We then classify the type of errors by the following pseudo-code:

[1] http://lang-8.com/

Table 1: Pseudo-code for error type classification.

```
Input: learner sentence l, system correction c
Output: error type t
---
(i, d, r) ← get_operations(l, c)
if  d > 0 and i > 0
  t ← "Disorder"
else if r > 0
  t ← "Selection"
else if d > 0
  t ← "Redundant"
else if i > 0
  t ← "Missing"
else
  t ← "correct"
end if
return t
```

If a sentence contains only one error, this algorithm correctly returns the "Disorder" error type, while it may fail to classify "Selection" error type and output "Redundant" or "Missing" error types. In a preliminary experiment, we found that this confusion can be negligible. We did not explore the use of machine learning-based classification method because the training corpus provided by the organizer contains only 1,000 instances.

4. Experiments

In this section, we describe the experimental settings and results for the NLP-TEA1.

4.1 Data and Tools

We obtained the Lang-8 Learner Corpora v2.0. The corpora come with "blog id", "sentence id", "learning language", "native language", "learner sentences" and "corrected sentences". We extracted blog entries whose "learning language" is set to "Mandarin". The Chinese portion of the Lang-8 Learner Corpora consists of 29,595 blog entries (441,670 sentences). We discarded following sentences and kept 95,706 sentences at last.

- Too long (more than or equal to 20 words) or too short (less than or equal to 3 words).
- Not written in Chinese.
- Any corrected sentence 1.3 times longer or more than the original one.[2]

We used Moses 2.1.1 as a statistical machine translation toolkit with its default parameter. The training and testing was done using the scripts distributed as KFTT Moses Baseline v1.4 (Neubig, 2011). We did not perform minimum error rate training (Och, 2003). We trained an SMT system with two training corpora: the Lang-8 Chinese Learner Corpus with and without word segmentation. In other words, we built a grammatical error correction

[2] Some corrected sentences contain comments and annotations, which may harm word alignment for SMT.

system trained on a character-based phrasal SMT in addition to a word-based phrasal SMT. We used jieba[3] 0.32 for Chinese text segmentation.

4.2 Results

Table 1 summarizes the false positive rate, accuracy, precision, recall and F1 scores for the formal run. Character-based approach outperformed word-based approach in all evaluation metrics. This confirms the hypothesis that word segmentation errors damage grammatical error correction for CFL.

We ranked the 2nd at the false positive rate and accuracy out of six groups participated in the shared task. However, these evaluation metrics alone cannot verify the effectiveness of our approach, since there is a trade-off between these metrics. Note that we only reports the scores at detection level, since the performance at identification level is almost the same.

Table 1: Experimental results for the formal run at NLP-TEA1. Accuracy, precision, recall and F1 scores are at the detection level.

	False Positive Rate	Accuracy	Precision	Recall	F1
TMU-Run1: Character-based	0.1977	0.5171	0.5399	0.2320	0.3245
TMU-Run2: Word-based	0.1691	0.5103	0.5287	0.1897	0.2792

5. Discussion

Our system achieved the worst (6/6) performance in terms of F1 score. The main reason is that we did not perform any parameter tuning at all, even though the error distribution of the test corpus is very skewed (half of the sentences contain errors). In a preliminary experiment, we ran the minimum error rate training using BLEU (Papineni, Roukos, Ward, & Zhu, 2002), but after the optimization the system outputs almost no corrections. This is because the BLEU score will become higher if the system does not change the learner sentence. Although BLEU is used to evaluate grammatical error correction as in (Park & Levy, 2011), it may not adequate to assess the quality of error correction systems. One possible direction is to optimize the SMT system using the F1 score with Z-MERT[4].

Note that the shared task only requires participants to determine whether a given sentence contains an error or not, our system is capable of locating the position of errors. In addition, our system can identify multiple errors in a sentence (although it is out of scope of this shared task).

One of the side effects of using the Lang-8 corpus is that the error correction system misclassifies correct sentences as "Missing" errors since it tends to use commas where applicable. However, commas often make more natural Chinese expressions than original. For instance, consider the following example. The system output is more fluent than the original, but it is different from the gold standard annotation, which deteriorates performance.

Gold: 今天的天氣很好不怎麼熱 （Today's weather is good, not very hot.）
System: 今天的天氣很好，不怎麼熱

[3] https://github.com/fxsjy/jieba
[4] http://cs.jhu.edu/~ozaidan/zmert/

Also, we would like to emphasize that we did not use any resources provided by the organizer. It is interesting to use domain adaptation approach such as in (Imamura, Saito, Sadamitsu & Nishikawa, 2012) to better reflect error distribution of the given domain (for example, 50% of the given test corpus contains errors, which is not often the case in realistic setting).

If a sentence contains more than one error, the proposed error type classification algorithm will output only one error type. Since the test corpus is controlled to contain only one error, we opted for a simple rule for the shared task. However, it is possible that these error types are not identical in real setting, so our future work includes error type classification for each error.

6. Related Work

Lang-8 is considered as one of the invaluable resources for knowledge acquisition for second language learners. For example, Japanese learner corpus (Mizumoto, Komachi, Nagata, & Matsumoto, 2011; Kasahara, Komachi, Nagata, & Matsumoto, 2011) and English learner corpus (Tajiri, Komachi, & Matsumoto, 2012; Mizumoto, Hayashibe, Komachi, Nagata, & Matsumoto, 2012) can be extracted from Lang-8. It is not surprizing that we can extract a large corpus of Chinese learners since Chinese (Mandarin) is the third most popular learning languages in Lang-8[5], followed by English and Japanese.

The use of statistical machine translation techniques to grammatical error correction was pioneered by (Brockett, Dolan, & Gamon, 2006), and has been adopted to many researchers in grammatical error correction for ESL (Mizumoto, Hayashibe, Komachi, Nagata, & Matsumoto, 2012; Buys & van der Merwe, 2013; Yuan & Felice, 2013; Behera & Bhattacharyya, 2013; Junczys-Dowmunt & Grundkiewicz, 2014).

Recently, similar approach is applied to Chinese spelling error correction as well (Wu, Liu & Lee, 2013; Wu, Chiu & Chang, 2013; Liu, Cheng, Luo, Duh & Matsumoto, 2013). However, all of these methods use word-based statistical machine translation, even though some of them use character n-gram language model. One of our proposed models investigates character-wise segmentation rather than word-wise one, and indicates that character-based model can learn useful correction patterns if the training corpus is sufficiently large.

Error type classification has gained much attention, for example in English (Swanson & Yamangil, 2012) and Japanese (Oyama, Komachi & Matsumoto, 2013). Although these works use linguistically motivated annotation scheme proposed in previous work, the error type annotation scheme for the NTNU learner corpus is based on edit operations and it is more appropriate to use rules rather than machine learning.

7. Conclusion

In this paper, we described the TMU system for the Grammatical Error Diagnosis for CFL Shared Task at NLP-TEA1. To increase the number of training corpus, we explored the web for constructing a learner corpus of Chinese. We extracted 100,000 learner sentences paired with their correction from the language exchange SNS, Lang-8, and used it to train an SMT-based grammatical error correction system. We compared two types of segmentation for phrasal SMT and found that character-based SMT outperforms word-based SMT for CFL grammatical error correction. The system achieved moderate performance even though it did not use any language resources from the target domain.

[5] http://cl.naist.jp/nldata/lang-8/

Acknowledgements

We would like to thank Xi Yangyang for granting use of extracted texts from Lang-8.

References

Behera, B. & Bhattacharyya, P. (2013). Automated Grammar Correction Using Hierarchical Phrase-Based Statistical Machine Translation. *Proceedings of the International Joint Conference on Natural Language Processing*, (pages 937-941). Nagoya, Japan.

Brockett, C., Dolan, B. W., & Gamon, M. (2006). Correcting ESL Errors Using Phrasal SMT Techniques. *Proceedings of the International Conference on Computational Linguistics and Annual Meeting of the ACL*, (pages 249-256). Sydney, Australia.

Buys, J. & van der Merwe, B. (2013). A Tree Transducer Model for Grammatical Error Correction. *Proceedings of the Seventeenth Conference on Computational Natural Language Learning: Shared Task*, (pages 43–51). Sofia, Bulgaria.

Imamura, K., Saito, K., Sadamitsu, K. & Nishikawa, H. (2012). Grammatical Error Correction Using Pseudo-Error Sentences and Domain Adaptation. *Proceedings of the Annual Meeting of the ACL*, (pages 388-392). Jeju Island, Korea.

Junczys-Dowmunt, M. & Grundkiewicz, R. (2014). The AMU System in the CoNLL-2014 Shared Task: Grammatical Error Correction by Data-Intensive and Feature-Rich Statistical Machine Translation. *Proceedings of the Eighteenth Conference on Computational Natural Language Learning: Shared Task (CoNLL-2014 Shared Task)*, (pages 25-33). Baltimore, USA.

Kasahara, S., Komachi, M., Nagata, M., & Matsumoto, Y. (2011). Correcting Romaji-Kana Conversion for Japanese Language Education. *Proceedings of the Workshop on Text Input Methods*, (pages 38-42). Chiang Mai, Thailand.

Liu, X., Cheng, F., Luo, Y., Duh, K., & Matsumoto, Y. (2013). A Hybrid Chinese Spelling Correction Using Language Model and Statistical Machine Translation with Reranking. *Proceedings of the Seventh SIGHAN Workshop on Chinese Language Processing (SIGHAN-7)*, (pages 54–58), Nagoya, Japan.

Mizumoto, T., Hayashibe, Y., Komachi, M., Nagata, M., & Matsumoto, Y. (2012). The Effect of Learner Corpus Size in Grammatical Error Correction of ESL Writings. *Proceedings of the International Conference on Computational Linguistics*, (pages 863-872). Mumbai, India.

Mizumoto, T., Komachi, M., Nagata, M., & Matsumoto, Y. (2011). Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners. *Proceedings of the International Joint Conference on Natural Language Processing*, (pages 147-155). Chiang Mai, Thailand.

Neubig, G. (2011). The Kyoto Free Translation Task. http://www.phontron.com/kftt.

Och, J. F. (2003). Minimum Error Rate Training in Statistical Machine Translation. *Proceedings of the Annual Meeting of the ACL*, (pages 160-167). Sapporo, Japan.

Oyama, H., Komachi, M., & Matsumoto, Y. (2013). Towards Automatic Error Type Classification of Japanese Language Learners' Writing. *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation*, (pages 163-172). Taipei, Taiwan.

Papineni, K., Roukos S., Ward, T., & Zhu, W.-J. (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. *Proceedings of the Annual Meeting of the ACL*, (pages 311-318). Philadelphia, USA.

Park, Y. A., & Levy, R. (2011). Automated Whole Sentence Grammar Correction Using a Noisy Channel Model. *Proceedings of the Annual Meeting of the ACL*, (pages 934-944). Ohio, USA.

Swanson, B., & Yamangil, E. (2012). Correction Detection and Error Type Selection as an ESL Educational Aid. *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (pages 357-361). Montreal, Canada.

Tajiri, T., Komachi, M., & Matsumoto, Y. (2012). Tense and Aspect Error Correction for ESL Learners Using Global Context. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, (pages 198-202). Jeju Island, Korea.

Wu, J.-C., Chiu, H.-W., & Chang, J. S. (2013). Integrating Dictionary and Web N-grams for Chinese Spell Checking. *Computational Linguistics and Chinese Language Processing*, 18(4), (pages 17-30).

Wu, S.-H., Liu, C.-L., & Lee, L.-H. (2013). Chinese Spelling Check Evaluation at SIGHAN Bake-off 2013. *Proceedings of the SIGHAN Workshop on Chinese Language Processing*, (pages 35-42). Nagoya, Japan.

Yuan, Z. & Felice, M. (2013). Constrained Grammatical Error Correction Using Statistical Machine Translation. *Proceedings of the Seventeenth Conference on Computational Natural Language Learning: Shared Task*, (pages 52–61). Sofia, Bulgaria.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Detecting Grammatical Error in Chinese Sentence for Foreign

Jui-Feng Yeh, Yun-Yun Lu, Chen-Hsien Lee, Yu-Hsiang Yu, Yong-Ting Chen
Department of Computer Science and Information Engineering, National Chiayi University
No.300 Syuefu Rd., Chiayi City 60004, Taiwan (R.O.C.)
{ralph, s1020447, s1020443, s1002967, s1003008}@mail.ncyu.edu.tw

Abstract: Each language has its own unique grammar, Chinese is same as the other languages similarly. But each language has different type even there does not have any relations. So the foreigners learn the language not only need to learning the word pronouns and glyph, but also need to learn the grammar. This issue is very extensive, not only can help foreigners to learn Chinese, but also can detect the error grammar. This paper had proposed method can divide five sections of the structure: First sections are input sentence; second sections are parsed and word segmentation; third sections find the missing, wrong word; fourth sections find the redundant wrong word; fifth sections are final output. This paper has two parts, the first is how to detect the grammar error, and the second is how should we know the Chinese grammar error is what type. Finally, we can get the type of the grammar, and we can know how to correct.

Keywords: CFL, Chinese word correction, grammatical error, rule induction.

1. Introduction

Learning Chinese is more famous than before, not only there has more and more Chinese business, but also there has many tourist attractions csuse the foreigners love to go there to travel. So, there have many of foreigners beginning to learn Chinese. But Chinese is not easy to learn like the other languages, it is not only having many pronunciations and glyph of the word, but also have many grammar of the sentence. The Chinese although has subject, verbs and object too, but there has a combination of fixed text, if you do not to comply with these rules, the meaning of the sentence itself will be different.So how to learn Chinese grammar is very import research. This topic is extensive, not only for foreigners to learn Chinese, but also can help to detect the wrong grammar in the document.

In recent years, there has a lot of paper to research about Chinese learning grammar. Most of paper about learning Chinese paper not only talk about the sentence correct rate, but also talk about the Grammar correct rate. Ying Jiang (2012) proposes an arithmetic called “Rnture-Sentence” to segment the Chinese sentences, it can segment the sentence more completely, and solve the problem about the complicated Chinese grammar, proofreading method, their method to cross sentences is based on LanguageTool, their paper also presents to some method of new rule which can accomplish complicated Chinese grammar proofreading. These authors also propose another Chinese grammar, proofreading (2013), the presents an indexing method of a corpus base of the Chinese grammar, they can evaluate the rule of the accuracy and the frequency, each rule, they adopt an iterative approach to improve it, make sure its better performance in the real word, it also introduces the important role of the corpus of their method. Mei-Jen Audrey Shih et al. (2011) propose a Chinese online learning system, this online system is convenience and it is assembled to abound environment and had a broad content search opportunity, this paper is mained on how to learn Chinese language effectively in an online learning environment. Lee Jo Kim et al. (2011) propose a Chinese language teaching and learning system based on ICT-Base tool, this tool can help peer assisted learning environment. Ying Jiang et al. (2012) they provide a new method to deal with grammar error. They arrange a new rule for their new grammar system, this system has two combined characters with intention, then their grammar system can contain as well as the spelling error system, so their system can have perfect respect of precision and practicality. David Tawei Ku et al. (2012) proposes a situated learning for

Chinese learning, it is a trend that ubiquitous learning environment, and the feature main on real life learning situation, and problem solving practice, this learning system has two parts, one is integrating situated learning strategy and the other is context awareness technology. Yanwei Wang et al. (2011) proposes a discriminative learning method of MQDF (Modified Quadratic Discriminant Function), MQDF is based on sample importance weights, this method is investigated and compared other discriminative learning methods about MQDF. Lung-Hsiang Wong (2010) propose a Mobile-Assisted Language Learning (MALL), there have two case studies, and mained on "creative learner outputs", student in two studies language by one-to-one mobile devices, and capture the picture of the real life. Hui Yang et al. (2010) proposes a continuous prior polarity algorithm, their method reflects subtle changes of sentiment in contrast, its previous studies which expressed sentiment polarity discretely, they also proposed a method based on Chinese dependency grammar which can assess modified polarity, they can accurately identify subjective words and its modified according different Chinese dependency grammar, then predict the sentence by aggregate. Peng Li et al. (2012) proposes "A Hierarchy-based Constraint Dependency Grammar Parsing for Chinese", they mentioned the Constraint Dependency Grammar (CDG) is a famous formalism which about the grammatical rules, and they have successfully adopted in Chinese, they propose to develop a three schema which is based on a study of constraining in the corpus. Haiping Zhu et al. (2011) propose a analyze Chinese sentence with semantic dependency method, the correlation between words and phrases can calculation of the similarity from Hownet, between the sentence and sentence's similarity can analyze by formula, their method can be adopted to analyze the correct topic and to categorize.

This paper had proposed method can divide five sections of structure: First sections are input sentence; second sections are parsing and word segmentation; third sections are find the missing wrong word; forth sections are find the redundant wrong word; fifth sections are final output.In the third section, we classify the four of the grammar rule, the part of speech(POS) can classify four type(Shi, Neu, D, DA), and there have regular POS behind the four type of the POS.After find, we put the worng grammar part in the dictionary file which only for these error, the dictionary file name is Miss . In the forth section, find the redundant wrong word, there has a special rule,behind the POS of DE must a DE, if did not, it will write in the dictionary file which name is "Redundant". finally, we will use these dictionary files to detect the Chinese Grammar, if the error is bellowing "Missing" diIionary file, the error is bellowing "Redundant"dictionary file, if not miss error or redundant error, that means it it correct, output to correct the file.

2. Method

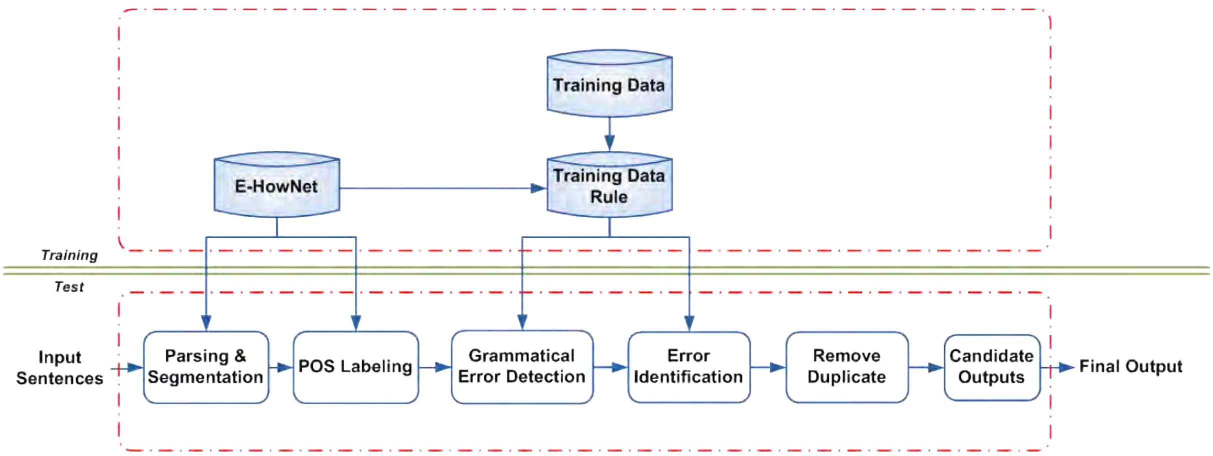


Figure 1. The framework of the proposed system.

In this section, we will introduce the framework of the proposed system and method. Our proposed method is aimed to detect and identify the sentence for learning Chinese as a foreign language (CFL). The sentences written by CFL may contain a variety of grammatical errors, such as word choice, missing words, word disorder and so on. It focuses on grammatical errors in this task. And the framework is divided into two parts: training phase and test phase that will describe in section 2.1 and section 2.2.

2.1 Training phase

The training phase shown in figure 1, we have some data that are contain some sentences can be trained to find some useful information. First, we do the pre-process to the data from the task organizer, that will be input file and we removed unneeded portions of each sentence in this file, such as SID number, then the treated results will be further inputted into the tool which is CKIP AutoTag, that is to do the word segmentation and part-of-speech (POS) tagging based on E-Hownet. The corresponding part-of-speech of each word is obtained in the sentences, which is given a part of speech at the end of a word in parentheses. Second, we are going to remove unessential blank spaces and parentheses, that will be more convenient in the following file operations. In the test phase, we are also adopted in this way. Then, we want to find some rules with training data which can be used in test phase. We construct the training data rule from the results of process which have part-of-speech. Finally, the candidate outputs are generated according to our training data rule.

2.2 Test phase

In the previous section, the training data rule is built in training phase. We will describe the test phase of the framework in this section. The word segmentation and part of speech (POS) labeling are the same as training phase. Then, we begin the processes with the third & fourth step, we have to detect and identify the wrong word. The following is focused on finding the Missing type of the wrong words.

- Behind the word with POS of “Shi” is not connected the word with POS of “Verb”.
- Behind the word with POS of “Neu” is not connected the word with POS of “De”.
- Behind the word with POS of “D” is not connected the word with POS of “Neqa”.
- Behind the word with POS of “Da” is not connected the word with POS of “Neu”.

According to above, the Missing type of the incorrect words will save in the file named missing. And the following is focused on finding the Redundant type of the wrong words.

- Behind the word with POS of “De” is not connected the word with POS of “De”.
- Behind the word with POS of “P” is not connected the word with POS of “P”.
- Behind the word with POS of “Cbb” is not connected the word with POS of “D”.
- Behind the word with POS of “Vh” is not connected the word with POS of “D”.
- Behind the word with POS of “De” is not connected the word with POS of “Neqa”.
- Behind the word with POS of “D” is not connected the word with POS of “D”.

According to above, the Missing type of the incorrect words will save in the file named missing. We will remove repeated SID. It is helping us to reduce the process time. We will output the result in the final step. The processes will run the test data, if the word If the word exists in the missing file, we will output the sentence with Missing. The Redundant type of the incorrect word is same as Missing. Others are identified as correct. For example, input: “(sid=C1-1876-2) 對社會國家不同的影響”, output: “C1-1876-2, Missing”, If the input contains no errors, the system should return “C1-1876-2, correct”.

3. Experiments

According to the grammatical error diagnosis for learning Chinese as a foreign language in NLP-TEA-1, this paper is dedicated to the detection and identification of errors in sentences. The evaluate is divided into two parts: Subtask 1 is detection level that is to check out the sentence which is incorrect or correct, then the subtask 2 is identification level, which is to identify the error type in sentences, i.e., Redundant, Missing, Disorder, and Selection. In section 3.1, we will describe the data sets, performance metrics, then we will show our evaluation in section 3.2.

3.1 Data sets

```
<ESSAY title="張愛文的一天">
<TEXT>
<SENTENCE id="A2-0101-1">下了課就去看他在學校對面的公車站</SENTENCE>
</TEXT>
<MISTAKE id="A2-0101-1">
<TYPE>Disorder</TYPE>
<CORRECTION>下了課就去在學校對面的公車站看他</CORRECTION>
</MISTAKE>
</ESSAY>
```

Figure 2. An example of the training data.

In this task, the evaluation is an open test. Participants can employ any linguistic and computational resources to develop the error diagnosis, and provide data of CFL's essays from the NTNU learner corpus for training purpose. The corpus was released in SGML format which is shown in figure 2. Moreover, there are at least 1000 different degrees of difficulty of testing passages for testing. In this paper, we use C++ to develop our proposed method.

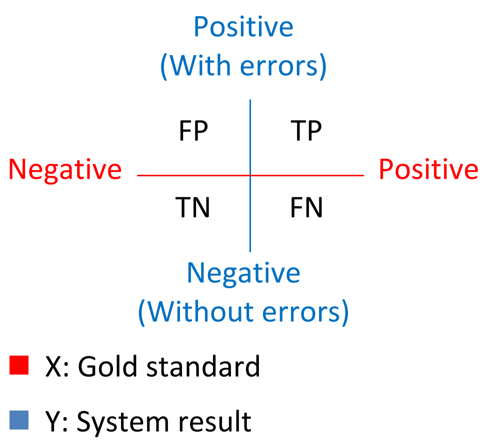


Figure 3. A quadrant map of performance metrics.

The judging correctness are divided into two parts: detection level and identification level. The following are showing some performance metrics and quadrant map shown in figure 3 that is measured in both levels of indicators:

- **TP**: System determines the character for errors related to the actual error, and the judgments the system is correct.
- **FP**: System determines the character for errors is not related to the actual error, and the judgments of the system is incorrect.
- **FN**: System determines the character for errors is related to the actual error, and the judgments of the system is incorrect.
- **TN**: System determines the character for errors is not related to the actual error, and the judgments of the system is correct.

The following of performance metrics are according to the quadrant map.

- **False Positive Rate** $= \frac{\boldsymbol{FP}}{(\boldsymbol{FP}+\boldsymbol{TN})}$

- $\mathbf{Accuracy} = \frac{(\boldsymbol{TP}+\boldsymbol{TN})}{(\boldsymbol{TP}+\boldsymbol{TN}+\boldsymbol{FP}+\boldsymbol{FN})}$
- $\mathbf{Precision} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP}+\boldsymbol{FP})}$
- $\mathbf{Recall} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP}+\boldsymbol{FN})}$
- $\mathbf{F1-Score} = \frac{2\times \boldsymbol{Precision}\times \boldsymbol{Recall}}{(\boldsymbol{Precision}+\boldsymbol{Recall})}$

3.2 Evaluation

According to the table 1, our false positive rate is the best in this task, which means that our proposed method is feasible, but our proposed method just focuses on identifying two error type, . There are two parts of performance evaluation: detection level and identification level which is shown in table 2 and table 3. In the identification, we can see that accuracy is the best. Then, accuracy and precision are also comparable to others, but our method in recall is relatively weaker than another. This performance evaluation shows that our method is viable, but our method is still much room for improvement.

Table 1: Participating teams of the false positive rate.

Participating teams	False Positive Rate
NCYU*	0.1189
TMU	0.1691
UDS	0.2491
KUAS&NTNU	0.2686
CIRU	0.496
NTOU	1

Table 2. Participating teams of performance evaluation in Detection Level.

Participating teams	Accuracy	Precision	Recall	F1
NCYU*	0.4983	0.4927	0.1154	0.187
TMU	0.5171	0.5399	0.232	0.3245
UDS	0.4914	0.4945	0.7749	0.6037
KUAS&NTNU	0.5006	0.5003	0.9051	0.6444
CIRU	0.6446	0.6128	0.7851	0.6884
NTOU	0.5	0.5	1	0.6667

Table 3. Participating teams of performance evaluation in Identification Level.

Participating teams	Accuracy	Precision	Recall	F1
NCYU*	0.4594	0.2409	0.0377	0.0652
TMU	0.4554	0.3545	0.1086	0.1662
UDS	0.2337	0.2467	0.2594	0.2529
KUAS&NTNU	0.2149	0.2696	0.3337	0.2983
CIRU	0.4589	0.4548	0.4137	0.4333
NTOU	0.2074	0.2932	0.4149	0.3436

4. Conclusions

This study proposes a method for Chinese text detect grammar error. The method in our study is focus on word classify to easy detect Chinese grammar error. The grammar error is classifying four type, the verbs was not add behind POS of Shi, the De was not add behind the POS of Neu, the Neqa was not add behind POS of D, and the Neu was not add behind POS of Da. The experimental result shows the performance it good, and we also apply this method in “grammatical error diagnosis for learning Chinese as a foreign language”, and the final result pretty good. In the feature, we hope can raise the performance and find the more grammar type . More grammar type can helpful to find the Chinese grammar error. After the Chinese grammar error, we will start to study the relationship between grammar and spelling errors, because in this paper we only care about the word pronouns and glyph, but in recent years some spelling error has been regularization, it most to understanding the context then detect it is right or wrong, so the issue about the relationship between grammar and spelling errors is need to study, if we can fine the relationship then the Chinese grammar detect correct rate must can raise higher.

Acknowledgements

This work is supported in part by the National Science Council, Taiwan, R.O.C., under the project grant numbers NSC 102-2221-E-415-006-MY3.

References

Qiu, X., Jia, W., and Li, H. 2012. A Font Style Learning and Transferring Method Based on Strokes and Structure of Chinese Characters. In Computer Science and Service System (CSSS) on pp. 1836-1839.

Syson, M. B., Estuar, M. R. E., and See, K. T. 2012. ABKD: Multimodal Mobile Language Game for Collaborative Learning of Chinese Hanzi and Japanese Kanji Characters. In Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 03 pp. 311-315.

Tam, V., and Cheung, R. L. 2012. An Extendible and Ubiquitious E-learning Software for Foreigners to Learn Chinese on iOS-Based Devices. InAdvanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on pp. 46-48.

Tam, V., and Huang, C. 2011. An Extendible Software for Learning to Write Chinese Characters in Correct Stroke Sequences on Smartphones. InAdvanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on pp. 118-119.

Li, K. H., Cheng, T. F., Lou, S. J., and Tsai, H. Y. 2012. Application of Game-based Learning (GBL) on Chinese language learning in elementary school. In Digital Game and Intelligent Toy Enhanced Learning (DIGITEL), 2012 IEEE Fourth International Conference on pp. 226-230.

Ku, D. T., and Chang, C. C. 2012. Development of context awareness learning system for elementary Chinese language learning. In Proceedings of the 2012 Sixth International Conference on Genetic and Evolutionary Computing on pp. 538-541.

Tam, V., and Luo, N. 2012. Exploring Chinese through learning objects and interactive interface on mobile devices. In Teaching, Assessment and Learning for Engineering (TALE), 2012 IEEE International Conference on pp. H3C-7.

Shih, M. J., and Yang, J. C. 2011. How to Learn Chinese through Online Tools? From the Perspective of Informal Learning to Culture Immersion. InAdvanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on pp. 305-306.

Kim, L. J., Lim, S. H., and Ying, L. T. 2011. ICT-based peer assisted learning environment: Using online feedback tools for Chinese Language writing tasks. In Electrical and Control Engineering (ICECE), 2011 International Conference on pp. 6612-6614.

Wong, L. H., and Looi, C. K. (2010, April). Mobile-assisted vocabulary learning in real-life setting for primary school students: Two case studies. In Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE), 2010 6th IEEE International Conference on pp. 88-95.

Chuang, S. J., Zeng, S. R., and Chou, Y. L. 2011. Neural Networks for the Recognition of Traditional Chinese Handwriting. In Computational Science and Engineering (CSE), 2011 IEEE 14th International Conference on pp. 645-648

Wu, Y., Yuan, Z., Zhou, D., and Cai, Y. 2013. Research of virtual Chinese calligraphic learning. In Multimedia and Expo (ICME), 2013 IEEE International Conference on pp. 1-5.

Ku, D. T., and Chang, C. C. 2012. Development of context awareness learning system for elementary Chinese language learning. In Proceedings of the 2012 Sixth International Conference on Genetic and Evolutionary Computing on pp. 538-541.

Zhao, Z., and Ma, X. 2012. Prediction of Prosodic Word Boundaries in Chinese TTS Based on Maximum Entropy Markov Model and Transformation Based Learning. In Computational Intelligence and Security (CIS), 2012 Eighth International Conference on pp. 258-261.

Lin, C. C., and Tsai, R. H. 2012. A Generative Data Augmentation Model for Enhancing Chinese Dialect Pronunciation Prediction. Audio, Speech, and Language Processing, Transactions on, 20(4), pp. 1109-1117.

Wang, Y., Ding, X., and Liu, C. 2011. MQDF discriminative learning based offline handwritten Chinese character recognition. In Document Analysis and Recognition (ICDAR), 2011 International Conference on pp. 1100-1104.

Shao, Y., Wang, C., Xiao, B., Zhang, R., and Zhang, Y. 2011. Multiple instance learning based method for similar handwritten Chinese characters discrimination. In Document Analysis and Recognition (ICDAR), 2011 International Conference on pp. 1002-1006.

Li, P., Liao, L., and Li, X. 2012 . A hierarchy-based constraint dependency grammar parsing for Chinese. In Audio, Language and Image Processing (ICALIP), 2012 International Conference on pp. 328-332.

Jiang, Y., Wang, T., Lin, T., Wang, F., Cheng, W., Liu, X., ... and Zhang, W. 2012,. A rule based Chinese spelling and grammar detection system utility. In System Science and Engineering (ICSSE), 2012 International Conference on pp. 437-440.

Jiang, Y., Zhou, Z., Wan, L., Li, M., Zhao, W., Jing, M., and Liu, X. 2012. Cross sentence oriented complicated Chinese grammar proofreading method and practice. In Information Management, Innovation Management and Industrial Engineering (ICIII), 2012 International Conference on Vol. 3, pp. 254-258.

Jiang, Y., Lin, Z., Wang, J., Dai, M., Zhen, L., Li, N., ... and Meng, Y. 2013. Corpus Based Chinese Grammar Error Detection Rules Evaluation Method and System. In Proceedings of the 2013 Third International Conference on Intelligent System Design and Engineering Applications on pp. 496-499.

Zhu, H., Yang, Y., Chen, Y., and Ma, Q. 2011, September. Chinese sentence correlation analyzing based on semantic dependency method. In Electronics, Communications and Control (ICECC), 2011 International Conference on pp. 1932-1935.

Ma, L. L., and Wu, J. 2012, June. On-line handwritten Chinese character recognition based on inter-radical stochastic context-free grammar. In Neural Networks (IJCNN), The 2012 International Joint Conference on pp. 1-5.

Wu, C. H., Liu, C. H., Harris, M., & Yu, L. C. (2010). Sentence correction incorporating relative position and parse template language models. Audio, Speech, and Language Processing, IEEE Transactions on, 18(6), 1170-1181.

Chang, R. Y., Wu, C. H., & Prasetyo, P. K. (2012). Error Diagnosis of Chinese Sentences Using Inductive Learning Algorithm and Decomposition-Based Testing Mechanism. ACM Transactions on Asian Language Information Processing (TALIP), 11(1), 3.

Yu, C. H., & Chen, H. H. (2012). Detecting Word Ordering Errors in Chinese Sentences for Learning Chinese as a Foreign Language. In COLING on pp. 3003-3018.

Academia Sinica CKIP. http://ckipsvr.iis.sinica.edu.tw/

Academia Sinica E-HowNet. http://ehownet.iis.sinica.edu.tw/

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Grammatical Error Detection with Limited Training Data: The Case of Chinese

Marcos ZAMPIERI, Liling TAN
Saarland University, Germany

Abstract: In this paper, we describe the UDS submission to the shared task on Grammatical Error Diagnosis for Learning Chinese as a Foreign Language. We designed four different experiments (runs) to approach this task. All of them are variations of a frequency-based approach using a journalistic corpus as standard corpus and comparing n-gram frequency lists to both the training and the test corpus provided by the shared task organizers. The assumption behind this approach is that comparing a standard reference corpus to a non-standard study corpus using frequency-based methods levels out non-standard features present in the study corpus. These features are very likely to be, in the case of this corpus, grammatical errors. Our system obtained 60.3% f-measure at the error detection level and 25.3% f-measure at the error diagnosis level.

Keywords: grammatical error detection, Chinese, error diagnosis, learner corpora

1. Introduction

Grammatical error detection and correction is a vibrant research area in NLP. In the last couple of years much effort has been concentrated on the detection, diagnosis and correction of errors in texts written both by native speakers and by foreign language learners. For foreign language learning the practical applications of grammatical error detection are manifold, ranging from spelling and grammar checkers to essay scoring and grading.

Given this interest, a number of shared tasks have been organized in recent years. This includes the HOO 2012 preposition and determiner error correction shared task (Dale et al., 2012) held at the 2012 edition of the BEA Workshop and the Grammatical Error Correction shared tasks held at CoNLL-2013 (Ng et al., 2013) and one year later at CoNLL-2014 (Ng et al., 2014).

Similar to the previous shared tasks, this year's Grammatical Error Diagnosis for Learning Chinese as a Foreign Language provided us the opportunity to explore computational methods on diagnosis of errors committed by foreign learners of Mandarin Chinese. The shared task was designed to evaluate systems' output in two stages:

1) Error detection level: identify whether a sentence contains an error or not.
2) Error diagnosis level: classification of errors types (redundant words, missing words, word order and word selection).

Participants were required to train their systems not only to identify errors, but also to classify error types making the task more challenging. As an example, a system trained for Chinese error diagnosis was recently presented by Lee et al. (2014) obtaining 68.9% F1 score.

Apart from the difficulty of the task itself, it is important to note that the computational processing of logographic languages such as Chinese poses several difficulties to researchers used to handling character-based languages. Trivial pre-processing tasks like tokenization and segmentation are much more challenging for Chinese than for example, for English. This issue will be discussed in more detail in this paper.

In the next sections we describe the UDS submissions to the shared task commenting on the results obtained and on the strengths and weaknesses of our approach. In our submissions we used a frequency-based approach using a reference corpus to compensate the small amount of training data available.

2. Related Work

Grammatical error correction and detection has been the subject of a number of research papers in recent years. Shared tasks such as the aforementioned Grammatical Error Correction at CoNLL-2013 (Ng et al., 2013) and CoNLL-2014 (Ng et al., 2014) have been organized to evaluate systems' performance in correcting errors in learner corpora.

Tetreault and Chodorow (2008) presented experiments for detecting preposition errors in English texts written by non-native speakers. The authors report 84% precision and 19% recall. Heilman et al. (2012) proposed a hybrid error correction approach to the HOO 2012 shared task (Dale et al., 2012) focusing on increasing recall and F-measure scores. The authors argue that most systems take only precision into account due to the high cost of false positives (e.g. labeling grammatical sentences as ungrammatical).

More recently, Yuan and Felice (2013) proposed the use of phrase-based statistical machine translation to grammatical error correction. The application of SMT techniques to the task is not new (Brockett et al., 2006) and the performance achieved by their approach is not particularly high. However, in Yuan and Felice (2013), authors contribute in other ways, as for example, in exploring methods of generating new artificial errors to increase the size of the dataset and therefore providing more training material. The generation of artificial errors has been the subject of other research papers such as in Foster and Anderson (2009) and Felice and Yuan (2014).

As for Chinese, Yu and Chen (2012) investigated the problem of word ordering errors in Chinese texts written by Chinese foreign language learners. Authors report 71.64% accuracy using word n-grams and POS tags. Chang et al. (2012) presented a rule-based learning algorithm (first order inductive learner (FOIL)) combined with a log-likelihood function to identify error types in Chinese texts.

In this section we briefly discussed a couple of recent papers that deal with error detection, correction and diagnosis. For a comprehensive overview about the topic see Leacock et al. (2014).

3. Methods

Given the task description presented in section 1, we received a training corpus from the organizers containing over 12,000 labeled instances in XML format. The corpus was annotated with a unique identifier for each sentence 'sentence id', the type of mistake that each sentence contained and its respective correction. A snapshot of the corpus provided by the organizers can be seen next:

```
<ESSAY title="寫給即將初次見面的筆友的一封信">
   <TEXT>
      <SENTENCE id="B1-0112-1">我的計畫是十點早上在古亭捷運站</SENTENCE>
      <SENTENCE id="B1-0112-2">頭會戴著藍色的帽子</SENTENCE>
   </TEXT>
      <MISTAKE id="B1-0112-1">
            <TYPE>Disorder</TYPE>
            <CORRECTION>我的計畫是早上十點在古亭捷運站</CORRECTION>
      </MISTAKE>
      <MISTAKE id="B1-0112-2">
            <TYPE>Missing</TYPE>
            <CORRECTION>頭上會戴著藍色的帽子</CORRECTION>
      </MISTAKE>
 </ESSAY>
```

In our preliminary experiments we observed that the corpus provided was not sufficiently large to build robust machine learning models for grammatical error detection or diagnosis. In a similar text classification shared task using learner corpora (Tetreault et al., 2013), the amount of training data was significantly larger than the test data which allowed researchers to build more robust models based only on the given training data.

In addition to that, we had a couple of problems with the Chinese segmentation tool that we used (Chang et al. 2008) and this returned us fewer segments than were actually in the training corpus. We unfortunately did not have enough time to perform error analysis on the segmentation and pre-processing tools available nor did we have time to use the most recent Chinese segmenters (Tan and Bond, 2014; Wang et al., 2014) before the shared task submission deadline. Given these difficulties, we had to search for new strategies to approach the task with limited training data that could still achieve results comparable to the state-of-the-art systems. Inspired by existing related work, we considered three alternatives to approach the task.

a) Use an external Mandarin Chinese as a foreign language corpus preferably containing similar tags to those of the training and test data.
b) Generate a list of artificial errors to increase the amount of instances in the training corpus as in Felice and Yuan (2014).
c) Use a frequency-based approach to compare the learner corpus to a standard general language corpus . The assumption is that this comparison would level out non-standard features of the training/test data that are more likely to be errors.

Given the time and resources we had, we decided to go with option (c) and leave the other two for future work. Option (a) seemed to be promising and straightforward in terms of performance, but we did not have suitable training data at our disposal. Acquiring and annotating new data is expensive and time consuming which made option (a) infeasible. As to option (b) we regard it to be a suitable and interesting alternative in cases where training data is not available. However, it is not currently possible to say much about the performance of these methods for Chinese. To our knowledge, previous work has only been done for Indo-European languages.

Option (c) proved to be the most adequate solution for our submission. A frequency-based approach, like the one used in our submission, requires only a large reference corpus (a general standard contemporary language one). We had a couple of suitable resources at our disposal (Chen et al., 1996 Graff and Chen, 2003) and we therefore decided to test this method.

The method works under a similar assumption to the keyword lists widely used in corpus linguistics (Scott, 1997; McEnery, 2009) and also applied on a similar scenario by Zampieri et al. (2013) on Internet data. Keyword lists are produced by comparing two corpora (a study corpus and a reference corpus) using association metrics such as log-likelihood, chi-square or mutual information. These keywords usually reflect salient features of the study corpus. In the case of the present comparison (standard corpus versus learner corpus), it is safe to assume that a reasonable amount of salient features from the learner corpus will be infrequent distributions of words which are very likely to be errors. This is the basic assumption of our approach.

3.1 Algorithm

If one assumes that a reference corpus is a portrait of standard language, lexical items that stand out in the study corpus in comparison to the reference corpus should deviate from what is considered to be 'the norm'. This is a relatively naive assumption and thematic bias may still occur when using unbalanced data. To avoid that, we used a large balanced journalistic corpus (Graff and Chen, 2003) as our standard corpus. From the reference corpus we sampled the first 50,000 sentences and extracted n-grams (1 to 5) using the KenLM Language Model Toolkit (Heafield, 2011).

We pre-processed the training, test and standard corpora using the Stanford tokenizer (Chang et al. 2008). As Chinese is a logographic language we treat every character in isolation. As previously mentioned, the Stanford segmenter yielded a number of errors in segmentation that worsened our system's performance. However, we were not able to evaluate the exact segmenter's performance for our dataset before this submission.

From the training and test corpora provided by the shared task organizers we proceeded to extract a list of ungrammatical n-grams that were not present in the subset of the reference corpus and treated them as key expressions. This calculation returned us a list of 35,000 ungrammatical n-grams not present in the reference corpus.

It is important to note here that the main difference betweem our approach and what is commonly used in corpus linguistics is that the latter uses the lexicon in the form of bag-of-words (or less often bigrams). In these experiments we used the complete set of n-grams (1 to 5) extracted from the corpus thus increasing the coverage of our method.

With these n-gram lists, we trained two classifiers to identify grammatical and ungrammatical instances: 1) a simple n-gram-based classifier to identify correct (grammatical) sentences using the formula below and 2) a Multinomial Naive Bayes (MNB) classifier to identify ungrammatical sentences along with their labels using the Scikit-learn package (Pedregosa et al., 2011).

$$\frac{\sum_{i}^{n} p(ng_i \mid ng_i \notin ng_{giga} \ , \ \dot{ng_i} \in ng_{traintest})}{\sum_{i}^{n} p(ng_i \mid ng_i \in ng_{giga})} > X \tag{1}$$

In the formula above, we tuned the *X* parameter value to optimize the results obtained by the first classifier. After a number of tests we found that the optimal value lies between 0.10 and 0.20. We therefore produced four submissions (runs) using different *X* values: 0.20 for the 1st run, 0.16 for the 2nd run, 0.15 for the 3rd run and 0.10 for the 4th run. The best results were obtained in our first run, using *X* = 0.20 and these are the results that will be reported and discussed next.

4. Results

According to the information provided by the organizers, 13 teams registered for the shared task and 6 of them submitted their final results. The results were calculated using standard metrics in text classification, namely: precision, recall, accuracy and F-measure as well as a false positive rate score. No limitation regarding the number of runs was set. The test set provided by the organizers contained 1,750 unlabeled test instances.

The UDS team submitted four runs changing the *X* parameter of our correct sentence classifier as explained in the previous section. In table 1 we present the best results obtained by all 6 groups at the error detection level. At this level, our approach was the fourth best with results reaching 60.37% F1 score and 49.14% accuracy.

Table 1: Error Detection Level: Results.

Team	Accuracy	Precision	Recall	F1
CIRU	0.6446	0.6128	0.7851	0.6884
NTOU	0.5000	0.5000	1	0.6667
KUAS&NTNU	0.5006	0.5003	0.9051	0.6444
UDS	**0.4914**	**0.4945**	**0.7749**	**0.6037**
TMU	0.5171	0.5399	0.232	0.3245
NCYU	0.4983	0.4927	0.1154	0.187

The top four systems obtained F1 scores between 60% and 69%; the 5th and 6th best system, however, obtained significantly lower F-scores. Our results were lower than the 3 best systems but still above the expect 50% baseline. In terms of recall, our system was also ranked as the 4th best and as to the accuracy results, our system was the 5th best. It obtained performance comparable to the 2nd, 3rd and 4th best systems: 49.14% against 50.00%, 50.06% and 51.71% accuracy. The best system obtained significantly higher accuracy scores compared to all other systems, 64.46% accuracy.

In table 2 we present the best results obtained by the six systems at the error diagnosis level.

Table 2: Error Diagnosis Level: Results.

Team	Accuracy	Precision	Recall	F1
CIRU	0.4589	0.4548	0.4137	0.4333
NTOU	0.2074	0.2932	0.4149	0.3436

KUAS&NTNU	0.2149	0.2696	0.3337	0.2983
UDS	**0.2337**	**0.2467**	**0.2594**	**0.2529**
TMU	0.4554	0.3545	0.1086	0.1662
NCYU	0.4594	0.2409	0.0377	0.0652

The error diagnosis level is more difficult than the error detection step. This is due to the multiple tags (e.g. missing words, word order) that could be attributed to each instance. At this stage, the performance of all systems was substantially lower than the error detection step. Once again our system was ranked 4th in terms of both F-score and accuracy achieving 23.37% F1 and 25.29% accuracy. The best system achieved 43.33% f-measure and 45.89% accuracy.

The dataset itself was to our understanding very challenging for the frequency-based methods we proposed. We found that some instances were virtually impossible to correctly tag. Examples of instances that were difficult to classify include single words: 老師 (EN 'teacher'), short expressions: 又很貴 (EN 'also very expensive') and instances that without context were difficult to understand even for native speakers: 姓本多 (EN literally: 'nature', 'by itself', 'many')

The results we obtained were consistently ranked in the middle of the table and they are, to our understanding, comparable to the state-of-the-art performance for the task. By looking at the performance obtained by the CIRU team, we see, however, room for improvement, as will be discussed in the next section.

5. Conclusion

This paper described the UDS submission to the shared task on Grammatical Error Diagnosis for Chinese as Foreign Language. We approached the task using frequency information and report results comparable to other state-of-the-art systems. The task is by no means trivial and the almost 9 percentage points behind the best system (CIRU team) showed us that there is still room for improvement. Even so, considering the lack of suitable training data, we believe that the results we obtained are still interesting to report.

We believe that better results can be obtained, for example, by integrating spell checkers (Lin and Chu, 2013) to our algorithm, particularly those that take phonetics into account (Zampieri and de Amorim, 2014). Another issue that should be taken into account in future experiments is the question of segmentation. Very good performance in tokenization is paramount when dealing with logographic languages and this was unfortunately not obtainable with the methods we used.

Finally, in the future we would like to perform experiments to increase the size of the training corpus using artificial errors as proposed by Felice and Yuan (2014). We believe that this is an effective way of producing more data for this task. The performance of these methods when applied to Chinese is still an open question.

Acknowledgements

We would like to thank the shared task organizers, especially Lung-Hao Lee, for providing prompt replies to all our inquiries. We acknowledge that no *LaTeX* template was available for this submission and we did our best to improve the quality of the formula and tables presented in this paper.

References

Brockett, C., Dolan, W. B., Gamon, M. (2006). Correcting ESL errors using phrasal SMT techniques. In Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. p. 249-256.

Chang, P. C., Galley, M., Manning, C. (2008) Optimizing Chinese Word Segmentation for Machine Translation Performance. In: Proceedings of WMT.

Chang, R. Y., Wu, C. H., Prasetyo, P. K. (2012). Error Diagnosis of Chinese Sentences Using Inductive Learning Algorithm and Decomposition-Based Testing Mechanism. ACM Transactions on Asian Language Information Processing, 11(1).

Chen, K. J., Huang, C. R., Chang, L. P., Hsu, H. L. (1996). Sinica corpus: Design methodology for balanced corpora. In: Proceedings of PACLIC. p. 167-176.

Dale, R., Anisimoff, I., Narroway, G. (2012). HOO 2012: A report on the preposition and determiner error correction shared task. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP* (pp. 54-62). Association for Computational Linguistics.

Felice, M., & Yuan, Z. (2014). Generating artificial errors for grammatical error correction. Proceedings of the EACL Student Workshop.

Felice, M., Yuan, Z., Andersen, Ø. E., Yannakoudakis, H., Kochmar, E. (2014). Grammatical error correction using hybrid systems and type filtering. In: Proceedings of CoNLL-2014.

Foster, J., Andersen, Ø. E. (2009). GenERRate: generating errors for use in grammatical error detection. In Proceedings of the fourth workshop on innovative use of NLP for building educational applications. Association for Computational Linguistics. p. 82-90.

Graff, D., Chen, K. (2003) Chinese Gigaword. Philadelphia: Linguistic Data Consortium.

Heafield, K. (2011). KenLM: Faster and smaller language model queries. In: Proceedings of the Sixth Workshop on Statistical Machine Translation. Association for Computational Linguistics. p. 187-197.

Heilman, M., Cahill, A., Tetreault, J. (2012). Precision isn't everything: a hybrid approach to grammatical error detection. In: Proceedings of the Seventh Workshop on Building Educational Applications Using NLP (pp. 233-241). Association for Computational Linguistics.

Leacock, C., Chodorow, M., Gamon, M., Tetreault, J. (2014). Automated grammatical error detection for language learners – Second Edition. Synthesis lectures on human language technologies. Morgan & Claypool.

Lee, L. H., Yu, L. C., Lee, K. C., Tseng, Y. H., Chang, L. P., Chen, H. H. (2014). A Sentence Judgment System for Grammatical Error Detection. Proceedings of COLING 2014.

Lin, C. J., Chu, W. C. (2013). NTOU Chinese Spelling Check System in SIGHAN Bake-off 2013. In Sixth International Joint Conference on Natural Language Processing.

McEnery, T. (2009) Keywords and moral panics: Mary Whitehouse and media censorship. in D. Archer (ed.) What's in Word-list? Investigating Word Frequency and Keyword Extraction. Oxford.

Ng, H. T., Wu, S. M., Wu, Y., Hadiwinoto, C.,Tetreault, J. (2013). The CoNLL-2013 Shared Task on Grammatical Error Correction. In: Proceedings of CoNLL-2013.

Ng, H. T., Wu, S. M., Briscoe, T., Hadiwinoto, C., Susanto, R. H., Bryant, C. (2014). The CoNLL-2014 shared task on grammatical error correction. In: Proceedings of CoNLL-2014.

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E. (2011) Scikit-learn: machine learning in Python. Journal of Machine Learning Research. 12. p. 2825-2830

Scott, M. (1997) PC Analysis of key words and key key words. System. n. 25. Elsevier. p. 233-245.

Tan, L., Bond, F. (2014). NTU-MC Toolkit: Annotating a Linguistically Diverse Corpus. In: Proceedings of COLING 2014. Dublin, Ireland.

Tetreault, J., Blanchard, D., Cahill, A. (2013). A report on the first native language identification shared task. In Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications. p. 48-57.

Tetreault, J. R., & Chodorow, M. (2008) The ups and downs of preposition error detection in ESL writing. In Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1. Association for Computational Linguistics. p. 865-872.

Wang, M., Voigt, R., Manning, C. (2014) Two Knives Cut Better Than One: Chinese Word Segmentation with Dual Decomposition. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, USA.

Yu, C. H., Chen, H. T. (2012). Detecting Word Ordering Errors in Chinese Sentences for Learning Chinese as a Foreign Language. Proceedings of the 24th International Conference on Computational Linguistics (COLING), pp. 3003-3017.

Yuan, Z., Felice, M. (2013). Constrained grammatical error correction using Statistical Machine Translation. In: Proceedings of CoNLL-2013.

Zampieri, M., de Amorim, R. C. (2014) Between Sound and Spelling: Combining Phonetics and Clustering Algorithms to Improve Target Word Recovery. Proceedings of the 9th International Conference on Natural Language Processing (PolTAL). Lecture Notes in Computer Science (LNCS). Springer.

Zampieri, M., Hermes, J., Schwiebert, S. (2013) Identification of Patterns and Document Ranking of Internet Texts: A Frequency-based Approach. Non-Standard Data Sources in Corpus-based Research. ZSM-Studien Series - Vol. 5. Shaker. p. 25-39.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Description of NTOU Chinese Grammar Checker in CFL 2014

Chuan-Jie Lin and Shao-Heng Chan
Department of Computer Science and Engineering,
National Taiwan Ocean University, Taiwan, R.O.C
cjlin@ntou.edu.tw

Abstract: This paper describes our first Chinese grammar checker participating in CFL 2014. Several features related to grammatical errors were proposed, including numbers of infrequent word bigrams and POS bigrams. Two SVM classifiers were trained and two formal runs were submitted, where the best F-scores were 66.67% in detection level and 34.36% in identification level.

Keywords: Chinese grammar checking, foreign language learning, machine learning

1. Introduction

Grammar checking for learning Chinese as a foreign language is a new challenge. Mistakes made by foreign students may greatly differ from the ones made by native speakers. It is necessary to study how to build a grammar checker for text written by students who learn Chinese as a foreign language.

As a shared task of ICCE, CFL 2014 (Grammatical Error Diagnosis for Learning **C**hinese as a **F**oreign **L**anguage) attempts to provide a benchmark to develop techniques on Chinese grammar checking. Four types of errors were defined in this task: redundant word, missing word, word disorder, and word selection problems. In this year, the task only focused on error detection and classification.

- **Redundant word**: a word should be deleted from this sentence
- **Missing word**: a missing word should be added into this sentence
- **Word disorder**: at least one word should change its location in this sentence
- **Word selection**: a word should be replaced into another word

This paper is organized as follows. Section 2 expresses some ideas we have got after observing examples in the training set. Section 3 gives definitions of features. Section 4 delivers experimental results and Section 5 concludes the paper.

2. Observation in Training Data

After observing example sentences in the training set, we found that the occurrence of low-frequency bigrams in the sentence is helpful. We used Google Web 1T 5-grams[1] (Google N-grams for short hereafter) as the resource of bigram frequencies. Some examples selected from the training set are provided here to illustrate our hypothesis.

(1) Example of redundant word problem

A redundant word is often unlikely to appear in that context. Moreover, removing this redundant word will create a higher-frequency bigram. For example,

[1] https://catalog.ldc.upenn.edu/LDC2006T13

[Sentence A2-0019-1]
可是 現在* 最近 我 工作 很 忙
(But now* recently my work is-very busy)

The word “現在” (now) has similar meaning with “最近” (recently), thus it is redundant. As an evidence, the bigram “現在+最近” is not collected in Google N-grams but the frequency of the bigram “可是+最近” is 218250.

(2) Example of missing word problem

Some examples provided in the training set are more like “a missing characters”, not “a missing word”. For example,

[Sentence A2-0026-1]
聽說 你 準備 開 一個 祝 會
(It-is-said-that you prepare to-have a wish* meeting*)
[Correct sentence]
聽說 你 準備 開 一個 慶祝會
(It-is-said-that you prepare to-have a celebration)

The character “慶” is missing, so the word “慶祝會” (celebration) cannot be correctly identified and is broken into two words “祝” and “會”. As an evidence, the bigram “祝+會” is not collected in Google N-grams.

(3) Example of word disorder problem

Word disorder means that order of the words should be re-arranged into a correct sentence. For example,

[Sentence A2-0027-1]
你* 很 久 以前 找 工作 很 幸* 苦
(You very long ago found jobs very lucky* difficult)
[Correct sentence]
很 久 以前 你 找 工作 很 辛苦
(Very long ago you found jobs very not-easily)

In Chinese, a long temporal phrase (“很久以前”, “very long ago” in this example) often appears in front of a complete sentence or, in another word, in front of a subject (“你”, “you” in this example). As an evidence, “以前+找” is not collected in Google N-grams but the frequency of the bigram “你+找” is 305477.

(4) Example of word selection problem

Word selection problem is that at least one word should be replaced with another, more appropriate word. For example,

[Sentence A2-0047-1]
我 真的 高興 你 找到 一* 新 的 工作 了
(I really am-happy you found one* new DE job LE)
[Correct sentence]
我 真的 高興 你 找到 一個 新 的 工作 了
(I really am-happy you found a new DE job LE)

(Note that DE and LE are function words without carrying much meaning)

When mentioning a countable noun in Chinese, quantifiers (量詞) should be used. For example, to say "a job", you use "一 個 工作" (one+GE+job), not "一 工作" (one+job). The character "個" (GE) in this example serves as a quantifier.

However, according to CNS14366, the Segmentation Standard for Chinese Natural Language Processing (中央標準局中文分詞標準, Huang *et al.*, 1997) in Taiwan, a number and a succeeding quantifier are segmented into two words, not grouped as one word. Such example is more like a missing problem rather than a word selection problem to us.

3. Error Detection Features

According to the observations described in Section 2, we defined several features to detect grammar errors as follows.

f_{bi-}:**number of infrequent bigrams** appearing in the sentence, where "infrequent bigram" is defined as a bigram NOT collected in Google N-grams. We expect that an erroneous sentence containing more infrequent bigrams. We are also interested to see if the number of infrequent bigrams is related to error types.

f_{POS-}: **number of infrequent POS bigrams** appearing in the sentence, where "infrequent POS bigrams" were trained from ASBC, a large POS-tagged corpus. Considering a POS bigram p_1p_2, if the probability P(p_2 | p_1) is less than 0.01, this bigram is an infrequent POS bigram.

f_{Nf-}: a Boolean feature denoting the **occurrence of a number without a succeeding quantifier**, where quantifiers are POS-tagged as Nf.

f_{stop}: a Boolean feature denoting the **occurrence of a stop POS bigram**. We defined a stop list of POS bigrams. POS bigrams in the stop list are:

- **VH + T**, a stative intransitive verb (mostly adjective in English) followed by a particle
- **Cbb + DE**, a correlative conjunction followed by a function word "的"
- **VC + Nd_DATE**, an active transitive verb followed by a date expression

f_D: a Boolean feature denoting the **occurrence of a key POS,** where key POS includes adverbs (D) and temporal nouns (Nd). Examples of disordered words often fall into these two POS classes.

$f_{bi=}$: **normalized number of infrequent bigrams,** i.e. f_{bi-} divided by the length of this sentence.

4. Run Submission and Results

Two runs were submitted to the CFL shared task this year. They were classification results from two different classifiers. System01 uses 5 features, f_{bi-}, f_{POS-}, f_{Nf-}, f_{stop}, and f_D. System02 also uses 5 features, $f_{bi=}$, f_{POS-}, f_{Nf-}, f_{stop}, and f_D. The only difference is the normalization of the first feature. Classifiers were trained by using LIBSVM (Chang and Lin, 2011).

Table 1 shows the performance of these two classifiers on training sets. The performances of the two systems are quite similar. Unfortunately, none of the classifiers can identify any word disorder case. System02 can correctly identify 4 word selection cases thus outperforms System01 a little.

Table 1: Performance in Training.

	System01			System02		
#	Prec.	Recl.	F-1	Prec.	Recl.	F-1
Redundant word	43.53	41.73	42.61	41.59	41.73	41.66
Missing word	43.71	75.51	55.43	44.71	75.51	56.22
Word disorder	0.00	0.00	0.00	0.00	0.00	0.00

Word selection	0.00	0.48	0.00	100.00	0.48	0.96
All	21.81	29.48	24.51	46.58	29.48	24.71

Table 2 shows the performance of two formal runs predicted by these two classifiers. The two systems have the same ability to detect errors. In fact, all sentences were predicted as "YES" but only half of them were correct. However, System02 achieved better performance in error-type classification thus outperforms System01 again.

Table 2: Performance of formal runs.

Submission	FP Rate	Detection Level				Identification Level			
		Acc.	Prec.	Recl.	F-1	Acc.	Prec.	Recl.	F-1
NTOU-Run1	100	50	50	100	66.67	16.00	24.24	32.00	27.59
NTOU-Run2	100	50	50	100	66.67	20.74	29.32	41.49	34.36

5. Conclusion

This paper describes our first Chinese grammar checker participating in CFL 2014. Six features related to grammatical errors were proposed, including numbers of infrequent word bigrams and POS bigrams. F-scores of formal runs were 66.67% in detection level and 34.36% in identification level. Normalized features seem outperform original numbers.

Because it was our first attempt to build a Chinese grammar checker, the performance was not satisfied. More studies and more features are needed for building a better system in the future.

References

Chang, C.-C. and C.-J. Lin (2011) LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, 27:1-27.

Huang, C.-R., K.-J. Chen, and Lili Chang (1997). Segmentation Standard for Chinese Natural Language Processing. *Computational Linguistics and Chinese Language Processing*, 2(2), 47-62.

Harnessing ICT for Educational Development in Emerging Developing Countries within the Asia-Pacific Region

Mas Nida MD KHAMBARI[a*], Ahmad Fauzi Mohd AYUB[b] & Mohammad LUKMAN[c]

Organizers of WICTTEE 2014[a,b,c]
Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia[a,b]
University of Muhammadiyah Prof. Dr. Hamka, Indonesia[c]
*nidakhambari@gmail.com

ICT trends in the educational landscape of emerging developing countries are growing and will continue to advance as technology plays a huge role in driving the development of economic and social sector (Kozma, 2005). Infact, the development of these sectors are often used as grounds to justify investments for ICT in education. Reciprocally, globalization has brought center the importance of knowledge creation (Kozma, 2011) with the help of ICT, consequently leveraging it with the aforementioned sectors. Current and new ICT trends are revolutionizing the ways knowledge are delivered and transferred, thus bringing several benefits as well as challenges to educators and learners. Especially in developing countries with diverse communities, the benefits and challenges can come in a multitude of forms. Now that "information and knowledge are the new forms of wealth and are the driving force for development" (Anderson, 2010, p. 10), it is interesting to learn how the citizens of developing countries go about their journey of harnessing ICT into the education system.

In response to the growing research diversity among emerging developing nations within the Asia-Pacific region, the Third International Workshop on ICT Trends in Emerging Economies (WICTTEE 2014) is held in conjunction with the 22nd International Conference on Computers in Education, Nara, Japan. WICTTEE 2014 is organized by the SIG on Development of Information and Communication Technology in the Asia Pacific Neighbourhood—DICTAP. The visions of DICTAP are to:

1. Share ideas and best implementation practices related to government policies and incentives aimed at promoting human resource development, technology transfer, effective e-learning strategies and implementation, software and content development suitable for each member of the Asia-Pacific neighborhood;
2. Coordinate and promote community-based e-learning activities, global sharing and management of information and knowledge. Examples of such communities are the Asia-Pacific Society on Computers in Education (APSCE) and the Association of South East Asian Nations (ASEAN); and
3. Coordinate and promote student and staff exchange among Asia-Pacific neighborhood member nations to promote more effective sharing of knowledge and practices.

The missions of DICTAP are to:
1. Connect researchers from emerging developing countries within the Asia-Pacific region to share scholarly findings and professional insights in ICT development in the field of education;
2. Establish networking opportunities among researchers, reduce the research gap between the researchers from more developed and less developed countries; and
3. Foster, enhance and sustain collaborations among these researchers.

WICTTEE 2014 is the third workshop that we are organizing in the hope to realise the aforementioned visions and missions. The workshop is a continuation of our relentless effort to provide a dynamic platform for practitioners and researchers alike to come together to share their country experiences.

We are extremely pleased that practitioners and scholars with university affiliations from Taiwan, Malaysia, Indonesia, Thailand, and Philippines will be congregating in Nara, Japan to present their research findings and share their views at WICTTEE 2014. A total of eight papers will be presented in a full day workshop.

We would like to take this opportunity to thank all the authors who submitted their papers to WICTTEE 2014. We would like to record our sincerest appreciation to our Program Committee Members who dedicated their time and expertise to the most challenging and demanding task of reviewing the paper submissions. Last but not least, we would like to thank DICTAP's Advisory Committee Members for their wisdom and guidance in making WICTTEE 2014 a reality.

References

Anderson, J. (2010). ICT transforming education: a regional guide. UNESCO: Bangkok.

Kozma, R.B. (2011). The Technological, Economic, and Social Contexts for Educational ICT Policy. In Transforming Education: The Power of ICT Policies. Paris: UNESCO.

Kozma, R.B. (2005). National policies that connect ICT-based education reform to economic and social development. Human Technology, *1*(2), 117-156.

Learning and Affect Trajectories Within Newton's Playground

Juan Miguel L. ANDRES[a*] & Ma. Mercedes T. RODRIGO[a]
[a]*Department of Information Systems and Computer Science,*
Ateneo de Manila University, Philippines
*miglimjapandres@gmail.com, mrodrigo@ateneo.edu

Abstract: Learning trajectories are typical, predictable sequences of thinking that emerge as students develop understanding of an idea. They have been used principally for research on instructional decision-making, but have also played a significant role in conducting research on learning. Affect refers to experiences of feelings or emotions. The affective states of boredom and confusion, in particular, have been of interest to researchers due to their significant relationships with student learning. In an attempt to account for gains or lack thereof among users of educational software, this study investigates these areas to monitor how students in the Philippines use software or how they are feeling compared to parallel studies conducted in the United States. In particular, this study investigates the relationships between learning trajectories and affect among students in the Philippines using Newton's Playground, a learning game for physics.

Keywords: Learning trajectories, affect trajectories, Newton's Playground

1. Theoretical Framework: Learning Trajectories and Affect

Learning trajectories (LTs) represent the "paths by which learning might proceed (Simon, 1995)." They are "typical, predictable sequences of thinking that emerge as students develop understanding of an idea (Daro, Mosher, & Corcoran, 2011)." LTs have been a topic of interest in recent years, enabling researchers to gain a better understanding of student learning. The study of LTs is less then two decades old (Clements & Sarama, 2004), and has only recently been getting attention in the field of learning sciences. Studies show that as researchers and teachers make sense of learning trajectories, they in turn can support growth in knowledge and further student learning.

Studying learning in terms of affective factors has also been a topic of interest in recent years. Two affective states of interest to researchers are confusion and boredom. Confusion or cognitive disequilibrium is the uncertainty about what to do next (D'Mello, Craig, Gholson, Franklin, Picard, & Graesser, 2005). It is interesting because it has a positive and negative dimension (D'Mello & Graesser, 2012), wherein it either spurs learners to exert effort deliberately and purposefully to resolve cognitive conflict, or leads learners to become frustrated or bored, and may lead to disengagement from the learning task altogether (D'Mello & Graesser, 2012).

Boredom, on the other hand, is defined by Fisherl (1993) as an "unpleasant, transient affective state in which the individual feels a pervasive lack of interest in and difficulty concentrating on the current activity." It has been a topic of interest because of the negative effects usually associated with it, like poorer learning and problematic behaviors, such as gaming the system (Baker, D'Mello, Rodrigo, Graesser, 2010; Rodrigo, Baker, & Nabos, 2010).

The extent to which students learn from educational software is influenced by the effectiveness with which they use software and how they feel while using the software. However, we do not always monitor how students use software or how they are feeling, so we cannot always account for gains or lack of gains. This study is an in-depth examination of software usage and affect and their interactions. Specifically, this study seeks to investigate the relationships between learning trajectories and the affective states of boredom and confusion.

2. Methodology

2.1 Participant Profile

We conducted a study to measure the relationship between a variety of affective and cognitive variables. Data was gathered from 60 eighth grade public school students in Quezon City, Philippines. Students ranged in age from 13 to 16. As of 2011, the school had 1,976 students, predominantly Filipino, and 66 teachers. Of the participants, 31% were male and 69% were female. Participants were asked to rate how frequently they played video games and watched television on a scale of 1 (not at all) to 7 (everyday, for more than 3 hours), and the resulting average frequency of gameplay is 3.2 (in between a few times a month, and a few times a week), and the resulting average frequency of watching television is 5.9 (in between everyday, but for less than 1 hour, and everyday, for 1-3 hours). Participants were asked for their most frequent grade on assignments, and on a scale of 0 (F) to 4 (A), the average most frequent grade of the participants is 3.1 (B).

2.2 Newton's Playground

Newton's Playground (NP) is a computer game for physics patterned after Crayon Physics Deluxe. It was designed to help secondary school students understand qualitative physics (Shute, Ventura, & Kim, 2013). Qualitative physics is a nonverbal conceptual understanding of how the physical world operates, along the lines of Newtonian physics. Qualitative physics is characterized by an implicit understanding of Newton's three laws: balance, mass, and conservation and transfer of momentum, gravity, and potential and kinetic energy (Shute et al., 2013).

NP is a two-dimensional computer-based game that requires the player to guide a green ball to a red balloon by drawing simple machines on the screen with colored markers controlled by the mouse. An example level is shown in Figure 3.1. The player uses the mouse to nudge the ball to the left and right (if the surface is flat), but the primary way to move the ball is by drawing or creating simple machines on the screen with the mouse and colored markers. The objects come to life once the object is drawn. Everything obeys the basic rules of physics relating to gravity and Newton's three laws of motion (Shute et al., 2013).

Figure 1. Example level of Newton's Playground.

The 74 levels in NP require the player to solve the problems via drawing different simple machines, representing agents of force and motion: inclined plane/ramps, levers, pendulums, and springboards. Again, all solutions are drawn with colored markers using the mouse. A ramp is any line drawn that helps to guide a ball in motion. A ramp is useful when a ball must travel over a hole. A lever rotates around a fixed point, usually called a fulcrum or pivot point. Levers are useful when a player wants to move the ball vertically. A swinging pendulum directs an impulse tangent to its direction of motion. The pendulum is useful when the player wants to exert a horizontal force. A springboard (or diving board) stores elastic potential energy provided by a falling weight. Springboards are useful when the player wants to move the ball vertically.

Gold badges versus silver badges. Some levels in NP have multiple solutions, which means a player can solve the level using different agents. Gold badges are awarded when a player solves a problem "under par", that is, under a limit set for a specific solution. For example, a level may be solved using a ramp, with a par of 1 object, or a pendulum, with a par of 3 objects. If a player solves the level with more objects than par, he receives a silver badge. Gold badges suggest that the player has mastered the agent relevant to the given level. Silver means the player may not have fully mastered the agent yet.

2.3 NP Interaction Logs

We collected two types of data during the study: interaction logs and human observations. During gameplay, NP automatically generates interaction log files. Each level a student plays creates a corresponding log file, which tracks every event that occurs as the student interacts with the game. The events per level and their respective attributes that are relevant to this study are:

- Level Start,
- Level Restart,
- Level End, an event that signals the player solved the level,
    - Badge, an attribute that states the type of badge (i.e. gold or silver) awarded to the player
    - Agent, an attribute that states for which agent the badge was awarded for
- Menu Focus, an event that signals the player gave up and quit the level without solving it,
- Drawing of any of the four agents,
- Object Limit, an event that is triggered by the player reaching the maximum number of objects drawn, and
- Stacking, an event that signals the player is gaming the system.

Each of these features provides useful information about students' gameplay behaviors, which can then be used to make inferences about how well they are doing in the game (Shute et al., 2013).

2.4 The Baker-Rodrigo-Ocumpaugh Monitoring Protocol

The Baker-Rodrigo-Ocumpaugh Monitoring Protocol (BROMP) is a protocol for quantitative field observations of student affect and behavior. BROMP is a holistic coding procedure that has been used in thousands of hours of field observations of students, from kindergarten to undergraduate populations. It has been used for several purposes, including to study the engagement of students participating in a range of classroom activities (both activities involving technology and more traditional classroom activities) and to obtain data for use in developing automated models of student engagement with Educational Data Mining (EDM) (Ocumpaugh, Baker, & Rodrigo, 2012). Within BROMP, each student observation lasts 20 seconds, and the observers move from one student to the next in a round robin manner during the observation period.

The affective states observed within Newton's Playground were concentration, confusion, frustration, boredom, happiness, delight, and curiosity. The behaviors observed were on-task, off-task, stacking, and a behavior called without thinking fastidiously (WTF), a behavior in which, despite a student's interaction with the software, "their actions appear to have no relationship to the intended learning task (Wixon, Baker, Gobert, Ocumpaugh, & Bachmann, 2012)."

The inter-coder reliability for affect was acceptably high with a Cohen's (1960) Kappa of 0.67. The typical threshold for certifying a coder in the use of BROMP is 0.6, established across dozens of studies as well as the previous affective computing literature.

2.5 The Human Affect Recording Tool

The Human Affect Recording Tool, or HART, is an Android application developed to guide researchers in conducting quantitative field observations according to the BROMP protocol. The application synchronizes the coded observations to Internet time, allowing for more precise synchronization with log file data from the educational software under study.

HART asks for input regarding school and classroom information, coding schemes to be used, and the student IDs of the students to be observed during the session. The application then presents the student IDs in the order entered, allowing BROMP observers to more conveniently code affect and behavior until the session is manually terminated. All observations are logged on a text file that is locally stored on the device used to run HART. The application and all its functions are discussed in more detail in (Ocumpaugh et al., 2012).

2.6 Data Gathering Process

Before playing NP, students completed a 16-item multiple-choice pretest for 20 minutes. Students were then assigned a computer on which they would play NP. Students played the game for two hours, during which, two trained observers used BROMP to code student affect and behavior. A total of 36 observations per participant per observer were collected. Videos of participants' faces were also recorded during gameplay. After completing the two hours of gameplay, participants completed a 16-point multiple-choice posttest for 20 minutes. The pretest and posttest were designed to assess knowledge of physics concepts, and has been used in previous studies involving NP (Shute et al., 2013).

In order to investigate learning within Newton's Playground, we made use of the interaction logs recorded during gameplay to analyze student performance. Of the 60 participants, data from 12 students were lost because of faulty data capture and corrupted log files. Only 48 students had complete observations and logs. The analysis that follows is limited to these students.

The BROMP observations were tabulated, and the percentage of each affective state per student was calculated. Boredom, confusion, and frustration were three of the more commonly observed affective states, besides concentration.

All interaction logs were passed through a parser to arrange log events neatly in tab-delimited text files. These text files were then run through a filter to get per student, per level, per attempt summaries, such as total time spent, total number of restarts, total number of objects drawn, etc. Finally, the information was collapsed to form per student vectors that summarized the students' entire interactions with the game. All statistical analyses conducted within this study are limited to the computation of percentages and result visualization.

3. Findings

We collected pre-test and post-test data from each student (N=60). Scores were generally poor. Students averaged 6.02 correct answers on both the pre-test and the post-test, out of a highest possible score of 16. This indicates that no learning improvements were detected. What follows is a descriptive analysis of the gathered data using methods described in the following subsections.

We operationalize learning trajectories on two levels:

1. On a coarse-grained level, learning trajectories are the performances of students in terms of gold and silver badges earned during gameplay, and
2. On a fine-grained level, learning trajectories are the sequences of students' interaction behaviors while solving or not solving a level.

As mentioned previously, affect coders followed BROMP, which resulted in 36 observations per student, per observer. For the purposes of this study, we define the incidence of affect as the percentage of students observed to be in a specific affective state during one observation count. We operationalize affect trajectories as the incidence of affect over time, that is, over the span of the 36 observations.

The findings in this section are from analyses conducted in finding:

1. The players' coarse-grained learning trajectories within NP,
2. The players' boredom and confusion trajectories, and
3. The relationships mined between the two.

3.1 Coarse-Grained Learning Trajectory Analysis

For the coarse-grained LT analysis, the percentages of students earning gold, silver, or no trophies were graphed over their opportunities to practice each of the four agents.

The three performance metrics (i.e. earning a gold trophy, earning a silver trophy, and earning no trophy) were used to track how well a student performed during gameplay, and in turn, see how well they understood each of the four agents used in the game. Every time a badge is awarded to a student, it is awarded for a specific agent. If a ramp was used to earn a gold badge, the student will get a gold badge for the ramp for that level. This is especially important for levels wherein any of the agents can be used to solve a level. Most levels, as the data showed however, only award badges for one of the four agents.

Using the logs generated by NP, trophies were grouped by level and by agent. In doing so, we were able to track which agents were awarded medals per levels, thus determining which agents were needed to solve each of the levels. Table 1 shows the tally of the first ten levels.

Table 1: Tally of trophies, by level and by agent for the first ten levels.

Level	Ramp	Lever	Pendulum	Springboard
P01L01	60			
P01L02	47			
P01L03		51		
P01L04			48	
P01L05				51
P02L01	47			
P02L02			46	
P02L03	46			
P02L04		46		
P02L05		22	46	30

This table shows which agents levels awarded badges for, whether gold or silver, which in turn gives us an idea of what agents were intended to solve the levels. The solution to Playground 1, Level 1, for example, is a ramp, with all sixty attempts on the level being awarded with ramp badges.

Using this table, we selected the first ten levels in which each of the four agents was used as a solution. Each of these first ten levels was then treated as an opportunity to practice one of the four agents. An opportunity to practice refers to a chance given to the student to exercise a specific skill, e.g. constructing a pendulum that pushes the ball to the target. Every level is solvable using one or more of these agents; therefore every level has one or more opportunities to practice possibly a variety of skills.

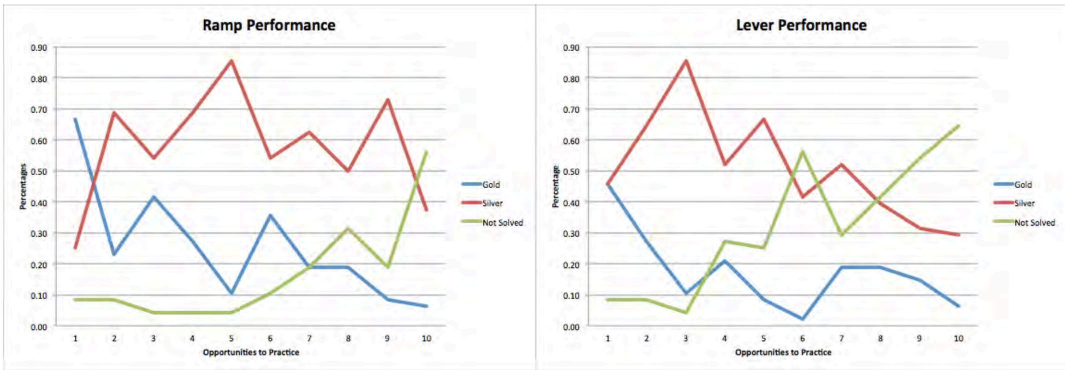


Figure 2. Learning trajectories of ramp and lever levels.

It is important to note that students were free to choose the levels they wanted to solve. The software did not force them through the material in a stepwise fashion. Furthermore, levels were not grouped thematically, by agent, so even if a student solved the levels sequentially, he would have opportunities to practice different agents.

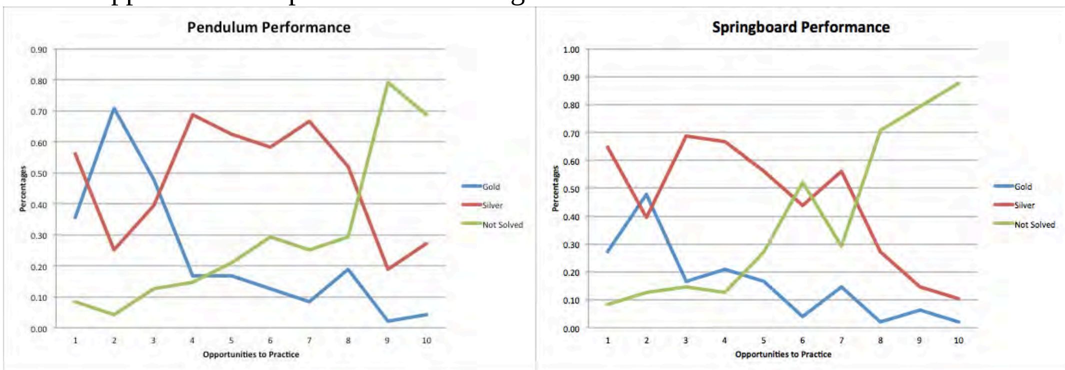


Figure 3. Learning trajectories of pendulum and springboard levels.

The percentage of students that earned gold trophies, silver trophies, and no trophies for each of the opportunities was then graphed. Figures 2 and 3 show the students' learning trajectories for each of the four agents. However, these graphs need to be unpacked further because students could choose what levels to solve, the actual levels corresponding to opportunities to practice 1 through 10 varied per student. That is for example, student 1's first opportunity to practice may be different from student 2's.

A consistent pattern can be observed across all four graphs, that is, as students progress through each of their ten opportunities to practice each of the four agents within NP, the perform more poorly over time. The percentage of students earning gold and silver trophies decreases over time, while the percentage of students unable to solve levels increases.

3.2 Fine-Grained Learning Trajectories

For the fine-grained LT analysis, a sequence mining analysis is to be conducted, taking into consideration the common sequences of actions student took in trying to solve each level.

As previously mentioned, NP generated interaction log files per level attempt that track every event that occurs as the player tries to solve the level. A filter was developed to pull only the relevant events from the log files. This filter is to be run on each level, arranging events chronologically on an output text file, divided by student. Events are to then be placed on a previous state-current state table to track transitions and transition frequencies between states. Using frequency calculations, common paths can easily be graphed and tracked through interaction network diagrams. This analysis is still in progress.

3.3 Affect Trajectory Analysis

As seen in the results above, students performed more poorly as they progressed through the levels in the game, and that at a certain point, the number of students earning trophies would just continuously decrease. The hypothesis this analysis sought to prove was that affect experienced by the students during gameplay could somehow be related to the students' eventual poor performance.

Using the logs generated by HART, all human observations per student were lined up on an Excel sheet. Each observer had a total of 36 observations per student. Using both of the observers' logs, a total of 72 observations per student were recorded. The percentage of students who were observed to be bored and the percentage of students who were observed to be confused per observation were calculated. An average percentage between the two coders was then calculated for each of the 36 observations. Figure 4 shows the affect trajectories of both confusion and boredom over time.

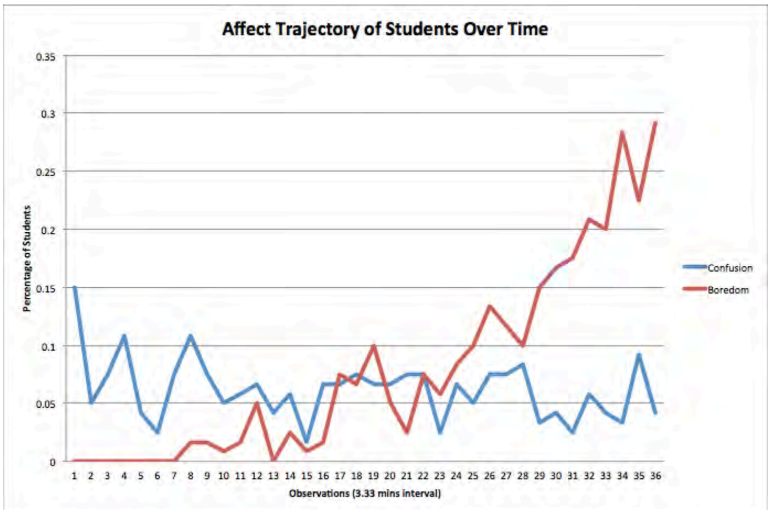


Figure 4. Affect trajectories for boredom and confusion.

It is interesting to note that while confusion was experienced by a steady number of students during the entire 2-hours of gameplay, the number of students experiencing boredom increased as the session progressed. The increase in percentage of boredom begins at observation 21, which is about one hour and ten minutes into the session.

3.4 Relationships

This study hopes to examine two relationships:

1. Between the coarse-grained LTs and affect, finding relationships between overall in-game student performance and incidences of both boredom and confusion, and
2. Between the fine-grained LTs and affect, finding relationships between common sequences and incidences of both boredom and confusion.

3.4.1 *Coarse-Grained LTs and Affect*

The BROMP observations were tabulated, and the percentage of each affective state per student was calculated. All interaction logs were passed through a parser to arrange log events in tab-delimited text files. These text files were then run through a filter to get per student, per level, per attempt gameplay features, such as total time spent, total number of restarts, total number of objects drawn, etc. Finally, the information was collapsed to form per-student vectors that summarized the students' entire interactions with the game. Each vector included the following performance metrics:

- Gold badge – percentage of level attempts solved, earning the student a gold badge
- Silver badge – percentage of level attempts solved, earning the student a silver badge

These two metrics were correlated with the students' respective percentages of boredom. The analysis, however, found no significant relationships. A previous study that ran the same methodology, however, found significant correlations between these metrics and confusion (Andres et al., in press). The study reported confusion to be negatively correlated with earning a gold badge, but positively correlated with earning a silver badge.

3.4.2 *Fine-Grained LTs and Affect*

This sequence mining analysis will take into consideration the common sequences mined in the previous analysis (in 3.2), and correlate them with the percentages of time the students were observed to be either confused or bored. We hypothesize that some sequences will be characteristic of either affective state, and can then be used as indicators within the game. As with the analysis in 3.2, this analysis is still in progress.

4. Discussion, Conclusions, and Future Work

The study attempted to identify learning and affect trajectories among students using an educational game for Physics, called Newton's Playground. In each level of the game, players are made to get a green ball to a red balloon using one or a combination of these four simple machines: lever, ramp, springboard, and pendulum.

The study operationalized learning trajectories (LTs) on two levels: coarse-grained LTs track the students' performance in terms of gold and silver badges earned, and fine-grained LTs track in-game events that occur as students interact with the game. Four coarse-grained learning trajectories were analyzed, one for each of the four simple machines in the game. All four coarse-grained LTs showed a common pattern of students performing more poorly as time progressed, earning less badges, and solving less levels. The fine-grained LT analysis is still in progress.

The study also looked at boredom and confusion trajectories among the students. Results showed that while confusion was experienced by a steady amount of students throughout the 2 hours of gameplay, the percentage of students experiencing boredom increased over time.

The study attempted to find relationships between the learning trajectories and affect, and in doing so, found no significant relationships between performance and boredom. A previous study found significant correlations with confusion, however, where confusion was negatively correlated to earning a gold badge, and positively correlated with earning a silver badge (Andres et al., in press). The analysis between fine-grained LTs and affect is still in progress.

We speculate that there are a number of relationships that are worth further exploration. ICT has not successfully penetrated the education system in the Philippines. Several infrastructural, financial, and implementation hindrances still exist, and despite the government's best efforts to work around them, programs and projects still fall through the cracks. Several government projects are currently in place, however, that aim to 1) ease the integration of ICT in the classroom for both teachers and students, 2) help alleviate poverty, most of which harness technology to maximize outcomes, and 3) utilize technology to reach potential learners who don't have immediate access to any form of formal learning.

On the student level, poor prior knowledge (as evidenced by students' poor pre-test results) might have made the game daunting, causing the students' poor performance in the game over time. The game interaction time of two hours may have been too long, leading to the increase in boredom. Indeed, the researchers noticed that the students rushed through the post-test, implying that they wanted to leave the testing area as quickly as possible. Boredom might have led to systematic guessing and other similar non-learning behaviors, leading in turn to poor post-test scores (Baker et al, 2010). In future work, we intend to verify which among these hypotheses the data support. In doing so, we hope to contribute to principles that guide the development of good educational games.

Acknowledgements

We would like to thank the Ateneo Center for Educational Development, Carmela C. Oracion, Christopher Ryan Adalem, the officials at Krus Na Ligas High School for making this study possible. We would also like to thank Ryan Baker, Valerie Shute, Luc Paquette, Matthew Ventura, Matthew Small, and the Gates Foundation Grant #OP106038 for collaborating with us.

References

Andres, J.M.L., Rodrigo, M.M.T., Sugay, J.O., Baker, R.S., Paquette, L., Shute, V.J., Ventura, M., & Small, M. (in press). An Exploratory Analysis of Confusion Among Students Using Newton's Playground. 22nd International Conference on Computers in Education

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. International Journal of Human-Computer Studies, 68(4), 223-241.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20 (1960), 37-46.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. Mathematical thinking and learning, 6(2), 81-89.

D'Mello, S., Craig, S. D., Gholson, B., Franklin, S., Picard, R., & Graesser, A. C. (2005). Integrating affect sensors in an intelligent tutoring system. In Affective Interactions: The Computer in the Affective Loop Workshop at (pp. 7-13).

D'Mello, S., Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2): 145-157.

Daro, P., Mosher, F. A., & Corcoran, T. (2011). Learning Trajectories in Mathematics: A Foundation for Standards, Curriculum, Assessment, and Instruction. CPRE Research Report# RR-68. Consortium for Policy Research in Education.

Fisherl, C. D. (1993). Boredom at work: A neglected concept. Human Relations, 46(3), 395-417.

Ocumpaugh, J., Baker, R.S.J.d., Rodrigo, M.M.T. (2012) Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0. Technical Report. New York, NY: EdLab. Manila, Philippines: Ateneo Laboratory for the Learning Sciences.

Rodrigo, M. M. T., Baker, R. S., & Nabos, J. Q. (2010). The relationships between sequences of affective states and learner achievement. In Proceedings of the 18th International Conference on Computers in Education (pp. 56-60).

Shute, V. J., Ventura, M., & Kim, Y. J. (2013). Assessment and Learning of Qualitative Physics in Newton's Playground. The Journal of Educational Research, 106(6), 423-430.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. Journal for Research in Mathematics Education, 26, 114–145.

Wixon, M., d Baker, R. S., Gobert, J. D., Ocumpaugh, J., & Bachmann, M. (2012). WTF? detecting students who are conducting inquiry without thinking fastidiously. In User Modeling, Adaptation, and Personalization (pp. 286-296). Springer Berlin Heidelberg.

Assessing Organizational Support and System Characteristics of Learning Management System : Views from Malaysian Higher Education Undergraduate Student

[a]Sousan BALEGHI ZADEH, [b]Ahmad Fauzi MOHD AYUB, [c]Rosnaini MAHMUD, [d]Shaffee MOHD DAUD
[abcd]*Faculty of Educational Studies, Universiti Putra Malaysia*
[bc]*Institute of Mathematical Research, Universiti Putra Malaysia*
s_baleghi@yahoo.com, afmy@upm.edu.my, rosnaini@upm.edu.my; shaffee@upm.edu.my

Abstract: Learning management system (LMS) is a kind of software that supports teaching and learning activities. Today, the number of educational institutions equipped with LMS is increasing. However, there are many factors that make students reject or accept this kind of technology. The present study assessed two characteristics of LMS (organizational support and system characteristics) in views of students. These two characteristics were assessed by three variables (technical support, system interactivity and system functionality). The respondents of the present study were 216 undergraduate students of faculty of education in Universiti Putra Malaysia. The result of the study revealed that in views of the respondents, organizational support was moderate, while system characteristics were high.

Keywords*:*Learning Management System, System Characteristics, Organizational Support

1. Introduction

One of the popular concepts that ICT has produced in the realm of education is e-learning (Hernandenz et al., 2011;Ŝumak, Heričko, & Pušnik, 2011). There are many definitions for e-learning. Some of them are broader and encompass different types of ICT, while the others are narrower. For example, Hill and Wouters (2010) have defined e-learning as "use of ICTs (e.g. Internet, Intranet, CD-Rom, interactive TV, teleconferencing, computer conferencing and char) to deliver instruction to learner" (p.204). However, some definitions of e-learning are narrower. For example, according to O'Mahony (2004) and Chang (2008) e-learning refers to any form of instruction delivered just through the web. Systems that conduct e-learning are different and have various names, such as online systems, virtual systems, learning management systems and so on (Piotrowski, 2010). To avoid getting confused, in the present study the term learning management system (LMS) is used for any kind of e-learning systems. LMS is one of the most popular software in that its usefulness in higher education institutions is widely increased (Chang, 2008; Dutta, Roy, &Seetharaman, 2013). Learning management system is a kind of information system that supports teaching and learning (Dutta et al., 2013). In fact, it organizes and provides tools through which students will be able to download learning contents, build, and deliver online learning environments (Piña, 2012). One of the most important benefits of LMS is to generate and manage reports on learners and assessment results (Theis, 2005). Besides, through the features of LMS, instructors and students can convey instructional materials, send notice to class, submit assignments, and interact with students (Lonn & Teasley, 2009). In fact, this information system combines technology features and pedagogy (Ioannou & Hannafin, 2008).

Although investing on LMS in institutional education is enhancing, research has reported that faculty and teachers are not interested in using technology (Chang, 2008; Hadjipavli, 2011).There are many factors which may affect LMS utilization of students and lecturers. However, Davis, Bagozzi, and Warshaw (1989) suggest that organizational characteristics and system characteristics are the crucial factors that may influence information system utilization. Technical support which is sometimes called facilitating support (Venkatesh & Bala, 2008) belongs to organizational characteristics. Technical support assists users to solve problems they encounter when they are working with an information system (Ngai et al., 2007). This factor enhances user satisfaction and has a critical influence on beliefs of users in accepting or rejecting an information system (Igbaria, Guimaraes & Davis, 1995). There are several studies in the LMS environment which indicate that technical support had a significant effect on LMS utilization. For example, Ngai et al. (2007) investigated the acceptance of LMS among undergraduate and postgraduate students of seven universities in Hong Kong with a sample size of 1263 and found out that technical support had a significant effect on LMS utilization. In another study, Sánchez and Hueros (2010) also examined LMS acceptance (Moodle) among students of business management and educational sciences in the University of Huelva (Spain) with a sample size of 226. The results of this research revealed that technical support had a significant effect on system utilization.

System characteristics encompass different variables such as system functionality and system interactivity (Davis et al., 1989; Pituch & Lee, 2006). Indeed, system functionality assesses the flexibility and quality of LMS features from the users' point of view, e.g. whether LMS is equipped with features through which students can send their assignments, download the contents of syllabus, take quizzes, use a variety of media such as audio, text and video and so on (Pituch & Lee, 2006; Selim, 2003). System interactivity refers to interaction among instructors and students in the process of learning and teaching (Pituch & Lee, 2006). The LMS should be equipped with features such as forum, email and chat room, through which students and teachers can interact with each other. Indeed, this factor assesses the interaction between lecturers and students (Pituch & Lee, 2006). There are several studies which indicate that system functionality and system interactivity had a significant effect on system usage. For example, Pituch and Lee (2006) investigated the influence of system characteristics (system interactivity and system functionality) on LMS utilization among 251 Taiwanese college students and found out that both system interactivity and system functionality had a significant effect on LMS utilization. Moreover, these variables obtained a high mean value. Wang and Wang (2009) also investigated the influence of system characteristics among 268 university instructors of Taiwan and revealed that system characteristics had a significant effect on LMS utilization. The outcomes of this research also indicated that system characteristics obtained a high mean value. The main purpose of the present study is to assess organizational characteristics and system characteristics of learning management system of Universiti Putra Malaysia (PutraLMS) in views of undergraduate students.

2. Research Objectives

i. To assess technical support of PutraLMS in views of students.
ii. To assess system interactivity of PutraLMS in views of students.
iii. To assess system functionality of PutraLMS in views of students.

3. Research Methodology

The population of the present study was full-time undergraduate students of faculty of educational studies of Universiti Putra Malaysia (UPM) in the second semester of the academic year 2012-2013. The students were selected through cluster sampling with a sample size of 216. The design of the present research is also descriptive.

4. Research Instrument

The variables of the present study were technical support, system interactivity, and system functionality and they were measured through a questionnaire with 23 items. Technical support is intended to measure services assisting undergraduate students of education at UPM to solve hardware and software problems with PutraLMS. The six items used to measure this construct were adopted from Sánches and Huerous (2010), and Ngai et al. (2007). The second construct is system interactivity. This construct is used to measure the ability of PutraLMS in providing facilities for interaction. This entails interaction among undergraduate students themselves, interaction between lecturers and students, and collaboration in learning which results from these interactions. This construct includes seven items adopted from Pituch and Lee (2006) and also self-developed items. It should be noted that self-developed items refer to items which were created by authors according to the definition of the variables. The last construct is system functionality which assesses undergraduate students' perception of flexibility of PutraLMS (UPM) in accessing instructional and media. This construct consists of 10 items adopted from Selim Ahmed (2010), Pituch and Lee (2006), and self-developed items. All items for this construct were measured through 5-point Likert-scale items labelled as 1 (strongly disagree), 2 (disagree), 3 (not sure), 4 (agree) and 5 (strongly agree). Two experts of the faculty of education at UPM examined the content validity of the questionnaire and their comments were followed. Reliability of the questionnaire was measured by Cronbach's alpha. Reliability of each variable was: technical support (.90), system interactivity (.91) and system functionality (.82). Since some items of the questionnaire were self-developed, its content validity was examined by four experts of education at Universiti Putra Malaysia.

5.0 Research findings

The results of the present study are based on descriptive study. The variables of the study are technical support, system interactivity, and system functionality.

5.1 Overall Mean

Table 1 indicates the overall mean of the variables of the study (technical support, system interactivity and system functionality). Among the three variables, system functionality obtained the highest mean (Mean = 3.83, SD = .53). This is followed by system interactivity (Mean = 3.53, SD = .74) and technical support (Mean = 3.35, SD = .67). The next section will discuss in detail items used to measure all the variables in this study

Table 1. Overall Mean of Variables

Factors	Number of items	Mean	SD
Technical Support	6	3.35	.67
System Interactivity	7	3.53	.74
System Functionality	10	3.83	.53

S.D.: standard deviation

5.2 Technical Support

When using any information system, technical support refers to assist users to solve problems they encounter when they are working with an information system (Ngai et al., 2007). Table2 indicates the mean and standard deviation of six items of technical support whereby the highest mean refers to PutraLMS offers good technical support (Mean = 3.42, S.D. = .820) followed by the manual on the operation of PutraLMS is sufficient(Mean = 3.38, S.D. = .887). There are two items which shared the same mean, which refers to e-mail inquiries to the technical support group when facing technical problem while using PutraLMS (Mean =3.35, SD = .799) and there is a hotline for fixing user problems (Mean = 3.35, S.D. = .787). Overall, the respondents felt that technical support is sufficient for them to use PutraLMS during their study period.

Table 2: Technical Support

Items	Source	Mean	SD
A hotline for fixing user problems is available at any time in PutraLMS.	Sánches & Huerous, 2010	3.35	.787
I can rely on the technical support group while using PutraLMS.	Self-developed	3.28	.823

Email inquiries to the technical support group can be made when there is a technical problem while using PutraLMS.	Sánches & Huerous, 2010	3.35	.799
Web-based inquiries can be made when there is a technical problem while using PutraLMS.	Sánches & Huerous, 2010	3.34	.853
The manual on the operation of PutraLMS is sufficient.	Ngai, Poon & Chan, 2007	3.38	.887
PutraLMS offers good technical support.	Sánches & Huerous, 2010	3.42	.820

TS: technical support; S.D: standard deviation

5.3 System Interactivity

Interactivity is an important aspect when users interact with information system. Therefore, system interactivity is the ability of the system to provide opportunities for interaction among users (Pituch & Lee, 2006). This section will describe students' perception towards the ability of PutraLMS in providing facilities for interacting among students, the interactions between lecturers and students, and collaboration in learning which grows out of these interactions. Analysis towards seven items shows that the highest mean refers to students' perception towards PutraLMS enabling interactive communication between lecturers and students (Mean = 3.65, S.D. = .93) followed by PutraLMS enabling them to receive comments (Mean = 3.61, S.D. = .86) and features of collaborative learning in PutraLMS (Mean = 3.60, S.D. = .90) (refer to Table 3). Overall, the students felt that PutraLMS provide features that enable them to interact among the colleagues and lecturers.

Table 3: System Interactivity

Item	Source	Mean	S.D.
PutraLMS enables interactive communication between lecturers and students.	Pituch & Lee, 2006	3.65	.93
I can see the features of collaborative learning (e.g. group work) in PutraLMS	Self-developed	3.60	.90
The communication tools (email, forum, chatroom, etc.) in PutraLMS are effective.	Pituch & Lee, 2006	3.51	.90
PutraLMS enables me to receive my lecturers' comments.	Self-developed	3.61	.86
PutraLMS enables interactive communication among students.	Pituch & Lee, 2006	3.49	.90
I can share my knowledge with my classmates through PutraLMS.	Pituch & Lee, 2006	3.53	.87
My lecturers often communicate with us through PutraLMS.	Pituch & Lee, 2006	3.37	1.07

SI: system interactivity; S.D.: standard deviation

5.3 System Functionality

System functionality is flexibility of an information system (Pituch & Lee, 2006). Therefore, this section will investigate students' perception of flexibility of PutraLMS in accessing instructional and assessing media. The construct of system functionality was measured by 10 items. The highest mean refers to the capability of PutraLMS to print course materials (Mean = 4.33, S.D. = .714). Students also felt that they can access PutraLMS from any place (Mean = 4.14, S.D. = .851) and it offers flexibility in learning regarding time (Mean = 4.08, S.D. = .742)

Table 4: System Functionality

Items	Source	Mean	S.D.
I can print course materials through PutraLMS.	Selim, 2010	4.33	.714
PutraLMS offers flexibility in learning regarding time.	Pituch & Lee, 2006	4.08	.742
I can access PutraLMS from any place.	Pituch & Lee, 2006	4.14	.851

PutraLMS offers different types of material (e.g., audio, video, and text) for every kind of course content.	Pituch & Lee, 2006	3.62	1.019
PutraLMS provides opportunity for taking tests.	Pituch & Lee, 2006	3.27	.957
PutraLMS presents course material in a well-organized manner.	Pituch & Lee, 2006	3.70	.958
PutraLMS clearly presents course contents.	Self-developed	3.93	.804
PutraLMS facilitates groupwork.	Self-developed	3.50	.857
PutraLMS provides opportunity for sending assignments.	Pituch & Lee, 2006	3.72	.893
The course material in PutraLMS is in a readable format.	Pituch & Lee, 2006	4.05	.649

SF: system functionality; S.D.: standard deviation

5. Discussion

Learning management system sometimes called e-learning platform, e-learning, online learning, and virtual learning emerged when instruction via network was conducted (Chang, 2008; Piotrowski, 2010). LMS provides a variety of opportunities for instructors and learners to increase their educational experiences (Holmes & Gardner, 2006). Today, LMS is widely used in higher education (Dutta, 2013). Nevertheless, the outcomes of many studies indicate that the quality of using LMS by students and faculties is limited (The Campus Computing Survey, 2008). For example, Lam, Lo, Lee, and McNaught (2012) investigated using LMS (WebCT) by undergraduate and graduate students in the Chinese University of Hong Kong and found that only 14.8 percent of the students used features for online discussion. According to Almarashdeh, Sahari, Mat Zin, and Alsmadi (2010), in almost all colleges of Malaysia, many lecturers use LMS just for transferring materials and never use communication features such as discussion board, wiki, chat room and so on. In another research, Embi, Hamat,and Sulaiman (2012) examined LMS utilization among 26 Malaysian university lecturers. The results of their study revealed that only two-thirds of lecturers used LMS, such that 65 percent of utilization was restricted to course delivery.

There are many factors affecting utilization of information system. However, system characteristics and organizational support can be considered as crucial factors that may affect system utilization (Davis et al., 1989; Venkatesh & Bala, 2008). In the present study, the mean value of three variables of technical support, system interactivity, and system functionality were measured. Among these variables, system interactivity and system functionality belong to system characteristics and technical support belongs to organizational system. In the present study, technical support obtained the lowest mean. This suggests that organizational support should increase assistance of students when they encounter hardware or software problems. Although system functionality and system interactivity obtained higher mean, it is suggested that system flexibility and system interactivity of PutraLMS improve.

This study has its own limitations. First the respondents of the study were full time students of faculty of education; therefore, the generalization of the findings should be done by care, because the backgrounds of part time students as well as students from other faculties may be different from full time students. Second, there are many variables such as perceived usefulness, perceived ease of use and subjective norm which may affect LMS utilization of users, but the present study was limited to investigating three variables (technical support, system interactivity and system functionality).

References

Almarashdeh, I. A., Sahari, N., Mat Zin, N. A., & Alsmadi, M. (2010). The success of learning management system among distance learners in Malaysian universities. *Journal of Theoretical and Applied Information Technology 21*(2), 80-91.

Chang, C. L. (2008). *Faculty perceptions and utilization of a learning management system in higher education* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Theses database (UMI No. 3319031).

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003.

Dutta , A, Roy, R., & Seetharaman, P. (2013). Course management system adoption and usage: A process theoretic perspective. *Computers in Human Behavior, 29*, 2535–2545.

Embi, M. A., Hamat, A., Sulaiman, A. H. (2012). The use of learning management systems among Malaysian university lecturers. *The International Journal of Learning*, *18*(4), 61-70.
Hadjipavli, E. (2011). *An examination of Cypriot teachers' concerns regarding the adoption of a learning management system in secondary education* (Doctoral dissertation, North central University, Arizona, USA).Retrieved from ProQuest Dissertations and Theses database. (UMI No 3467400)
Hernandez, B., Montaner, T., Sese, F. J., & Urquizu, P. (2011).The role of social motivations in e-learning: How do they affect usage and success of ICT interactive tools? *Computers in Human Behavior*, *27*(6), 2224-2232.
Hill, N. S., & Wouters, K. (2010). Comparing apples and oranges: Toward a typology for assessing e-learning effectiveness. In H. Liao, J.J. Martocchio, & Joshi, A. (Eds.), *Research in personnel and human resources management: Volume 99* (pp. 201-242) Bingley, UK: Emerald.
Holmes, B., & Gardner, J. (2006).*E-learning: Concepts and practice*. London: Sage Publications.
Igbaria, M., Guimaraes, T., & Davis, G. B. (1995).Testing the determinants of microcomputer usage via a structural equation model. *Journal of Management Information Systems*, *11*(4), 87–114.
Ioannou, A. & Hannafin, R. (2008). Deficiencies of course management systems: Do students care*? The Quarterly Review of Distance Education*, *9*(4), 415–425.
Lam, P., Lo, J., Lee, J., & McNaught, C. (2012). Evaluations of online activities based on LMS logs. In R. Babo & A. Azevedo (Eds.), *Higher education institutions and learning management systems* (pp. 75-93). Hershey, PA: Information Science Reference.
Lonn, S., & Teasley, S. D. (2009). Saving time or innovating practice: Investigating perceptions and uses of learning management systems. *Computers & Education*, *53*(3), 686-694.
Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H.C. (2007).Empirical examination of the adoption of WebCT using TAM. *Computers & Education*, *48*(2), 250–267.
O'Mahony, C. (2004). E-learning component evolution and integration: A case study. *International conference on computers in Education2004*. Retrieved September 2012 from http://plum.yuntech.edu.tw/icce2004/Theme3/083_OMahany.pdf.
Piña, A. A. (2010). An overview of learning management systems. In Y. Kats (Ed.), *Learning management system technologies and software solutions for online Teaching* (pp. 1-19). Hershey, PA: Information Science Reference.
Pituch, K. A., & Lee, Y-K. (2006). The influence of system characteristics on e-learning use. *Computers & Education*, *47*(2), 222-244.
Piotrowski, M. (2010). What is an e-learning platform? In Y. Kats (Ed.), *Learning management system technologies and software solutions for online teaching: Tools and applications* (pp. 20-36). Hershey, PA: Information Science Reference.
Sánchez, R. A., & Hueros, A. D. (2010). Motivational factors that influence the acceptance of Moodle using TAM. *Computers in Human Behavior*, *26*(6), 1632–1640.
Selim Ahmed, H. M. (2010).Hybrid e-learning acceptance model: Learner perceptions. *Decision Sciences: Journal of Innovative Education*, *8*(2), 313- 346.
Selim, H. M. (2003). An empirical investigation of student acceptance of course websites. *Computers & Education*, *40*(4), 343-360.
Šumak, B., Heričko, M., & Pušnik, M. (2011). A meta-analysis of e-learning technology acceptance: The role of user types and e-learning technology types. *Computers in Human Behavior*, *27*(6), 2067-2071.
The campus computing survey 2008, Retrieved May 2012 from:http://www.campuscomputing.net/survey-summary/2008-campus-computing-survey.
Theis, J. G. (2005). *Perceived learning and performance needs of organizations seeking a learning management system* (Doctoral dissertation, The University of Minnesota, USA). Retrieved from ProQuest Dissertations and Theses database. (UMI No 3167659)
Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273-315.
Wang, W-T., & Wang, C-C. (2009). An empirical study of instructor adoption of web-based learning systems. *Computers & Education*, *53*(3), 761-774.

Exploring Deep Approach to Learning for Accounting through ICT-Supported Learning Environment in Malaysian Secondary Schools: A Preliminary Study

Boon See TAN[a*] & Su Luan WONG[b]
[a]*Faculty of Educational Studies,Universiti Putra Malaysia, Malaysia*
[b]*Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia*
*bstan_29@yahoo.com

Abstract: Under the commitment of the Malaysian Ministry of Education to transform the curriculum and assessment of Principles of Accounting in secondary schools where deep approach to learning is critical for attaining quality learning outcomes, this preliminary study aims to assess the instrument used to measure students' approaches to learning for Principles of Accounting and perceptions of ICT-supported learning environment. It also aims to explore the current status of students in terms of these two variables, and their relationship. A total of 33 secondary school students who are studying Principles of Accounting in an ICT-supported learning environment participated in this study. They responded to the questionnaire which consists of the subscales of Deep Approach, Surface Approach that measure students' approaches to learning, as well as subscales of Student Cohesiveness, Teacher Support, Cooperation, Involvement, Investigation, Task Orientation, Computer Usage, and Appropriate Assessment that assess for perceptions of ICT-supported learning environment. The findings indicate that the measures for Surface Approach need to be refined as it only obtained a satisfactory level of reliability value compared to other subscales. The descriptive statistics findings reveal that students exhibited higher deep approach than surface approach to learning. They also showed a rather low score in Appropriate Assessment which may reflect a more fact-and memory oriented assessment was perceived. Further Pearson correlation analysis suggests that Teacher Support and Computer Usage are the main variables to be significantly related to Deep Approach. They are simultaneously associated with other subscales of perceptions such as Student Cohesiveness, Cooperation, Involvement, Investigation, and Task Orientation that are also significantly related to Deep Approach. In addition, most of the subscales of perceptions are related to Deep Approach except Appropriate Assessment. This suggests that the perceived fact and memory-oriented assessment was not significantly associated with Deep Approach.

Keywords: Deep approach to learning, surface approach to learning, perceptions of ICT-supported learning environment, Principles of Accounting

1. Introduction

Accounting is widely referred to as the "language of business". Its functions are identifying, measuring, and communicating economic information to permit informed judgments and decisions by users of the information (Martin, 1994). The fundamental of accounting encompasses the basic knowledge of accounting in identifying and measuring financial information through the application of the double-entry book-keeping system. In Malaysia, the education of fundamental accounting starts at the upper secondary school level (i.e. Form 4 and 5 or Grade 10 and 11) where the accounting knowledge and skills are delivered through the subject of Principles of Accounting. This subject consists of the concepts, principles, and accounting methods complemented by the skills in recording, classifying, interpreting, and summarising financial data based on business transactions (Technical Education Department, 2000).

It was observed by few researchers that students always perceive that learning accounting is simply about learning a set of rules and evidences suggest that they tend to adopt a surface learning approach compared to other subjects (Eley, 1992; Beattie, Collins, & McInnes, 1997; Booth, Luckett, & Mladenovic, 1999; Lucas, 2001). Similarly, in the Malaysian context, students' learning for the subject of Principles of Accounting has yet to achieve deep approach as it was found that most of the accounting teachers tended to use the teacher-centred teaching methods such as lecture, drill and practice, and demonstration of problem solving without delving deeper into the knowledge (Suhaida Abdul Kadir, 2002; Hanuni Yusuf, 2003; Rohaila Yusof, 2006). Such methods could lead to surface learning where the lower-level procedural skills are acquired without processing information for meaning. This may be a reflection of the exam-oriented education system which is the common learning issue faced by Asian countries. It affects many teachers unwilling to take the risk of students' failure in examination by attempting innovative teaching. They would rather employ the teaching methods which comply with the requirements of the examination system which is mostly teacher-centred (Looi, Hung, Chen, & Wong, 2006). Moreover, students are short of the ICT skills to adopt deep approach to learning for managing the whole set of accounts by relating the processes of accounting cycle into a coherent whole (Arfah Salleh, 2001; Rashidah Hassan & Arfah Salleh, 2008; Tan & Wong, 2012).

In view of the aforementioned weaknesses in the Principles of Accounting education, the Malaysian Ministry of Education has strived to transform its curriculum and assessment from content-based to skilled-based from 2010 onwards where deep approach to learning is critical for attaining the desired learning outcomes such as the master of ICT skills and soft skills (e.g. higher order thinking skills, communication skills, problem solving, and decision making skills) (Bahagian Pembangunan Kurikulum [Curriculum Development Section], 2009). However, to what extent the revised curriculum and assessment are able to foster students' deep approaches to learning, particularly under the ICT-supported learning context? Thus, a preliminary study was conducted with the following objectives:

- To assess the appropriateness and reliability of the instrument;
- To take the first step in exploring students' approaches to learning and perceptions of ICT-supported learning environment; and
- To conduct initial investigation on the relationship between students' perceptions of ICT-supported learning environment and approaches to learning.

2. Literature Review

2.1 Students' Approaches to Learning (SAL)

SAL is defined as the ways in which how a student perceived a particular academic task and then handle it (Marton & Säljö, 1976). In addition, SAL is seen as a contextually dependent response rather than an enduring characteristic of the individual (Meyer, Parsons, & Dunne, 1990). It is further identified into two contrasting approaches i.e. deep approach and surface approach. A deep approach entails learners' intrinsic motivation and interest to attempt to understand the meaning of the learning material and relate parts to each other, new ideas to previous knowledge or to personal meaningful context; whereas a surface approach is characterised by extrinsic motivation to acquire only sufficient knowledge to complete the task or pass the subject and thus, learners tend to memorise separate facts and/or view a particular task in isolation from other tasks and real life as a whole (Marton & Säljö, 1976; Biggs, 1985; 1987a; Biggs and Moore, 1993; Wong, Lin, & Watkins, 1996; Biggs, Kember, & Leung, 2001; Kember, Biggs, & Leung, 2004).

2.2 The Influence of ICT-Supported Learning Environment Perceptions on Approaches to Learning

There have been many studies reporting that significant relationships exist between students' perceptions of learning context and approaches to learning. It was found that approaches to learning are influenced by different perceptions of students studying different subject areas (Ramsden, 1979). In this vein, accounting students were especially influenced by their learning context which perceived as being tensed up with pressure and demands from the professional accounting bodies and there was evidence that most of them adopt the surface approaches to learning (Eley, 1992; Gow, Kember, & Cooper, 1994; Sharma, 1997; Booth et al., 1999; Jackling, 2005; Lord & Robertson, 2006; Abraham 2006). On the other hand, deep approach to learning was found to be associated with perceived quality teaching support (Eley, 1992; Chan & Watkins, 1994; Lizzio, Wilson, & Simons, 2002), appropriate pedagogy which encourages independence, interaction, and inquiry (Eley, 1992; Abraham, 2006), and appropriate assessment (Abraham, 2006; Watty, Jackson, & Yu, 2010).

Furthermore, the ICT-supported learning environment is especially contributing to students' perceptions of interactivity and involvement (Maor, 2000; Law, Lee, & Chow, 2002; de Lange, Suwardy, & Mavondo, 2003; Jebeile & Abeysekera, 2010; Premuroso, Tong, & Beed, 2011), inquiry and investigation (Basu & Cohen, 1994; Siragusa, 2002; Jones, Scanlon, Gaved, Blake, Collins, Clough et al., 2013), authencity of learning (Basu & Cohen, 1994; Green, Reinstein, & Mc Williams, 2000; Murphy & Hoeppner, 2002; Marriott, 2004; Stanley & Edwards, 2005; Neal 2005), cooperation (Rumpagaporn, 2007), differentiation (Jebeile & Abeysekera; 2010) as well as the perceptions of teacher support (Rumpagaporn, 2007; Lillie & Wygal, 2011). These perceptions were found contributing to students' deep approaches to learning. Meanwhile, the ICT-supported learning environment perceived by students to have replaced them by producing the accounting reports automatically (Green et al., 2000) and being a safety net for absence (Wells, de Lange, & Fieger, 2008) were associated to surface approaches to learning.

3. Methodology

This study was conducted through adopting the quantitative descriptive research design to explore students' perceptions of ICT-supported learning environment and approaches to learning as well as their relationship. It was conducted through a survey where questionnaires were developed and distributed to the targeted respondents to collect the related information.

3.1 Subject

The subjects were 33 Form 4 (or Grade 10) students who are studying Principles of Accounting under an ICT-supported learning environment which is defined by Aldridge, Dorman, and Fraser (2004) as an environment where students could utilise various ICT tools to assist their learning. For example, using ICT to complete and submit assignments, search information, obtain notes, and conduct on-line discussion and communication with teacher and peers.

3.2 Measures

Students were asked to complete a questionnaire assessing their approaches to learning and perceptions of ICT-supported learning which was adapted from the Revised Two-Factor version of the Learning Process Questionnaire (R-LPQ-2F) (Kember, Biggs, & Leung, 2004), Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI) (Aldridge et al., 2004), and Course Experience Questionnaire (CEQ) (Wilson, Lizzio, & Ramsden, 1997). They were requested to respond to a five-point Likert scale for all the items in the questionnaire from "strongly disagree" to "strongly agree". The questionnaire was back-to-back translated into the Malay language as most of the students in public secondary schools in Malaysia are proficient in this language.

The R-LPQ-2F was adapted to determine students' approaches to learning. It consists of 22 items where both Deep and Surface Approach are measured by 11 items respectively. This instrument is chosen as it is specially designed to measure the approaches to learning of secondary school students and it is a shorter version that is suitable for use as a classroom evaluation tool which was developed from the original version of Learning Process Questionnaire (LPQ) designed by Biggs (1987b). Moreover, this simple two-factor instrument was verified by the authors as valid and reliable with good psychometric properties. It contains scales that scored good Cronbach alpha values (Deep Approach = .82 and Surface Approach = .71) and exhibited good construct validity (Comparative Fit Index, or CFI = .967, Standardised Root Mean Squared Residual, or SRMR = 0.036).

On the other hand, students' perceptions of ICT-supported learning environment were assessed through the instrument adapted from the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI) (Aldridge et al., 2004) and Course Experience Questionnaire (CEQ) (Wilson et al., 1997). The TROFLEI provides the scales of Student Cohesiveness, Teacher Support, Cooperation, Involvement, Investigation, Task Orientation, and Computer Usage where each subscale was measured by seven to nine items; while the scale of Appropriate Assessment was developed from CEQ which is measured by five items. TROFLEI is renowned in terms of its reliability and validity for being extensively cross-validated in various countries such as Taiwan and Australia (Aldridge, Fraser, & Huang, 1999); and the United Kingdom, Canada, and Australia (Dorman, 2003). It is also advocated by the authors that the instrument showed good reliability where its scales yielded high Cronbach's alpha values (ranging from .77 to .95) with sound construct validity (CFI=0.98). In addition, the Appropriate Assessment scale from CEQ was validated by the authors as it attained consistently good Cronbach's alpha values ranging from .69 to .75. for three consecutive years of study. It also yielded high loadings for each item ranging from .56 to .85 which reflected good construct validity.

3.3 Procedures

The questionnaire was directly administered by the main author of the present study to all the subjects in order to ensure the genuineness of the data collection process and to obtain higher response rate. In addition, respondents were given prompt assistance should they encounter any difficulties in answering the questionnaire. Permission was first acquired from the school principal followed by discussion with the Principles of Accounting teachers on the suitability of date, time, and place for conducting the survey. The survey process started with a 10-minute briefing on the purpose of the research and instructions on how to answer the questionnaire.

4. Results

The first stage of this study employed the reliability test to assess the internal consistency of each subscale, followed by descriptive statistics in order to understand the overall status of respondents' perceptions. Pearson correlation was used for analysis of relationships among the subscales in the second stage by reporting the inter-correlational matrix to provide an overview of the univariate relationships.

4.1 Reliability and Descriptive Statistics

Each subscale of the instrument achieved a high level of reliability, except for Surface Approach. Thus, a few items were dropped from this subscale in order to secure a satisfactory level of reliability. The reliability values i.e. Cronbach's alpha values (α) together with descriptive statistics for each subscale are shown in Table 1.

All means are greater than 3.5, except the Surface Approach and Appropriate Assessment. It indicates that students generally have positive perceptions towards their learning environment and have low tendency to adopt surface approaches to learning but are more inclined towards using deep approach. However, the rather low mean value in Appropriate Assessment may reflect a more fact- and memory-oriented assessment for the subject of Principle of Accounting. In addition, the standard deviations for all subscales were less than 1 and this suggests that the scores tend to be very close to the mean scores.

Table 1: Reliability values and descriptive statistics of the subscales.

Subscale	Mean	Standard Deviation	Items	Cronbach Alpha (α)
Approaches to Learning				
Deep Approach (DA)	3.83	.69	11	.91
Surface Approach (SA)	3.01	.66	6	.71
Perceptions of ICT-Supported Learning Environment				
Student Cohesiveness (SC)	3.83	.73	8	.91
Teacher Support (TS)	3.79	.84	8	.92
Cooperation (CO)	4.06	.71	8	.93
Involvement (IV)	3.72	.64	8	.85
Investigation (IVT)	3.62	.76	7	.88
Task Orientation (TO)	3.88	.67	8	.89
Computer Usage (CU)	3.72	.94	9	.93
Appropriate Assessment (AA)	2.68	.89	5	.80

4.2 *Inter-correlation Matrix*

Table 2 shows the inter-correlation matrix for scores on the 8 subscales of Perceptions of ICT-Supported Learning Environment, Deep Approach, and Surface Approach.

Table 2: Inter-correlation matrix between Deep Approach, Surface Approach, and subscales of Perceptions of ICT-supported Learning Environment (N=33).

	DA	SA	SC	TS	CO	IV	IVT	TO	CU
SA	.43*								
SC	.43*	.03							
TS	.73**	.25	.57**						
CO	.35*	.07	.66**	.50**					
IV	.42*	-.05	.83**	.65**	.76*				
IVT	.44*	.16	.43*	.47**	.72**	.64**			
TO	.56**	.23	.54**	.58**	.59**	.65**	.68**		
CU	.51**	.03	.67**	.34	.34	.48**	.36*	.53**	
AA	.02	-.10	-.29	-.04	-.42*	-.19	-.29	-.25	-.16

Notes: * = $p < .05$; **= $p < .01$

Deep Approach has significant positive relationships with all the subscales of perceptions, except Appropriate Assessment (r=.02). Meanwhile, Appropriate Assessment does not have significant relationships with most of the other subscales of perceptions. Furthermore, Teacher Support has the strongest positive relationship (r=.73) with Deep Approach which reflects the strong link between teacher's role and students' adoption of the deep approach. In addition, Teacher Support is significantly related to other subscales of perceptions such as Cooperation (r=.50), Involvement (r=.65), Investigation (r=.47), and Task Orientation (r=.58). It further suggests the importance of teacher's role in the learning environment.

On the other hand, none of the subscales is significantly related to Surface Approach. This could suggest that the ICT-supported learning environment which forms students' perceptions towards Student Cohesiveness, Teacher Support, Cooperation, Involvement, Investigation, Task Orientation, Computer Usage, and Appropriate Assessment is not associated to Surface Approach but more towards Deep Approach.

Furthermore, it is found that Computer Usage has significant relationships with many other subscales such as Student Cohesiveness (r=.67), Involvement (r=.48), Investigation (r=.36), and Task Orientation (r=.53). This suggests that ICT is related to many dimensions of the learning environment.

Lastly, it was found that many of the subscales of perceptions are strongly related, e.g. Student Cohesiveness and Involvement (r=.83), Involvement and Cooperation (r=.76), and Cooperation and Investigation (r=.72). It could suggest a phenomenon of multicollinearity which one should pay attention when conducting multiple regression analysis. Multicollinearity is an undesired statistical phenomenon in which two or more independent variables in a multiple regression model are highly correlated and will affect the accuracy of the model (Hair, Anderson, Tatham, & Black, 1995).

5. Discussion

From the findings, it reveals that the instrument measures for Surface Approach needed to be further refined as a few items had to be dropped in order to obtain a satisfactory but relatively lower reliability value compared to other subscales.

On the other hand, though there are many studies claiming that students adopt the surface approach for accounting by rote learning a set of rules (Eley, 1992; Beattie et al., 1997; Booth et al., 1999; Lucas, 2001), the present study preliminarily found that students demonstrated deep approach for learning Principles of Accounting. This phenomenon was found strongly associated to the perceptions of teacher support and it is consistent with many studies that found the significant relationship between teacher support and deep approach (Eley, 1992; Chan & Watkins, 1994; Lizzio et al., 2002), as well as teacher support and deep approach under the learning environment that supported by ICT (Rumpagaporn, 2007; Lillie & Wygal, 2011). Furthermore, the significant positive relationships between perceptions of teacher support and cooperation, involvement, investigation, or task orientation could suggest that the right pedagogy employed by a teacher which encourages independence, interaction, and inquiry is able to foster deep approach to learning which was proposed by Eley (1992) and Abraham (2006).

Moreover, this study also shows the significant positive relationships between perceptions of computer usage and many other perceptions that are simultaneously related to deep approach significantly such as student cohesiveness, involvement, investigation, and task orientation. These results are coherent with many studies which advocate the impact of these perceptions on deep approach to learning in the ICT-supported learning environment (Basu & Cohen, 1994; Maor, 2000; Law et al., 2002; Siragusa, 2002; de Lange et al., 2003; Jebeile & Abeysekera, 2010; Premuroso et al., 2011; Jones et al., 2013). Thus, the findings could suggest that ICT plays an important role in fostering deep approach to learning.

Finally, the perceptions of assessment which were found to be unrelated to deep approach may reflect that the perceived fact- and memory-oriented assessment could hardly be related to deep learning. The result is in contrast with the studies conducted by Abraham (2006) and Watty et al. (2010) which advocated that deep approach to learning is influenced by the perceptions of assessment.

6. Conclusion and Future Research

This preliminary study found that students demonstrated deep approaches to learning Principles of Accounting in an ICT-supported learning environment. In this vein, deep approach to learning was found to be positively related to various perceptions of ICT-supported learning environment, particularly the perceptions of teacher support and computer usage. It may reflect that a good teaching environment that employs appropriate pedagogy and technology could encourage students to adopt deep approaches to learning. This supports the concept of Technological Pedagogical Content Knowledge (TPCK) (Mishra and Koehler, 2006) which advocates that good teaching require an understanding of how technology constructively relates to pedagogy and content. Thus, technology is not just an “add on” in the teaching and learning process; rather, the skill and art of using technology and the context of its use are the key determinants for stimulating deep approach to learning.

On the other hand, more effort has to be put in to improve the assessment of Principles of Accounting as the current findings show a more fact- and memory-oriented assessment as perceived by students. As advocated by researchers, deep approach to learning is especially important in the learning of accounting because most of the concepts must be mastered through understanding and not memorising (Sukumaran, 1991; Borthick & Clark, 1986). In this vein, students must see the general principles which organize all procedures of accounting into a coherent whole rather than rote learning the procedural rules. Therefore, appropriate assessment which evaluates students' level of meaningful understanding is important for fostering deep approach to learning.

Further research has to be conducted to consolidate the research instrument, especially the items measuring Surface Approach, in order to obtain a more reliable result. On one hand, more samples have to be included for future studies to ensure the representativeness and generalisability of the sample. In addition, more variables have to be involved in predicting approaches to learning as suggested by Biggs (1985) because both the students' personal and situational factors could influence learning approaches. Since the perceptions of ICT-supported learning environment which were studied in the present research are the situational factors, more personal factors have to be examined in future such as students' academic abilities, prior educational experience, and personalities. Finally, by involving more samples and variables in this study, a more sophisticated analysis method such as Multiple Regression Analysis (MRS) or Structural Equation Modelling (SEM) (Bentler, 1980; 1983; Bollen, 1989) can be employed. The MRS is able to identify the best independent variable which can contribute to the deep or surface approach to learning, while SEM has the capability in estimating and testing hypothesised interrelationships among the observed and latent variables.

References

Abraham, A. (2006). Teaching and learning in accounting education: Students' Perceptions of the Linkages between teaching context, approaches to learning and outcomes. In Juchau, R. and Tibbits, G. (Eds.), *Celebrating accounting* (pp.9-21). New South Wales, Australia: University of Western Sydney.

Aldridge, J. M., Dorman, J. P., Fraser, B. J. (2004). Use of Multitrait-Multimethod Modelling to Validate Actual and Preferred Forms of the Technology-Rich Outcomes-Focused Learning Environment Inventory (Troflei). *Australian Journal of Educational & Developmental Psychology*, *4*, 110-125.

Aldridge, J. M., Fraser, B. J., & Huang, T. C. I. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research*, *93*, 48–62.

Arfah Salleh (2001). Integration of computers into accounting curriculum: A revisit. *Akauntan Nasional, Journal of the MIA*, *14*(5): 42-47.

Bahagian Pembangunan Kurikulum [Curriculum Development Section] (2009). *Sukatan Pelajaran Kurikulum Bersepadu Sekolah Menengah: Prinsip Perakaunan* [Syllabus for Integrated Secondary School Curriculum: Principles of Accounting]. Putrajaya: Bahagian Pembangunan Kurikulum, Kementerian Pelajaran Malaysia.

Basu, P., & Cohen, J. (1994). Learning to learn in the accounting principles course: Outcome assessment of an integrative business analysis project. *Journal of Accounting Education*, *12*(4), 359-374.

Beattie, V., Collins, B., & McInnes, B. (1997). Deep and surface learning: A simple or simplistic dichotomy? *Accounting Education: An International Journal*, *6*(1), 1–12.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, *31*, 419-456.

Bentler, P. M. (1983). Simultaneous equations as moment structure models: With an introduction to latent variables models. *Journal of Econometrics*, *22*, 13-42.

Biggs, J. B. (1985). The role of meta-learning in study process. *British Journal of Educational Psychology*, *55*, 185-212.

Biggs, J. B. (1987a). *Student approaches to learning and studying.* Hawthorn, Victoria: Australian Council for Educational Research.

Biggs, J. B. (1987b). *Study Process Questionnaire manual*. Hawthorn, Victoria: Australian Council for Educational Research.

Biggs, J. B., & Moore, P. J. (1993). *The process of learning* (3rd ed.). New York: Prentice Hall.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The Revised Study Process Questionnaire. *British Journal of Educational Psychology*, *71*, 133-149.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Booth, P., Luckett, P., & Mladenovic, R. (1999). The quality of learning in accounting education: The impact of approaches to learning on academic performance. *Accounting Education: An International Journal*, *8*, 277–300.

Borthick, A. F., & Clark, R. L. (1986). The role of productive thinking in affecting student learning with microcomputers in accounting education. *The Accounting Review*, *LXI*(1), 143-157.
Chan, Y. Y. G., & Watkins, D. (1994). Classroom environment and approaches to learning: An investigation of the actual and preferred perceptions of Hong Kong secondary school students. *Instructional Science, 22,* 233-246.
de Lange, P., Suwardy, T., & Mavondo, F. (2003). Integrating a virtual learning environment into an introductory accounting course: Determinants of student motivation. *Accounting Education: An International Journal*, *12*(1), 1-14.
Dorman, J. P. (2003). Cross-national validation of the What Is Happening In this Class? (WIHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research, 6,* 231–245.
Eley, M. G. (1992). Differential adoption of study approaches within individual students. *Higher Education, 23,* 231–254.
Gow, L., Kember, D. & Cooper, B. (1994). The teaching context and approaches to study of accountancy students. *Issues in Accounting Education*, *9*(1), 118-130.
Green, B. P., Reinstein, A., & Mc Williams, D. (2000). Using interactive courseware to teach the procedural components of introductory financial accounting. *Advances in Accounting Education: Teaching and Curriculum Innovations, 3,* 179-197.
Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate Data Analysis* (3rd ed.). New York: Macmillan.
Hanuni Yusuf (2003). *Masalah pembelajaran dan pengajaran Prinsip Akaun* [Teaching and learning problems in Principles of Accounting]. Unpublished master’s thesis. Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia.
Jackling, B. (2005). Perceptions of the learning context and learning approaches: Implications for quality learning outcomes in accounting. *Accounting Education*, *14*(3), 271-291.
Jebeile, S., & Abeysekera, I. (2010). The spread of ICT innovation in accounting education. *International Journal of Teaching and Learning in Higher Education, 22*(2), 158-168.
Jones, A., Scanlon, E., Gaved, M., Blake, C., Collins, T., Clough, G. et al. (2013). Challenges in personalisation: Supporting mobile science inquiry learning across contexts. *Research and Practice in Technology Enhanced Learning, 8*(1), 21-42.
Kember, D, Biggs, J. B., & Leung, D.Y.P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology, 74*, 261–280.
Law, N., Lee, Y., & Chow, A. (2002). Practice characteristics that lead to 21st century learning outcomes. *Journal of Computer-Assisted Learning, 18,* 415-426.
Lillie, R. E., & Wygal, D. E. (2011). Virtual Office Hours (VOH) in accounting coursework: Leveraging technology to enhance an integrative learning environment. *Journal of Accounting Education, 29,* 1-13.
Lizzio, A., Wilson, K., & Simons, R. (2002). University students’ perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education, 27*(1), 27-52.
Looi, C. K., Hung, W. L., Chen, W. L., & Wong, L. H. (2006). A research agenda for fostering deep learning mediated by technologies: Perspectives from Singapore. *The Official Journal of Global Chinese Society for Computers in Education*, *4*(1&2), 85-100.
Lord, B. R., & Robertson, J. (2006). Students' experiences of learning in a third-year management accounting class: Evidence from New Zealand. *Accounting Education*, *15*(1), 41-59.
Lucas, U. (2001). Deep and surface approaches to learning within introductory accounting: a phenomenographic study. *Accounting Education, 10*(2), 161–184.
Maor, D. (2000). A teacher professional development program on using a constructivist multimedia learning environment. *Learning Environments Research*, *2,* 307-330.
Martin, C. (1994). *An introduction to accounting* (4th ed.). Sydney: McGraw Hill.
Marriott, N. (2004). Using computerized business simulations and spreadsheet models in accounting education: A case study. *Accounting Education, 13*(S1), 55-70.
Marton, F., & Säljö, R. (1976).On qualitative differences in learning: II. Outcome as a function of the learner’s conception of the task. *British Journal of Educational Psychology, 46,* 115-127.
Menard, S. (1995). *Applied logistic regression analysis: Sage University series on quantitative applications in the social sciences*. Thousand Oaks, CA: Sage.
Meyer, J. H. F., Parsons, P., & Dunne, T. T. (1990). Individual study orchestrations and their association with learning outcome. *Higher Education*, *20*(1), 67–89.
Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.
Murphy, E. A., & Hoeppner, C. J. (2002). Using technology and library resources in financial accounting courses. *Journal of Accounting Education, 20*, 331-346.

Neal, G. (2005). *Student reflections on the effectiveness of ICT as a learning resource.* Retrieved December 18, 2010, from http://www.aare.edu.au/05pap/nea05582.pdf

Premuroso, R. F., Tong, L., & Beed, T. K. (2011). Does using clickers in the classroom matter to student performance and satisfaction when taking the introductory financial accounting course? *Issues in Accounting Education*, *26*(4), 701–723.

Ramsden, P. (1979). Student learning and perceptions of the academic environment. *Higher Education, 8*, 411–428.

Rashidah Hassan & Arfah Salleh (2008, July). *Issues and challenges in preparing teachers for ICT integration in classroom teaching and learning: The case of Principles of Accounting Sijil Pelajaran Malaysia – SPM (Malaysian Education Certificate – MCE).* Paper presented at the Conference of Vocational and Technical Education and Training, Bali, Indonesia.

Rohaila Yusof (2006). *Keberkesanan strategi pembelajaran pengalaman berasaskan Model Kolb terhadap pembangunan kompetensi pelajaran perakaunan di Institut Pengajian Tinggi Awam* [Effectiveness of Kolb Model-based experiential learning strategy on competency development of accounting students in Public Higher Education Institutions]. Unpublished master's thesis. Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia.

Rumpagaporn, M. W. (2007). *Students' critical thinking skills, attitude to ICT and perceptions of ICT classroom learning environments under the ICT schools pilot project in Thailand* (Doctoral dissertation, University of Adelaide, Australia). Retrieved from http://digital.library.adelaide.edu.au/dspace/bitstream/2440/37896/1/02whole.pdf

Sharma, D. (1997). Accounting students' learning conceptions, approaches to learning, and the influence of the learning-teaching context on approaches to learning. *Accounting Education: An International Journal, 6*(2), 125–146.

Siragusa, L. (2002). *Research into the effectiveness of online learning in higher education: Survey findings*. Retrieved May 6, 2010, from http://www.waier.org.au/forums/2002/siragusa.html

Stanley, T., & Edwards, P. (2005). Interactive multimedia teaching of Accounting Information System (AIS) cycles: Student perceptions and views. *Journal of Accounting Education*, *23*(1), 21-46.

Suhaida Abdul Kadir (2002). *Perbandingan pembelajaran koperatif dan tradisional terhadap prestasi, atribusi pencapaian, konsep kendiri akademik dan hubungan sosial dalam pendidikan perakaunan* [Comparison between cooperative and traditional learning towards performance, performance attribution, academic self-concept, and social relationship in accounting education]. Unpublished doctoral dissertation, Universiti Putra Malaysia, Serdang, Selangor, Malaysia.

Sukumaran, S. (1991). *Strategi pengajaran simpan kira dan perakaunan KBSM* [Teaching strategy for book-keeping and accounting of Integrated Secondary School Curriculum]. Selangor, Malaysia: Fajar Bakti.

Tan, B. S., & Wong, S. L. (2012). Fostering deep approach to learning for Principles of Accounting through ICT integration in Malaysian secondary schools. In G. Biswas et al. (Eds.), *Workshop Proceedings of the 20th International Conference on Computers in Education* (pp.112-118). Singapore: Asia-Pacific Society for Computers in Education.

Technical Education Department (2000). *Huraian sukatan pelajaran Kurikulum Bersepadu Sekolah Menengah: Prinsip Perakaunan Tingkatan 4 dan 5* [Curriculum specifications for Integrated Secondary School Curriculum: Principles of Accounting Form 4 and 5]. Kuala Lumpur: Jabatan Pendidikan Teknikal, Kementerian Pendidikan Malaysia.

Watty, K., Jackson, M., & Yu, X. (2010). Students' approaches to assessment in accounting education: The unique student perspective. *Accounting Education: An International Journal, 19*(3), 219–234.

Wells, P., de Lange, P., & Fieger, P. (2008). Integrating a virtual learning environment into a second-year accounting course: determinants of overall student perception. *Accounting and Finance, 48*, 503–518.

Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and application of the Course Experience Questionnaire. *Studies in Higher Education*, *22*(1), 33-53.

Wong, N. Y., Lin, W. Y., & Watkins, D. (1996). Cross-cultural validation of models of approaches to learning: An application of confirmatory factor analysis. *Educational Psychology*, *16*, 317–327.

Technology-enhanced Chemistry Learning and Students' Perceptions: A Comparison of Microcomputer-based Laboratory and Web-based Inquiry Science Environment

Kulthida KAMTOOM[a], Niwat SRISAWASDI[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*

**niwsri@kku.ac.th*

Abstract: With current advancement of information and communication technology, researchers have indicated the necessity and challenges of developing effective instructional strategies by the use of technological tools to promote chemistry learning in school science. Consequently, the researchers has developed two kinds of technology-enhanced chemistry learning environment regarding rate of reaction concepts: MBL-based guided-inquiry learning, as face-to-face instruction; web-based inquiry science environment (WISE), as online instruction. The purpose of this study was to investigate effects of both learning environment on students' perceptions. The participants were tenth graders; they are assigned randomly into two groups for interacting with different learning environment, MBL (N= 37) and WISE (N= 34). To assess students' perceptions in the different groups, pre- and post-questionnaires were used. The results of the study (N = 71) show that there was statistically significant difference between pre - and post-questionnaire for both groups. No significant difference was found between the MBL and WISE groups in terms of students' perceptions. The result suggested that both technology-enhanced learning environment progressed students' perceptions of learning and the progression was not indifferent. This finding highlight the benefits of face-to-face MBL and WISE to promote students' perceptions of learning.

Keywords: MBL, WISE, chemistry learning, perception, rate of reaction

1. Introduction

To advance the practice of science education, technologies have become commonplace because of their potential of bringing about change in ways of teaching and learning (Srisawasdi, 2012). For this reason, in research and development of the student learning process in science, the effective use of technology in the classroom teaching process has become a common place in science-based education. In the community of science educator and educational researcher, it has been widely established that contemporary science learning environments should foster inquiry-based experiences and investigations. These inquiry-based scientific practices should take place in the laboratory in order to interact directly with naturally occurring phenomena (Srisawasdi, 2014; Pyatt & Sims, 2011). In recent years, many researchers have applied microcomputer-based laboratory into inquiry-based science education for teaching and learning (Friedler et al., 1990; Gunhaart & Srisawasdi, 2012; Liao & She, 2009; She & Liao, 2009) that can promote authenticity and application of scientific process and practice and facilitate concrete-to-abstract conceptualization of real-world phenomena (Pyatt & Sims, 2011). Likewise, Web-based Inquiry Science Environment (WISE) is an open-source digital learning platform that supports student inquiry in classrooms for support students visualization and understand in sub-microscopic levels (Chiu, Jennifer & Linn, 2011). Researchers reported that students could be provided effective learning experience of science and student can link the idea of three levels of chemistry representation (i.e. macroscropic, sub-microscropic, and symbolic) (Chang & Linn, 2013). As abovementioned, both science learning environments were recognized as effective technological tool for inquiry-based learning process, and there is a plenty of challenge to think about how to use them in science teaching and learning pedagogically.

Accordingly, the aim of this study was to investigate students’ perceptions delivered in face-to-face MBL and online-mediated WISE. Specifically, the following questions were answered:

(1) Do the students engaged in face-to-face MBL and online-mediated WISE learning environment perform significantly better by students’ perceptions?

(2) Is there a difference between students' perceptions between face-to-face MBL and online-mediated WISE learning environment?

Following the questions, it is hypothesized that both face-to-face MBL and online-mediated WISE will lead greater students' perception after their participations, and there are no different on students' perceptions on experience with both learning environments.

2. Literature review

2.1 Microcomputer- based Laboratory (MBL)

Recently, In teaching and learning about of science-based education, computer technology has been widely used for modeling scientific knowledge structures or learning patterns, developing more in-depth and integrated understanding of concepts and process, enhancing the development of scientific skills, visualizing complex and dynamic scientific phenomena, promoting collaborative network in the community of learning for the construction of knowledge, and sharing of data, support access to a variety of information, support collecting various types of scientific data, test underlying theories through diagnostic or tutorial strategies, and enhance characteristics of inquiry as the way scientists work (Srisawasdi, 2014). In recent years, the reviewing of empirical evidence it is clear that microcomputer can improve the learning process of students in science education. Moreover, many researchers have applied microcomputer-based learning to inquiry-based science education for teaching and learning (Friedler et al., 1990; Gunhaart & Srisawasdi, 2012; Liao & She, 2009; She & Liao, 2009). The result of researchers shown that microcomputer-based inquiry learning can support student to learn science efficiency and make student can observe phenomena and understand in macroscopic level of representation.

2.2 Web-based Inquiry Science Environment (WISE)

WISE is the multimedia learning, in which the teachers can insert the lessons into web-based learning environment in order to aid and follow the learning results from students. WISE has four steps following Knowledge Integration (KI) framework (Chiu, Jennifer, & Linn, 2011): (i) eliciting ideas is to emphasize the native knowledge of the students; (ii) adding ideas is to add the new knowledge to the students; (iii) distinguishing ideas is to promote critical thinking of the students; and (iv) sorting out ideas is to rearrange the ideas of the students. In an addition, researcher reported impact of the WISE that students can link the idea of three levels of chemistry knowledge (i.e. sub-microscropic, macroscropic, and symbolic (Chang & Linn, 2013).

3. Method

3.1 *Participants*

The participants in this study were tenth-grade students (16-17 years old) in the Northeastern region of Thailand. The selection of samples performed randomly. They were divided into two groups which received different learning environment: WISE (N=34) and MBL (N=37) groups. Both groups were assigned to learn a chemistry lesson on rate of reaction. For online-mediated WISE class, the students were enabled to learning by self-regulated learning process with the WISE, which was designed the rate of reaction lesson followed KI framework. In contrast, the students in face-to-face MBL class were engaged into guided-inquiry learning process with the support of MBL tools. However, two technology-based learning environments were equivalent in term of chemistry concepts on rate of reaction. The students had not had any formal education on rate of reaction before the study took place. Figure 1 illustrated information about participants and learning environment.

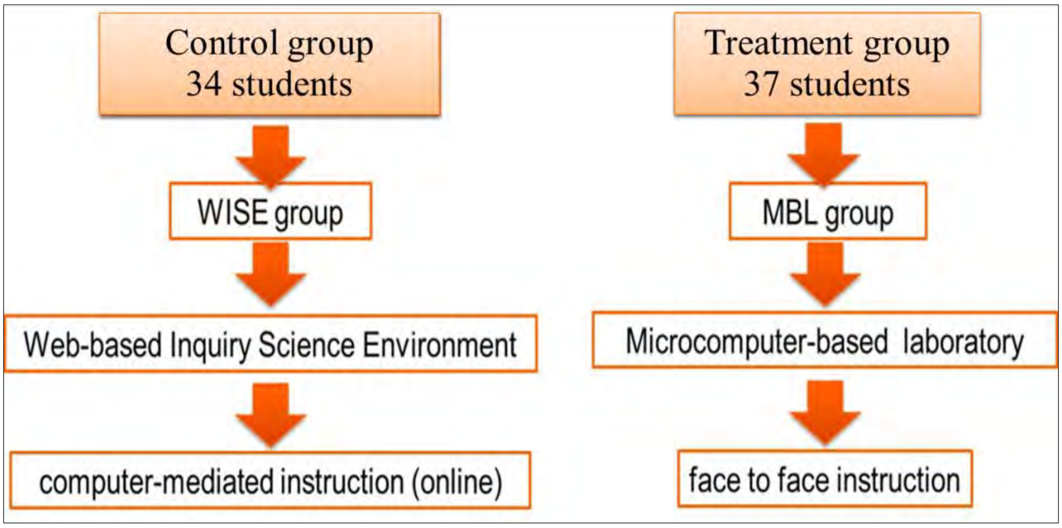


Figure 1. Diagram of participants and learning environment

3.2 *Learning Environments*

3.2.1 *Face-to-face MBL*

To engage student into face-to-face MBL, this study employed microcomputer-based laboratory (MBL) by Vernier and software technology. Vernier mini-LabQuest is a digital interface used to collect sensor data and then transfer data into Logger Pro software building graph and analysis application. For this study, students were provided opportunity to conduct an investigation of effects of reaction rate using carbon dioxide gas sensor. Figure 2 displays an example of hands-on MBL for guided-inquiry learning process in chemistry class.

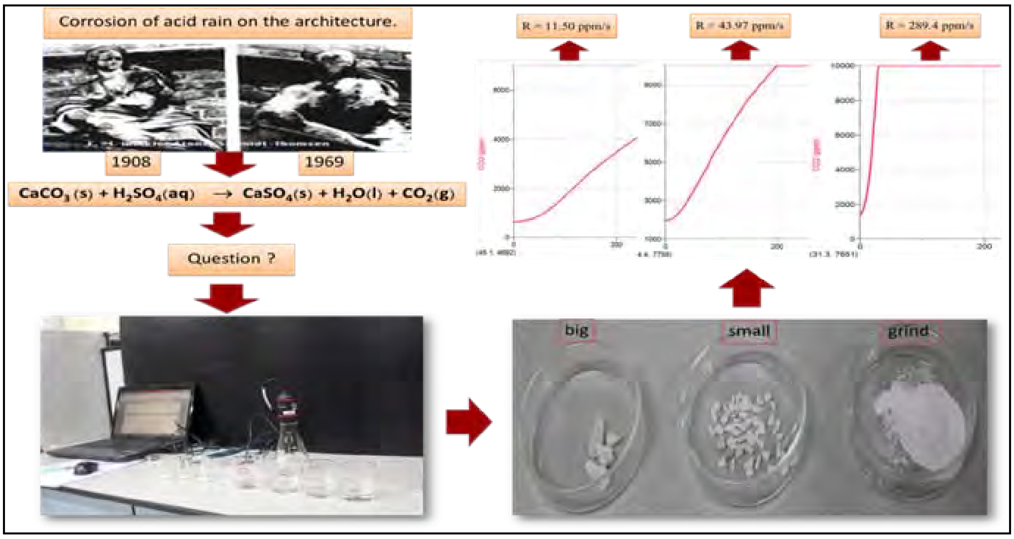


Figure 2. Example of hands-on guided-inquiry MBL on the effects of rate of reaction

As mentioned, guided-inquiry learning process was used to promote student's chemistry learning of rate of reaction. The simulation-based open inquiry with dual-situated learning model (Srisawasdi and Kroothkeaw, 2014; Srisawasdi and Sornkhatha, 2014) was applied for the face-to-face MBL learning process in a manner of guided inquiry. Table 1 displayed an example of the learning process on rate of reaction concept.

Table 1: The six-stage guided-inquiry learning process for face-to-face MBL environment

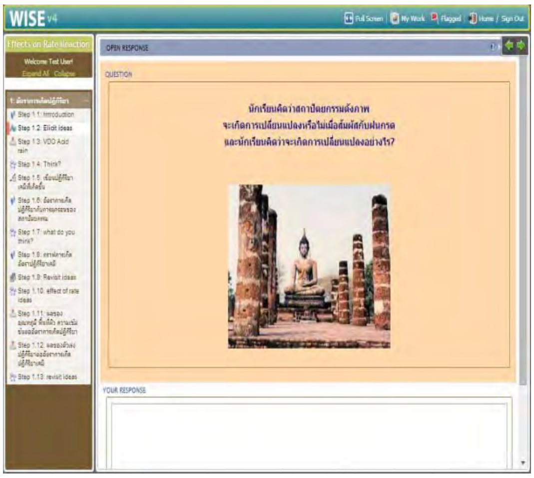
Stage	Learning process
Pre-lab	
Open-ended inquiry question	Teacher provides an open-ended inquiry question: ‘‘How about rate of reaction if difference surface areas of sculpture react with acid rain is concentrated static?’’
Scientific background/ information	Teacher induces collaborative discussion toward the definitions and pictorial diagram of acid rain reaction ;surface area ;CO_2 gas
Lab practice	
Procedure/design	Teacher introduces the overall procedure of the laboratory, explains how to perform a predetermined experiment step-by-step, and guides how to record necessary data. Then, student conducted the experiment about acid rain (H_2SO_4) with sculpture ($CaCO_3$) interaction for collecting the experimental data from carbon dioxide gas when they change surface area of $CaCO_3$.
Data and result analysis	After the interacting with laboratory, students make a decision to analyze obtained experimental data from their own design and interpret it into results.
Post-Lab	
Result communication	Students have to select the way to present, communicate, and discuss the meaning of data and experimental results to others, for an example, writing experimental question, experimental design, results, and discussion on a newsprint paper and then present to the class.
Conclusion	Students have to collaboratively make a relationship between each group results and then draw it into a conclusion as the best answer to the provided inquiry question. For an example, teacher induces each group make a conclusive answer by using an integration of their own and other results.

In this learning process, student will collaborative work together in small groups of three to five members. This pedagogy begins with an open-ended driving question targeted to alternative conceptions about effects of rate of reaction commonly found in students. To assist the process of hypothesis generation addressed the question, essential scientific backgrounds are provided to students. Then, students are required to perform generating testable hypotheses. Information about scientific process and procedures were given to guide them performing experimentation with MBL tools. During experimenting with the MBL, instructor came to probe them with a series of formative assessment question. When they finished the experiment, all group has to communicate findings among groups by creating a chart of results and explaining findings. Finally, instructor induces students into a forum for drawing a conclusion based on evidence and collaborative explaining the result of hypotheses testing. Figure 3 displays an example of students’ interaction in the context of hands-on guided-inquiry MBL.

Figure 3. An illustrative example of classroom learning activity through face-to-face microcomputer-based laboratory with guided-inquiry process.

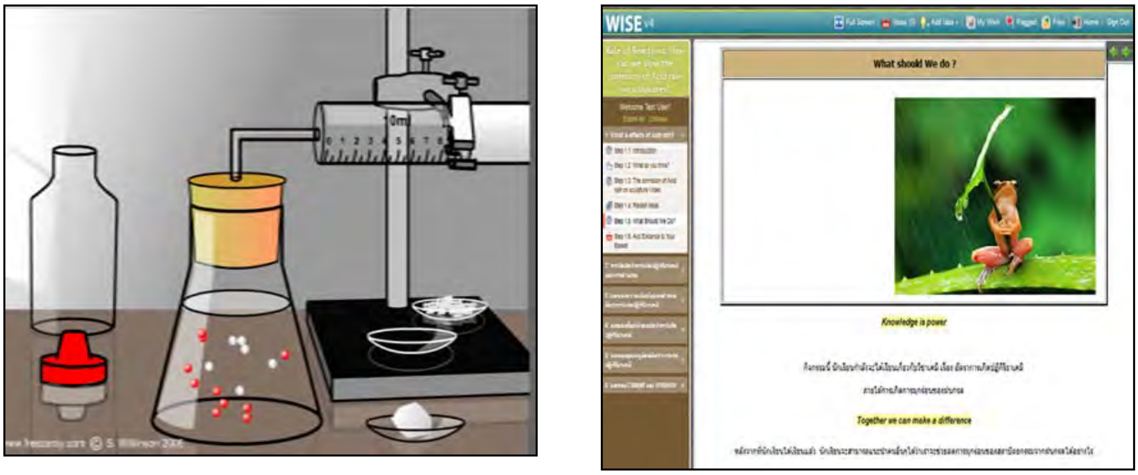
3.2.2 *Online-mediated WISE*

WISE curriculum materials is an online-mediated environment which features powerful, interactive scientific visualizations to illustrate unobservable phenomena such as chemical reactions at the atomic level, large-scale phenomena such as climate change, and phenomena that happen quickly such as collisions. The online environment logs student interactions and responses to embedded assessments. WISE offers teachers grading and classroom management tools to monitor student progress, provide feedback to students, and personalize instruction. In this study, the researchers created a chemistry learning unit on rate of reaction on WISE system by emphasizing the interplay between macro-, sub-micro-, and symbolic representations of chemistry. The learning unit was constructed following Knowledge Integration (KI) framework. Figure 4 displays screen interface of WISE system for the chemistry learning unit of rate of reaction.



(A) Elicit Ideas: asking students what happen in this picture? for eliciting their existing knowledge of the students

(B) Add idea: giving student necessary information for adding new related knowledge to existing ideas



(C) Distinguishing idea: using of simulation and animation to promote students' scientific thinking and reasoning

(D) Sorting out idea: Asking students a question, “What factors affect the weathering of architecture from acid rain?” to rearrange conceptual ideas

Figure 4. An illustrative examples of screen interface of the chemistry learning unit of rate of reaction in online-mediated WISE following KI framework, from (A) to (D)

In this learning process, student will collaborative work together in small groups of two or three members. The teacher gave project ID to students in order to access the WISE and then facilitate students' group learning in the class and monitor progression of learning process during their participations. Finally, instructor induces students into a forum for drawing a conclusion based on evidence and collaborative explaining the rate of reaction phenomena. Figure 5 displays an example of students’ interaction in the context of online-mediated WISE.

Figure 5. An illustrative example of classroom learning activity through online-mediated WISE.

3.3 Instruments

A 16-item Likert-scale questionnaire was developed to use in this study for examining students’ perceived ease of use (3 items), perceived usefulness (3 items),and perceived satisfaction (3 items), obtained from Barzilai & Blau (2014), while an instrument by Cheng (2014) was used to measure perceived learning (4 items), and enjoyment (3 items). To develop a Thai version of the questionnaire, the original English version was translated identically in Thai language. Two experts were recruited to identify communication validity of the items. On each item, respondents were assigned to rate how much the respondent agree with into five scale, from 1-strongly disagree to 5-strongly agree. The instrument had been established validity and reliability. Table 2 displayed sample items of the questionnaire used in this study and its reliability.

Table 2: Sample items of the Likert-scale questionnaire measured students' perceptions

Dimension	Sample items		α
	WISE	MBL	
Perceived ease of use	▪ It is easy for me to learn how to use WISE. ▪ I can easily accomplish what I need to do in WISE.	▪ It is easy for me to learn how to use MBL. ▪ I can easily accomplish what I need to do in MBL.	0.737
Perceived usefulness	▪ WISE can help me learn more effectively. ▪ WISE can improve my course performance.	▪ MBL can help me learn more effectively. ▪ MBL can improve my course performance.	0.842
Perceived satisfaction	▪ I feel comfortable to use WISE. ▪ I enjoy the experience of using WISE.	▪ I feel comfortable to use MBL. ▪ I enjoy the experience of using MBL.	0.774
Perceived learning	▪ I learned a lot from the WISE. ▪ The WISE added to my knowledge.	▪ I learned a lot from the MBL. ▪ The MBL added to my knowledge.	0.803
Enjoyment	▪ I enjoyed the WISE. ▪ Interacting the MBL was pleasant.	▪ I enjoyed the MBL. ▪ Interacting the MBL was pleasant.	0.754

3.4 Data Collection and Analysis

Students were investigated perceptions by the 5-point Likert-scale questionnaire before giving face-to-face MBL and online-mediated WISE intervention for 10 minutes. Both learning environments, students participated to interact with them for 60 minutes. After the instruction, students were administered by the same questionnaire again as post-test. The statistical data techniques selected for analyzing students' perceptions was repeated-measures MANOVA in SPSS 21.0.

4. Results and Discussion

The results for the repeated-measures MANOVA indicated no significant main effect for group (Wilks' lambda=0.878, $F(5, 65) = 1.798$, $p=0.126$, partial $\eta^2 = 0.122$). There was no significant difference on students' perceptions between face-to-face MBL and online-mediated WISE groups. The univariate results on two group revealed none of the five subscales on perception reached a statistical significance between both groups. That is, both face-to-face MBL and online-mediated WISE group performed indifferently with regard to perceived ease of use, perceived usefulness, perceived learning, perceived satisfaction, and enjoyment. Also, there was no significant interaction effect between group and time (Wilks' lambda=0.934, $F(5, 65)=0.923$, $p=0.472$, partial $\eta^2 = 0.066$). This means that both face-to-face MBL and online-mediated WISE has similar effects on students' perceptions. However, there was a significant main effect for time (Wilks' lambda=0.567, $F(5, 65)=9.940$, $p<0.001$, partial $\eta^2 = 0.433$). This suggests that, on average, students' perceptions have changed over the learning environments. Univariate analyses of variances (ANOVA) on each subscale were conducted as follow-up tests to the one-way MANOVA. The results of the univariate test for face-to-face MBL and online-mediated WISE groups are summarized in Table 3.

Table 3: The students' subscale means of perceptions by group and time and univariate MANOVA by time

Subscale	MBL		WISE		F	Sig.	η^2
	Pre-test	Post-test	Pre-test	Post-test			

Perceived ease of use	8.38(1.97)	11.19(1.71)	9.71(1.95)	11.35(1.97)	43.741	0.000	0.388
Perceived usefulness	8.87(2.64)	11.54(2.00)	9.97(2.36)	11.62(1.84)	34.515	0.000	0.333
Perceived learning	11.35(3.13)	15.00(2.43)	13.47(3.06)	15.15(2.41)	31.124	0.000	0.311
Perceived satisfaction	11.14(3.13)	14.84(2.27)	13.09(3.06)	15.47(2.43)	42.716	0.000	0.382
Enjoyment	8.30(2.52)	10.87(2.11)	9.94(2.53)	11.32(1.89)	25.807	0.000	0.272

As can be seen in Table 3, The univariate MANOVA on the five subscale scores of perceptions were significant differences across time, from pre-test to post-test. The univariate results revealed a significant effect on perceived ease of use ($F_{1,69} = 43.741$, $p < 0.01$, partial $\eta^2 = 0.388$), perceived usefulness ($F_{1,69} = 34.515$, $p < 0.01$, partial $\eta^2 < 0.333$), perceived learning ($F_{1,69} = 31.124$, $p < 0.01$, partial $\eta^2 = 0.311$), perceived satisfaction ($F_{1,69} = 42.716$, $p < 0.01$, partial $\eta^2 = 0.382$), and enjoyment ($F_{1,69} = 25.807$, $p < 0.01$, partial $\eta^2 = 0.272$). According to aforementioned results, the overall result suggested that the increase of students' perceptions regarding perceived ease of use, perceived usefulness, perceived learning, perceived satisfaction, and enjoyment from the pre-test to post-test was homogeneous both face-to-face MBL and online-mediated WISE group. This result consistent with findings that students perceived online learning environment easy to use, to learn, and to accomplish what need to be done (Barzilai & Blau, 2014). In an addition, it confirm with previous studies (Voogt, Tilya and van den Akker, 2009) that students were satisfied with hands-on MBL experience for getting more investigative and open-ended learning, and user friendly.

5. Conclusion

The results revealed that there is no interaction effect on group (face-to-face MBL and online-mediated WISE) and time (pre- and post-test) regarded perceived ease of use, perceived usefulness, perceived learning, perceived satisfaction, and enjoyment. This means that the increasing of students' perceptions did not depend on different technology-enhanced learning environments e.g. face-to-face MBL and online-mediated WISE. This finding highlights the importance of the need for considering effective instructional design and environment to enhancing chemistry teaching and learning. For the chemistry class, both face-to-face MBL and online-mediated WISE group could be used to facilitate teaching and learning of chemical concepts and support emotional practice in chemistry education. For the next study, the authors have a plan to implement both learning environments for investigating whether if it would be more beneficial to combine WISE and MBL activities than to use them separately in teaching rate of reaction.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, *70*, 65-79.

Caspi, A., & Blau, I. (2011). Collaboration and psychological ownership: how does the tension between the two influence perceived learning?. *Social Psychology of Education*, *14*(2), 283-298.

Chang, H., & Linn, M. (2013). Scaffolding learning from molecular visualizations. *Journal of Research In Science Teaching*, *50*(7), 858-886.

Cheng, G. (2014). Exploring students' learning styles in relation to their acceptance and attitudes towards using Second Life in education: A case study in Hong Kong. *Computers & Education*, *70*, 105-115.

Drennan, J., Kennedy, J., & Pisarski, A. (2005). Factors affecting student attitudes toward flexible online learning in management education. *The Journal of Educational Research*, *98*(6), 331-338.

Friedler, Y., Nachmias, R., & Linn, M. (1990). Learning scientific reasoning skills in microcomputer-based laboratories. *Journal Of Research In Science Teaching*, *27*(2), 173-192.

Gunhaart, A., & Srisawasdi, N. (2012). Effect of integrated computer-based laboratory environment on students' physics conceptual learning of sound wave properties. *Procedia - Social and Behavioral Sciences, 46*, 5750-5755

Johnson, M., & Lawson, A. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes?. *Journal of Research in Science Teaching*, *35*(1), 89-103.

Lawson, A.E. (1994). Science Teaching and the Development of Thinking. Belmont, California, United States: Wadsworth Pub Co.

Lawson, A.E.;1; (1995). *Science Teaching and Development of Reasoning. Belmont, CA: Wadsworth.*

Liao, Y., & She, H. (2009). Enhancing Eight Grade Students' Scientific Conceptual Change and Scientific Reasoning through a Web-based Learning Program. *Educational Technology & Society*, *12*(4), 228-240.

Linn, M., Layman, J., & Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Contemporary Educational Psychology*, *12*(3), 244-253.

Pyatt, K., & Sims, R. (2012). Virtual and physical experimentation in inquiry-based science labs: Attitudes, performance and access. *Journal of Science Education and Technology*, *21*(1), 133-147.

Srisawasdi, N., & Kroothkeaw, S. (2014). Supporting students' conceptual development of light refraction by simulation-based open inquiry with dual-situated learning model. *Journal of Computers in Education, 1*(1), 49-79.

Srisawasdi, N. (2014). Developing Technological Pedagogical and Content Knowledge in Using Laboratory Environment: An Arrangement for Science Teacher Education program. *Research and Practice in Technology Enhanced Learning, 9*,123-143.

Voogt, J., Tilya, F. & van den Akker, J. (2009). Science Teacher Learning of MBL-Supported Student-Centered Science Education in the Context of Secondary Education in Tanzania. Journal of Science Education and Technology, 18(5), 429-438.

Weld, J., Stier, M., & McNew-Birren, J. (2011). The Development of a Novel Measure of Scientific Reasoning Growth among College Freshmen: The Constructive Inquiry Science Reasoning Skills Test. *Journal of College Science Teaching, 40*(4), 101-107.

The Design of Instructional Scaffolds to Facilitate Online Project-Based Learning

Chun-Ping WU[a*], Ching-Chiu YEH[b], Shu-Ling WU[c] & Hao Jie YONG[c]
[a]*Department of Educational Technology, TamKang University, Taiwan*
[b]*Department of Educational Technology, TamKang University, Taiwan*
[c]*Graduate Institute of Human resource and eLearning Technology, National Hsinchu University of Education, Taiwan*
[c]*Department of Educational Technology, TamKang University, Taiwan*
*cpwu303@gmail.com

Abstract: The study designed scaffolds to facilitate online project-based learning. The question-generation, worked examples and peer-assessment were adopted and transformed into the scaffolds. The integration of the scaffolds into the project process was exemplified in a distance course, entitle as *instructional design*. Thirty college students participated in the study for 18 weeks. The Moodle as well as the Adobe-connect learning system were adopted to engage learners with the contents, scaffolds and team collaboration. Data analysis found that the scaffolds facilitated participants' engagement with the learning contents and team interaction. Empirical significance of the study as well as suggestions for instructional implementation and future study are provided.

Keywords: Online project-based learning, Worked examples, Scaffolds, Peer-assessment

1. Introduction

Creating instructional products with virtual teams to effectively engage learners in active knowledge creation by maximizing the merits of advanced interactive learning technology is very complex (Perez & Emery, 1995). Instructional designers' cognitive abilities to integrate multiple domains of knowledge and collaboration skills to bring together the diverse expertise of each team member at the distant are highly demanded. The online project-based learning strategy is frequently adopted to engage student designers in deliberately practicing the above-mentioned process (Tynjala, 1998).

However, the inherent openness of project-based learning environment might introduce even more complex cognitive tasks which might increase student designers' cognitive loads. Without systematic guidance to facilitate learners in constructing essential knowledge bases, student designers often fail to execute relevant domain knowledge in analyzing the project scenarios. Moreover, they are often not able to efficiently manage project progress and exchange ideas among virtual team members, which might further result in the incompleteness of project tasks. Therefore, design appropriate scaffolds to provide support on a conceptual level and virtual team interaction are essential.

This study aims to design the scaffolds for implementing project-based learning in the distance course of instructional design. More specifically, the design of the worked example, question generation strategy and peer-assessment as the scaffolds is proposed. Students' engagement in the learning process and interaction among peers are investigated.

2. Literature Review

Project- based learning, which is rooted in the idea of learning by doing, enables student designers to practice solving diverse design problems, synthesize theories and skills, and further construct their knowledgebase (Tynjala, 1998). Its effect is influenced by three design elements: (1) The relevancy of the project scenario with the real-world design problems: The more the project simulates the real-world problem, the more complex of the project scenario will be. The complexity of the project brings student designers more challenge; meanwhile, it might also enhance their curiosity and intrinsic goal orientation (Bednar, Cunningham, Duffy, & Perry, 1992; Cobb, 1994; Karagiorgi, & Symeou, 2005). (2)Systematic guidance: The project usually presents learners with ill-structured design problems, requiring multiple thinking dimensions, which in turns result in student learners heavy cognitive loads. Without adequate and timely guidance, learns might easily give up their projects especially while working with peers at the distant on the virtual project. Moreover, student designers, lacking of the experience in working on ill-structure problems, tend to think of the design problem from single perspective and over-emphasize on the project product instead of the project process. (3) Team facilitation: Student designers' abilities to construct their cognitive structures are enhanced by social interaction with peers, which includes asking questions, explaining their reasoning for the design ideas, listening to peers' ideas and reconciling differences between their ideas and the ideas of others (O'Donnell & King, 1999). However, collaboration levels in the virtual projects are difficult to control, which leads to the lack of clarity of individual responsibilities and leadership and low intensity of interaction among the virtual group members. Therefore, a sound design of scaffolds is essential to engage student designers in reasoning through the problem from multiple perspectives, provide challenge and encourage more creative design ideas (Tynjala, 1998) and facilitate them in working on virtual projects.

Two essential types of scaffolds to facilitate online project-based learning are conceptual scaffolds and interaction scaffolds. The support of the conceptual scaffold using the worked example could effectively reduce a novice designer's cognitive load in analyzing and solving a new ill-structured problem (Sweller, 1998; Sweller, 2005; Turner, Meyer, Cox, Logan, DiCintio, & Thomas, 1998; van Merriënboer, Kirschner, & Kester, 2003). However, the reduction of the extraneous cognitive loads does not necessarily initiate their devoting more working memory to exploring the complex structures represented in the worked examples and the underlying principles and knowledge associated with the given problem. In other words, an effective worked example should engage student designers in executing working memory to simulate expert thinking process and reason through the solution path. Moreover, the cognitive load brought by the project tasks will change according to a designer's expertise. A designer, who is able to transfer the learned knowledge and skills into schema, could efficiently focus on the design/solution path and underlying knowledge structures presented in the worked example. Too much guidance might bring them extra cognitive load (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga & Sweller, 2004), Therefore, the conceptual support should gradually fade out with student designers' development of the expertise.

Scaffolding effects of question-generation and peer-assessment strategies can be reasoned based on social learning and constructivism (Yu, & Liu, 2005; van Gennip, Segers & Tillema, 2009). On one hand, the effects of these two strategies in enhancing learners' comprehension of the contents are evidenced in many empirical studies. On the other hand, the high cognitive abilities demanded by the collaborative question-generation task might contribute to high density of the interaction within virtual teams and encourage them to share and construct knowledge. Moreover, the peer-assessment process enables student designers to observe peers' products, compare their own design solutions with those designed by others, which might lead to modification of their products and their own cognitive structures.

Several research gaps after in-depth literature review were identified: First, the linkage between the worked examples and novice instructional designers' engagement in the activity of constructing their knowledge base in the distant learning environment remains less known. Second, the scaffolding effects of question-generation and peer-assessment strategies on cognitive abilities can be reasoned based on social learning and constructivism; however, their effects on the team interaction require more empirical evidence. Last, but the most importantly, reflection is importance in developing instructional designers' competency. Majority of the research emphasizes on the role of "self" in the reflection process (Bilinski, 2002; Johns, 1995; Schmidt, 2004). However, argued from the collective intelligence perspective, that peer assessment and dialogue could enhance more deliberate and collaborative reflection is reasonably anticipated (Sluijsmans, Moerkerke, van Merriënboer, & Dochy, 2001; Sluijsmans, Brand-Gruwel, & van

Merriënboer, 2002). Therefore sound design of these scaffolds emphasizing on peer collaboration in the distant course is worthy of more efforts.

3. Research Method

3.1 Research Design

Thirty college students, who registered in the distant course, entitle as *instructional design*, participated in the study for 18 consecutive weeks. The asynchronous features of the Moodle learning system and the synchronous features of the Adobe-Connect were adopted in this study to engage learners with the learning activities, contents, scaffolds etc. To ensure that participants possessed the fundamental skills of using the communication technologies embedded in the learning systems, a training session on integrating the technologies for project communication and the follow-up question-generation and peer assessment with hands-on activity was arranged.

For the duration of the study, students attended face-to-face sessions at the first two weeks to get familiar with the learning path of the overall course, their peers and the learning technologies. The learners were randomly assigned to six teams. In the following 16 weeks, the class moved on to an asynchronous format. During the weekly online learning activity, students watched the video explaining the learning contents individually, followed by a team mission to be completed. Each member within the team took turns on leading the team to accomplish the team mission. In addition to the weekly team mission, each team was required to an instructional product as their course project. At the last class, they moved on to individually assess the instructional product developed by other teams.

The design blueprint of the scaffolds was developed and implemented. Two data sources are collected to explore the impact of the scaffolds. First, the frequency of participants' accessing the systems was used to evidence levels of their engagement with the course content and the scaffolds. Second, the number and quality of the posting within each team were presented and analyzed weekly.

4. Results and Conclusions

4.1 Design of the Scaffolds

Two types of scaffolds to facilitate online project-based learning were designed and exemplified in Table 1. The purposes and implementation of the scaffolds associated with the four stages of online project-based learning were described.

Table 1. Design blueprint of the scaffolds

Learning stage	Aim	Scaffold	Levels
1.Project initiation	Associate the course progress with the project process	Learning Map	Individual
	Introduce the project	Project Descriptor	Group
2-1.Knowledge building	Ensure the knowledge base for the project	Question-Generation	Individual
2-2.Team Building	Enhance team interaction	Question-Assessing	Group
3-1.Problem Analysis	Exemplify the analysis process	Worked Examples	Group
3-2.Research /Development	Exemplify the research/development process	Worked Examples	Group
4.Solution Evaluation	Seek feedback	In-field testing	Group

	Enhance evaluation ability	Peer-assessment	Individual

1. Project initiation stage: Learning maps were designed to match the project tasks and requirement with the weekly learning contents. This scaffold does not only enable individual learner to link learning contents with the learning outcomes but also scaffold learners to plan their learning path. The second scaffold introduced is the project descriptor. This study created an animation to present different project scenarios and the e-tutor within the animation guided the learners to explore the information inherent in the scenario. Teams were asked to reach consensus in selecting one scenario as their term project based on their exploration of the scenarios.
2. Knowledge and team building stage: The scaffold of question-generation was implemented to ensure individual's comprehension of the course contents by asking each learner to generate one multiple-choice question. The weekly leaders of the teams have to facilitate their team discussion to modify the generated questions.
3. Problem analysis, research and development stage: The scaffold of worked example with e-tutor exemplified experts' reasoning process and guided learners to think of multiple perspectives and synthesize the required domain knowledge.
4. Solution evaluation stage: Each team conducted a field-testing of their design products and revised before in-class peer review. During the peer review process, each individual evaluated the assigned product according to the given criteria and provided written recommendation. After the peer-assessment, the team members were required to read and interpret the comments and write a reflection on product improvement.

Learners' engagement with the learning content is evidenced in 798 times of visiting the content weekly. More specifically, the frequencies of each learner interacting with the contents during the period of introducing the scaffolds of learning maps, project descriptors, question-generation and worked-examples are 26.6, 10.4, 13.2, and 10.53, respectively.

Table 2. The frequency of learners' interacting with course contents weekly

	Learning Maps	Project Descriptor	Question-Generation	Worked-Example
Weekly hits	798	312	396	316
Weekly hits per learner	26.6	10.4	13.2	10.53

Team interaction is revealed in 743 postings supplied by the learners and the 959 times of visiting the posting during 18 weeks. More specifically, the weekly conversations within teams occurring during the period of introducing question-generation scaffolds range from 8 to 39 iterations while the team conversations occurring during the period of introducing the worked example scaffold range from 1 to 25 iterations (Please see Table 3).

Table3. The weekly conversations within the team weekly

	No. of subjects	Question-Generation	Worked-Example	Peer-Assessment
Team1	6	20	7	12
Team2	4	12	2	11
Team3	4	28	8	14
Team4	5	8	1	18
Team5	5	12	2	11
Team6	6	39	25	12
Total	30	119	45	78

In addition to the number of postings, the team interaction should also account for the frequency of reading the postings, which can represent inherent cognitive interaction within the team. Table 4 presents the weekly frequency of reading the postings within teams during the period of introducing question-generation scaffolds, ranging from128 to 536 times. More interestingly, the weekly frequency of reading the comments provided by their peers during the peer-assessment activity ranges from92 to 217.

Table4. The weekly frequency of reading Postings and replied postings within the team

	No. of subjects	Question-Generation	Worked-Example	Peer-Assessment
Team1	6	300	86	217
Team2	4	233	243	114
Team3	4	298	82	181
Team4	5	128	21	192
Team5	5	150	33	92
Team6	6	536	348	185
Total	30	1644	613	981

4.2 Conclusion

The study designed scaffolds to engage student designers in conducting instructional projects with virtual teams. First, the results imply that the designed scaffolds could facilitate student designers' engagement with constructing their knowledge base in the distant learning environment. Second, the scaffolding effects of question-generation and peer-assessment strategies on the team interaction are also evidenced.

The findings have important empirical significance as well as implications for research on online project-based learning. First, the study substantiated the effect of question-generation on learners' mastery of the content Second, the worked-example scaffold which simulates experts' reasoning process, facilitates students in synthesizing knowledge associated with solving the given problem, which is the key factor to the success of project-based learning. Finally, the team interaction was enhanced by the scaffolds. In-depth and systematic analysis of the conversations between interacting parties and learners perceptions toward group dynamic via interaction process analysis is recommended for future study.

References

Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1992) Theory into practice: How do we link? In T. M. Duffy & D. H. Jonassen (Eds.) *Constructivism and the technology of instruction: A conversation* (pp.17- 34). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bilinski, H. (2002). The mentored journal. *Nurse Educator*, *27*, 37-41.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, *23*(7), 13-20.

Johns, C. (1995). Framing learning through reflection within Carper's fundamental ways of knowing in nursing. *Journal of Advanced Nursing*, *22*, 226-234.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. E*ducational Psychologist*, *38*, 23–31.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Education & Psychology, 96*, 558–568.

Karagiorgi, Y., & Symeou, L. (2005). Translating constructivism into instructional design: Potential and limitations. *Educational Technology & Society, 8*(1*)*, 17-27.

O'Donnell, A. M., & King, A. (Eds.) (1999). *Cognitive perspectives on peer learning*. Mahwah, NJ: Laurence Erlbaum.

Perez, R. S., & Emery, C. D. (1995). Designer thinking: How novices and experts think about instructional design. *Performance Improvement Quarterly, 8*(3), 80-95.

Schmidt, L. (2004). Evaluating the writing- to-learn strategy with undergraduate nursing students. *Journal of Nursing Education*, *43*, 466-473.

Sluijsmans, D. M. A., Moerkerke, G., van Merriënboer, J. J. G., & Dochy, F. J. R. C. (2001). Peer assessment in problem-based learning. Studies in Educational Evaluation, 27, 153–173. .

Sluijsmans, D. M. A., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, *27*, 443–454.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge, UK: Cambridge University Press.

Tynjala, P. (1998). Traditional Studying for Examination Versus Constructivist Learning Tasks: Do Learning Outcome Differ? *Studies in Higher Education*, *23*, p.173–189.

Turner, J. C., Meyer, D. K., Cox, K. E., Logan, C., DiCintio, M., & Thomas, C. T. (1998). Creating contexts for involvement in mathematics. *Journal of Education & Psychology, 90*, 730–745.

van Gennip, N. A. E., Segers, M. & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal and structural features. *Learning and Instruction, 4* (1), 41–54.

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load of a learners' mind: Instructional design for complex learning. *Educational Psychologist, 38*, 5–13.

Yu, F. Y. & Liu, Y. H. (2005). Potential values of incorporating multiple-choice question-construction for physics experimentation instruction. *International Journal of Science Education, 27*, 1319-1335.

GeMA-ICT Learning Effectiveness in Improving Student Mathematical Ability

Sigid Edy PURWANTO[a*], Wahidin[b] & Aidiyah Novian NISYAH[c]
[a]*Mathematic Education, UHAMKA University, Indonesia*
[b]*Mathematic Education, UHAMKA University, Indonesia*
[c]*Mathematic Education, UHAMKA University, Indonesia*
*sigidmath@yahoo.co.id

Abstract: This study examines and measures the effectiveness of using GeMA-ICT learning methods (Games, Manipulatives, and Activities with Information and Communication Technologies) on the students' mathematical abilities. GeMA-ICT is a new method that resulted from the theoretical studies of the Mathematics Education Study Program UHAMKA research team. Mathematics has abstract objects such as abstract facts, abstract concepts, abstract operations, and abstract principles. The abstract objects in mathematics education sought to be easily understood by students, by presenting them through concrete objects, props math, math games and math activities supported by the use of ICTs. Props, games, and activities developed with the help of ICTs are a blend of media and learning methods that can visualize math concepts. A set of concrete objects are intentionally drafted, designed, manufactured, assembled, and used for instilling or developing concepts or principles in mathematics. With GeMA-ICT methods, things that can be presented in the form of abstract models such as concrete objects can be viewed, held and manipulated so that they can be easily understood. In addition to presenting math games, learning math is made fun for students and learning activities are turned into a process of investigation and exploration of further mathematical concepts.

Keywords: GeMA-ICT, props math, math games, math activities

1. Introduction

Extant research has shown that effective learning seldom happens in a scenario where the teacher dominates the process and reduces the students to passive participants. Whereas learning ought to be made easy for students, lots of teachers are still making it difficult for them because they do not reinvent their classroom procedures according to the patterns of modern teaching approaches, Classroom activities such as, "one-way lectures" and "take-home-assignments" only teach students to perform certain symbolic procedures as well as to work but not to think. This is what characterizes most mathematics classrooms – students are often confined to a scenario where they simply watch the teacher solve mathematical problems on board and at the end of the class they would be given assignments to take home and solve them in worksheets according to the pattern that the teacher used, (Turmudi, 2008).

Another characteristic of the traditional teaching approach is that teachers emphasize on curriculum attainment targets rather than targeting at mastery of the material. Many math teachers teach by simply following routines that are critically ineffective. This eventually wears them out because the process is not only tedious to the teacher but also damaging to students' interest (Sobel and Maletsky, 2004). Furthermore, in the traditional classroom setting, teachers tend to use chalk and talk method often more than not. This was caused by several possibilities: 1) schools already have props but are not usin them optimally; 2) school do not have props; 3) school have had adequate props but are not good enough (Asyhadi, 2005).

To reinvent the current classroom scenario that still retains the traditional characteristics, and enhance the quality of national education; research must focus on teachers and educational facilities and infrastructure. Learning in the traditional classroom setting does not give room for maximal innovation and creativity and it hinders the opportunity of using a wide variety of methods and media. As a result, learners gain minimal knowledge little understanding of the content of the lessons. Hence, the learning process becomes unattractive and weakening to the national education standards. Adding up all these together, the standard of the competence of graduates becomes significantly undermined. Hence, there is need to test pilot the usage of contextual and humanistic learning approaches in today's mathematics classrooms. Accordingly, the attention of governments and experts of mathematics education in many countries is quite crucial at this point in time, so as to improve the mathematical ability of students and equally champion gigantic efforts for overcoming students' passiveness in mathematics classrooms.

Results in NAEP study show that students are still having difficulties when faced with problems that require reasoning and problem-solving abilities. Mathematics learning of the present day has not been able to develop students' mathematical problem solving abilities optimally. This causes the learning conditions of students to be only able to resolve problems only in accordance with the examples given by the teacher. But when students are given anther problem that is similar but not the same as the examples once given, they find it difficult or even impossible to handle. It was observed in NAEP study that the success rate of students in solving problems dropped dramatically when the context is replaced with things that are not known to the students (Suherman et al, 2003).

However, one of the new offers that can activate students' learning in today's Mathematics classrooms is learning with GeMA-ICT (Games, Manipulatives, and Activities with Information and Communications Technologies). This lesson allows students to be actively involved through hands-on activity. GeMA-ICT is an emphasis on the learning activity of students in the manipulation of concrete objects by exploiting the use of math games, math props, and activities with the help of ICTs. GeMA-ICT encourages students to learn facts, skills, concepts and theories through manipulative activities with concrete objects, models, props, or math games. GeMA-ICT can increase the desire to learn; enhance learning by doing; and enhance the application of scientific problem solving through making analyses and evaluations. GeMA-ICT insists on ensuring that mathematical principles are found, generalized, and proven. This learning approach takes off the abstract nature of mathematics, making it interesting through the integration of games and a variety of other ICT activities.

In this approach, students' thinking phases are made concrete and real, so as to allow for the execution of visual-kinesthetic activities a vriety of ICT tools. On the whole, GeMA-ICT is a good option for overcoming the dull routine of learning in Maths classrooms. It is outrightly appropriate learning for enhancing students' mathematical abilities. Therefore, the purpose of this study is to determine the differences that are likely to be found in mathematics learning outcomes of students using GeMA-ICT method and students using expository method.

2. Literature

2.1 GeMA-ICT

GeMA-ICT is an abbreviation for Games, Manipulatives, and Activities with Information and Communications Technologies. Author deliberately adopted GeMA-ICT in order to obtain the equivalent word that can represent the learning activities of students who will be using the games, props, and math activities with ICTs.

When referring to the percentage of the amount that can be remembered, GeMA is a very important lesson. Johnson and Rising in Ruseffendi (2006) posited that, "learners can remember about a fifth of the heard, half of which is seen, and three-quarters of the done". But in learning Mathematics, the manipulation of concrete objects is very important; yet in most cases the concrete objects used are ordinary local materials.

For learning to be more meaningful teachers should as much as possible avoid dominating the process. Literature has shown that in a teacher dominated classroom scenario, learning tend to be target-oriented-mastery in nature, and such learning has only proved successful in short term given competitions, while in the long run it doesn't provide the child with the desired problem solving ability.

Abstract mathematical objects have facts, concepts, principles and principles of operation, which are as well abstract. It seems that students do easily come to grasp with lessons that are backed up with an orientation of concrete phenomena more than those that are not.

The works of Piaget, Bruner and Dienes have supported the claim that, manipulation of concrete objects is an important activity in mathematics learning. In GeMA, students solve problems, explore mathematical concepts, formulate and experiment with mathematical principles, and make mathematical discoveries through manipulation of concrete objects that represent abstract mathematical ideas.

GeMA learning follows the principle of learning by observing and doing; and it starts from the concrete to the abstract, as is the case with inductive method. Hence, students learn the objects and then generalize while ignoring the special nature of mathematical abstraction. It can attract learners to abstract mathematics. According to Ernest in Turmudi (2008) that learning mathematics is first and foremost is active, with students learning through games, activities, investigations, projects, discussion, exploration, and discovery.

Students can learn facts, skills, concepts, postulates, or theories through manipulating concrete objects, models, props, or math games. Hence, GeMA can increase students' desire to learn, since it is based on the principle of learning by doing; and it also gives room for the students to appreciate and apply the scientific method of problem solving. However, application of learning aids such as laboratory equipment and other media becomes necessary for the implementation of learning with GeMA. But while the design of learning outside the classroom is directed to how students can play while learning, GeMA can be operationalized through the implementation of games, props, and math activities.

2.2 Math Games

Basically the students would love the games and puzzles, because play is indeed a world of children (Turmudi, 2008). Imam Al-Ghazali said, "Playing around for a child is something that is very important but banning him from playing around will turn him off and disturb his sense of belonging, intelligence and general rhythm of life". Children will find it easier to learn arithmetic by means of handing out apples to them than by abstract examples . Congruently, play is seen as a natural activity in helping the child to gain experiences, develop creativity and determination to succeed.

If a mathematical concept is presented through play, the understanding of the concept is expected to be steady, because learning in this way is in a natural pattern, which is in accordance with the child's instincts. Hence, the learning process is a psychological process, not a logical process. Therefore, the patterns should not be mere series of knowledge that have been previously defined in form of a mechanical process; but rather, through play, the students construct their mathematical patterns (Hudoyo, 1985).

Math game is a fun activity that can support the achievement of learning goals in mathematics cognitive, affective, and psychomotor domains. Math games help students to memorize basic facts, find the arithmetic operations and improve numeracy skills, as well as gain more understanding and problem solving ability (Ruseffendi, 2006).

Learning mathematics by games and puzzles was also emphasized by Turmudi (2008), as an approach for motivating and giving fun to both students and teachers alike in the learning process. This is important because games and puzzles have been widely recognized as a way of inspiring students to mathematical literacy. Ernest in Turmudi (2008) claims that games teach math effectively due to: 1) Provision of reinforcement and skills practice, 2) Provision of motivation, 3) Assistance, acquisition and development of mathematical concepts, and 4) Development of problem-solving strategies. Posamentier and Stepelman in Turmudi (2008) presented an analogy between game strategy problems solving strategy in the following table:

Table 1: The Comparison of Games Strategy and Problem Solving

Games Strategy	*No*	*Problem Solving Strategy*
Read the rules	1	Read the rules
Understand the rules	2	What is given and what to look for?
Develop a plan	3	Write the equation
Work the plan	4	Solve the equation
If you win, smile; if not, think about why it lost	5	Check the answer

Moerlands and Makkink (2003), reported that the play activity could help resolve the problem of the "unknown" number. Their study revealed improvement in child's learning outcome by 40%, from a mean baseline (2.6) to 3.7. Congruently, results in this study indicate that the learning using games makes abstract problems seem ordinary (Armanto, 2003). Meanwhile Benko & Maher (2006) have reported significant improvemnts in students' oral, written ability and physical representation ability to level 7 after learning through games that use dices.

2.3 Props Mathematics

Props are teaching media that contains or carries the characteristics of the concept being studied. A set of concrete objects are designed, manufactured, assembled, or prepared deliberately to help embed or develop concepts or principles in mathematics. With props, things that can be presented in the form of abstract models such as concrete objects that can be seen, held and distorted are used for teaching, so that lessons taught can be easily understood.

Since the 1950s until the 1970s, research into the use of props in teaching mathematics has been going on, and not less than 20 summaries of such researches have been recorded. Among these is the popular summary of Higgins and Suydam in 1976 (Lithanta, 2009), which among other things concluded as follows: 1) In general, the use of visual aids in the teaching of mathematics was successful or effective in promoting student achievement. 2) Approximately 60% vs 10% of the students sampled showed a convincing success of learning with the props than without. 90% of students that learnt with props recorded yet the same magnitude of 90% in their learning outcomes above those students that did not use props. 3) Manipulation of the visual aids is important for elementary students at all levels. 4) Only a little evidence was found showing that manipulation of props is only manageable at lower learning levels.

Slamet, Soenarto, and Wahidin (2008), reported that the ability to compute and factorize quadratic equations and increase students' learning outcomes becomes easy when lessons are administered with props, AEM (Block al-Khawarizmi). Congruently, studies have shown that learning with games could serve the needs of students at all levels; and even weak students could easily manipulate concrete rectangular objets using their prior knowledge of the broad concept of the rectangle, (Dienes AEM). Hence, a problem is often best solved (understood) by using sketches, folded pieces of papers, pieces of strings, or other simple props available. Invariably, strategic use of props can make the situation real to the students so as to motivate them to learn faster and better. Therefore, manipulation of geometry models can be a way of help the problem solving process as well as an activity for innovation, (Sobel & Maletsky, 2004).

Mathematics learning activities wherever possible involve all the senses of the students, especially hearing, seeing, and touching. In this case props bridge gap between the abstraction process and apprehension. In addition, by using the props, the child can be helped to find a strategy to solve problems. This is done by allowing the child to describe the problem in a simple concrete pattern; construct his or her own knowledge and understanding of the issues for the purpose of develop problem-solving strategies (Triyana, 2004).

2.4 Activities

In learning with games, activites are equated with experiments in a way that present lessons for students to conduct experiments and prove their own experiences and the things they learnt. In this case, students are given the opportunity to experience for themselves, learn on their own, following a process that allows them to observe objects, analyze issues and draw conclusions, (Djamarah and Zain, 2006). Through this practicum process, the students are able to discover facts and truths in the form of conjectures and theorems by themselves.

Learning mathematics through practical applications or hands-on experience is an activity within the framework of the invention and principles of mathematical concepts. The process is helpful in improving students' ability to explore, investigate, and the draw conclusions through physical activity, as well as through mental and emotional engagments, (Krismanto, 2003). For the full geometry of the material abstraction, hands-on mathematics is still a necessary experience for improving students' learnig outcomes.

With the mock objects (models) or concrete objects that are deliberately prepared to further stimulate the minds of students in constructing their own understanding, there are more elements of practical work on learning experiences for students to use the knowledge they gained (according constructivism) and not to solely depend on how their teachers teach math. This is an advocation for a paradigm shift in teaching mathematics.

From the analysis of data from the 1996 NAEP test, two samples of countries involving 15,000 students mentioned that the rate of 8 students whose teachers actively taught through the process of learning activities generated employments in mathematics achievement levels of more than 70% (Crawford, 2001).

Vui (2006-2007) reported that the goal of good practice in teaching mathematics is to help students make meaning of the contents and skills taught in the lessons (what is known) and the process involveds (what is done). Good Practicum must balance between the content and the process of lerning problem solving skills, because the two are entirely different aspects of the knowledges which the students must be equipped with, all their lives. When teachers use manipulative materials in teaching mathematics, they discover that their students are more active in learning. Students enjoy learning mathematics with dynamic models or motions. Teachers must learn how to create new mathematical models of problematic situations and prepare good manipulative materials for students to deal with. Students may also be inspired to build their own questions and activities.

2.5 Using ICTs

Using ICTs in this study refer to the use of Microsoft power point and GeoGebra Software. The use of Microsoft power point should be simplified in a way that will make it a preferred attraction to many. Currently, Microsoft PowerPoint is widely used by teachers in delivering course materials. Most teachers consider the use of Microsoft power point as an effective and efficient tool for practical learning.

In the use of GeoGebra many interesting things can be encountered and difficult mathematical problems such as in calculus lessons are easily resolved with the GeoGebra media software. Hence, apart from calculus problems, other mathematical problems, such as solving line equations, vectors, angles, algebras, geometry and many others can be easily resolved by the use of the GeoGebra media software. In addition, GeoGebra is designed to facilitate its use in an interactive way.

3. Research Method

3.1 Research Instruments

This research was conducted among the 150 Junior High Schools at Kramat Jati, East Jakarta on January 20 until February 4, 2014. This study uses a quasi-experimental design. In a quasi-experimental, the subjects are not taken at random but researchers do accept existing subjects. The use of quasi-experimental research is based on the consideration that there are classes that were not considered previously formed at random grouping of individuals who would disturb the Teaching and Learning Activities at schools.

The population in this study were students of 150 Junior High Schools, that comprised up of 8 eighth grade classes. This research used two classes as a sample, each of which has the same characteristics. The first class was used as an experiment class, and it was taught by using the GeMA-ICT. The other class was used as a control class, and it was taught without using the GeMA-ICT.

Data was drawn from the results of learning mathematics scores amanog 71 students sampled. The research instrument used to measure students' mathematics learning outcomes was desinged with multiple choices questions having four (4) alternative answers. The validity of the instrument was measured using the point biserial correlation formula, by which the instrument earned as much as 20 valid questions. Reliability of the instrument was measured by using the Kuder - Richardson (KR-20) formula, by which the instrument a very high reliability degree of $R_{11} = 0.836 > r_{table} = 0.329$.

3.2 Data Analysis

Students mathematical ability tests were analyzed by using the inferential normality test, homogeneity test and t-test, to see the effectiveness of a given learning GeMA-ICT method.

4. Result

4.1 Student Mathematical Ability

The descriptive statistics research data obtained from the study of mathematics students in the experimental class (i.e. the class taught with GeMA-ICT method) is presented in Table 2 below:

Table 2: Descriptive Statistics of Students Learning Outcomes in Math Experiment Class

Data	Maximum Score	Student Learning Outcomes Math Scores						
		Y_{min}	Y_{max}	$\bar{Y}$	s	s^2	Me	Mod
35	20	9	18	14,74	1,99	3,96	15	15

The descriptive statistics research data obtained from the study of mathematics students in the control class (the class that was taught without using GeMA-ICT method) is presented in Table 3 below:

Table 3: Descriptive Statistics of Student Learning Outcomes Math Control Class

Data	Maximum Score	Student Learning Outcomes Math Scores						
		Y_{min}	Y_{max}	$\bar{Y}$	s	s^2	Me	Mod
36	20	7	17	13,08	2,55	6,48	13	13

4.2 Results of Testing Data Analysis Requirements

Using the Lilliefors test, it was concluded that the experimental and the control class data derived from the sample had a normal distribution.

Using the Fisher test, it was concluded that the samples of the two classes of experimental class taught using GeMA-ICT method and control class taught without using GeMA-ICT method have the same conditions or homogeneous variance.

4.3 Results of Hypothesis Testing

Through the calculation of the average grade, experimental and control classes obtained t = 3.06 and t (0.95; (69)) = 1.67 with a significance level $\alpha = 0.05$ and (df) = 69. Since t = 3.06 > 1.67 = t (0.95; 69). This means the rejection of the research hypothesis H_0 which stated that there are significant differences in mathematics the learning outcomes of students taught with GeMA-ICT method and those taught without using GeMA-ICT method in the 150 Jakarta Junior High Schools.

5. Conclusion

GeMA-ICT methods improve students' mathematics learning outcomes owing to students' usage of manipulative activities, props and games in this study. The process also allows students to discuss with each other in the group. On the whole, the GeMA-ICT method improves students' learning outcomes better than other methods of teaching mathematics.

Acknowledgements

We would like to thank University of Muhammadiyah Prof. Dr. HAMKA of Jakarta for supporting this study.

References

Armanto, D. (2003). Katak Pemakan Kapur. Penelitian PMRI oleh Frans Moerlands dan Annie Makkink, SD kelas 2 di Bandung, 17 Februari 2003.

Asyhadi, A. (2005). Pengenalan Laboratorium Matematika di Sekolah. IHT Media Bagi Staf LPMP Pengelola Laboratorium Matematika Tanggal 5 s.d. 11 September 2005 di PPPG Matematika Yogyakarta.

Benko, P. and Maher, CA. (2006). Students Constructing Representations for Outcomes of Experiments. Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education, Vol. 2, pp. 137-143. Prague: PME

Crawford, M.L. (2001). Teaching Contextually. Research, Rationale, and Techniques for Improving Student Motivation and Achievement in Mathematics and Science. Texas: CCI Publishing. Inc. [online]. http://www.cord.org/contextual-teaching. [6 Oktober 2005]

Dahlan, J.A. (2004). Meningkatkan Kemampuan Penalaran dan Pemahaman Matematik Siswa Sekolah Menengah Lanjutan Pertama melalui Pendekatan Pembelajaran Open-Ended. Disertasi PPS UPI Bandung: Tidak diterbitkan.

Djamarah, SB. dan Zain, A. (2006). Strategi Belajar Mengajar. Jakarta: Rineka Cipta.

Hudoyo, H. (1985). Teori Belajar dalam PBM Matematika. Jakarta: Depdikbud.

Idris, N. (2006). Creativity in the Teaching and Learning of Mathematics: Issues and Prospects. Department of Mathematics and Science Education Faculty of Education University of Malaya.

Ismail, A. (2006). Education Games: menjadi cerdas dan ceria dengan permainan edukatif. Yogyakarta: Pilar Media.

Krismanto. (2003). Beberapa Teknik, Model dan Strategi dalam Pembelajaran Matematika. PPPG Matematika. Yogyakarta.

Lithanta, A. (2009). Alat Peraga Perkalian Model Matrik sebagai Media Pembelajaran Matematika yang Menyenangkan. SDN Mangunharjo V.

Maier, H. (1985). Kompendium Didaktik Matematika. Bandung: CV Remaja Karya.

Moerland, F. & Makking, A. (2003). The chalk frog: How playful activities can help solving 'point sums'. Retrieved from http://www.edumat.nl/English-content/PMRI-Chalkfrog.pdf

Mullis, I.V.S., Martin, MO. and Foy, P. (2005). IEA's TIMSS 2003 International Report on Achievement in the Mathematics Cognitive Domains. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Pujiati. (2004). Penggunaan Alat Peraga dalam Pembelajaran Matematika SMP. Disajikan pada Diklat Intruktur Matematika SMP Jenjang Dasar, 10–23 Oktober 2004, Dirjen Dikdasmen, PPPG Matematika Jogjakarta.

Puskur. (2007). Kajian Kebijakan Kurikulum Mata Pelajaran Matematika. Balitbang Depdiknas.

Riedesel, CA., Schwartz, JE. and Clements, DH. (1996). Teaching Elementary School Mathematics. Boston: Allyn & Bacon.

Ruseffendi, ET. (2003). Dasar-dasar Penelitian Pendidikan dan Bidang Non Eksakta Lainnya. Bandung: Penerbit Tarsito.

____________. (2006). Pengantar kepada Membantu Guru Mengembangkan Kompetensinya dalam Pengajaran Matematika untuk Meningkatkan CBSA. Bandung: Tarsito.

Shadiq, F. (2004). Penalaran, Pemecahan Masalah, dan Komunikasi Dalam Pembelajaran Matematika. Disajikan pada Diklat Instruktur Matematika SMP Jenjang Dasar, 10–23 Oktober 2004. Dirjen Dikdasmen PPPG Matematika Jogjakarta.

Slamet, Soenarto, Y. dan Wahidin. (2008). Upaya Meningkatkan Hasil Belajar Matematika Siswa SMA 97 Jakarta melalui Pembelajaran dengan Metode Laboratorium. Laporan Penelitian untuk Peningkatan Kualitas Pembelajaran (PPKP) Dikmenti DKI Jakarta tahun 2008: Tidak Diterbitkan.

Sobel, MA. Dan Maletsky, ME., terj. Suyono. (2004). Mengajar Matematika. Ed. 3. Jakarta: Erlangga.

Suherman, E. dkk. (2003). Strategi Pembelajaran Matematika Kontemporer. Bandung: Depdiknas-JICA-UPI.

Suherman, E. 2004. Model-Model Pembelajaran Matematika Berorientasi Kompetensi Siswa. Makalah disajikan dalam acara Diklat Pembelajaran bagi Guru-guru Pengurus MGMP Matematika di LPMP Jawa Barat tanggal 10 Desember 2004: Tidak Diterbitkan.

Sulastri, S. dan Dahlan, JA. (2007). Penerapan Model Pembelajaran Quantum Teaching untuk Mengurangi Kecemasan Matematika Siswa pada Topik Bangun Datar. Bandung: Prosiding Seminar Nasional, 8 Desember 2007 Jurusan Pendidikan Matematika FPMIPA UPI.

Sumarmo, U. (1987). Kemampuan Pemahaman dan Penalaran Matematika Siswa SMA dikaitkan dengan Kemampuan Penalaran Logik Siswa dan Beberapa Unsur Proses Belajar Mengajar. Disertasi IKIP Bandung: Tidak Diterbitkan.

__________. (2004). Pembelajaran Matematika untuk Mendukung Pelaksanaan Kurikulum Berbasis Kompetensi. Makalah pada Pertemuan MGMP Matematika SMPN I Tasikmalaya. [12 Februari 2005].

Suryadi, D. (2005). Penggunaan Pendekatan Pembelajaran Tidak Langsung serta Pendekatan Gabungan Langsung dan Tidak Langsung dalam Rangka Meningkatkan Kemampuan Berpikir Matematik Tingkat Tinggi Siswa SLTP. Disertasi PPs UPI: Tidak diterbitkan.

Triyana, J. (2004). Peran Alat Peraga dalam PMRI. Buletin PMRI Edisi kelima Oktober 2004.

Turmudi. (2008). Landasan Filsafat dan Teori Pembelajaran Matematika (berparadigma Eksploratif dan Investigasi). Jakarta: Leuser Cita Pustaka.

Vui, T. (2007). Helping Students Develop and Extend Heir Capacity to Do Purposeful Mathematical Works. Department of Mathematics, Hue College of Education, Vietnam.

Wittmann, E. Ch. (2004). Developing Mathematics Education in a Systemic Process. In Proceedings of the Ninth International Congress on Mathematical Education. 75-90.

Yunengsih, Y., Widiatmika, IMAA. dan Candrasari, A. (2008). Ujian Nasional: Dapatkah Menjadi Tolok Ukur Standar Nasional Pendidikan? (hasil kajian UN matematika SMP). Jakarta: Putra Sampoerna Foundation.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

THE EFFECTIVITY OF ASSOCIATON PICTURE MEDIA APPLICATION TOWARD THE KATAKANA LETTER READING COMPREHENSION OF GRADE TEN STUDENTS OF SMK MANAJEMEN (MANAGEMENT VOCATIONAL SCHOOL) JAKARTA

Restoe NINGROEM[a*] & Endy Sjaiful ALIM[b]**
[a] *State University of Jakarta, Indonesia*
[b]*University of Muhammadiyah Prof. DR. HAMKA (UHAMKA), Jakarta, Indonesia*
*restoe.ningroem@yahoo.com, **esaitb@yahoo.com

Abstract: The purpose of this research is to acknowledge the effectivity of association picture media application toward the students' katakana letter reading comprehension. The hypothesis proposed that the comprehension of students' katana letter reading, which are exposed to association picture in class is more effective compared to the students which aren't exposed with such media. The research was conducted over grade ten students of SMK Manajemen Jakarta. The research used quasi experiment. The sample was taken from 50 students, which categorized 25 students as experimental class and 25 students as controlled class. The data collection technique used an instrument of katakana alphabet reading comprehension test. The requirement test of data analysis used in this research was Lilliefors Test for normality and Fisher-test for homogeneity, it was revealed that the class was normal and homogeneous. Hypothesis assessment used t-test which showed tcount = 7.67 ≥ 1.67 = ttable that Ho was rejected It is revealed from the research that association picture media is effective to increase student's comprehension in reading katakana letter.

Keywords: Learning Media, Association Picture, Reading Katakana.

1. INTRODUCTION

Indonesia has had bilateral cooperation with other countries for a long time, In engaging cooperation with other country, it is necessary to comprehend the language, because it is the sound symbol system that articulated (produced by speech device) which is used by community members to cooperate, interact, and self recognition (TPKB, 2005)

One of the important evidences of language in engaging cooperation with Japan has been shown by the presence of Japanese curriculum in middle schools in Indonesia. Acquiring skill in Japanese is a basic step and access for our people to study, get knowledge, understand their culture, and have employment in Japan, a modern and developed country. A country that has rich natural resources, various cultures, and positive life values that deserve to learn from.

In order to have skill in Japanese, it is necessary to read researches of Japanese literature first hand. One of the components of the Japanese writing system is Katakana letter.

In reality, students find it difficult to memorize Katakana compare to Hiragana, because Hiragana letter appears more in the text book. Students' obstacle in learning Katakana is revealed to be psychologist one, such as: interest, attitude, confidence, and intelligence.

One of the efforts to overcome the issue is to use media in teaching Katakana letter. According to Hamalik, using media in the learning process can stimulate new will and interest, arouse motivation and eagerness in learning, and even affects psychologically to students. It can also ease students to comprehend, presents data interestingly and reliably, interprets data and information easily (Rohani, 1997).

One of teaching media that eases students in learning Katakana is association pictures, the benefit of such media can deliver messages, idea, etc. That involves little verbal languages but, leaves a deeper impression (Arsyad, 2003). With association picture media, learners will have long term memorization of *katakana* letters because association makes words easy to be stored in the memory and to be retrieved if needed since they already accustom to the words (Sameto, 2003).

2. LITERATURE REVIEW

2.1.Learning Media

Association of Education and Communication Technology (AECT) in USA identifies media as all forms and channels which are used to deliver messages or information (Arshad, 2005). Gerlach and Ely (1971) give the definition of media broadly and narrowly. Broadly, it means every individual, material, or event that gives the student opportunity to gain knowledge, skill, and manner. Therefore the media doesn't simply mean things or objects, it can also mean an individual and a learning event. Teacher, text book, school environment can be the media. While narrowly, media is non-personal media that teacher use as people in charge of the learning process to achieve goals. That way media tend to be looked as graphical tools, photograph, or electronic tools to grasp, reshapes visual or verbal information (Roshidi, 2009).

According to Hamalik (1986) educational media is something that can be digested by the senses, shape and things which are visual and audio, used as a means of communication in the learning process, as an aid in learning process and related to teaching method. Picture is a crucial visual aid and available everywhere. Crucial because it can replace verbal words, concretes the abstract. Pictures enable people to grasp the idea or information contained clearly, vividly more than words can say (Munadi, 2007). As the Chinese saying, pictures speak more than a hundred words (Adiman, 2009).

2.2 Association

The association is relating one event to another event, between someone and other people, which is considered as a related series and interconnected to each other.

In learning a foreign language, we must develop association as a mean that allows us to memorize words easily and brings full picture and any circumstances that embedded in that word (Sameto, 2003). The ability to associate in learning foreign language needs to be developed according to particular events.

Association is available to be developed for all life aspects and circumstance that related to our life (Sameto, 2003). Memorization in conventional ways is by applying left brain hemisphere by memorizing and repeating materials, while the association is done by applying right brain hemisphere to memorize symbols or pictures which is easier and quicker to absorb. From descriptions and benefits of association elaborated above. It can be concluded that the association picture media discussed in this research are pictures that serves as a noun or verb that resembles a Katakana letter as a series of meaning that interconnected to each other to smoothen long term memorization of Katakana letter shapes.

2.3.Reading Comprehension

Anderson (Syamri, 2011) defined reading as uttering written language symbols. While A.S.Broto and Syamri (2011) stated that reading is the uttering of sound symbols.
Aspects of reading:
1. Mechanical Skills, which is considered in the lower order. This aspect covers:
a. Identifying letter form
b. Identifying linguistic parts (phoneme, word, phrase, clause pattern, sentence, etc)
c. Identifying relation or correspondence of spelling and sound pattern
d. Low level of reading speed.
2. Comprehension skills, which is considered in the higher order. This aspect covers:
a. Comprehend meaning definition.
b. Comprehend meaning.
c. Evaluation or assessment (content, form)
d. Flexible reading speed, that adjustable to particular situation.

From both important reading aspects in reading, it can be inferred that reading letter is in the early introduction of mechanical skills, which takes place before one do comprehension skills.

2.4 Katakana Letters

Katakana letters are formed from straight and sketchy lines or *chokusenteki* (直線的). These straight and stroke lines that differ them from *hiragana* (Ang, 2005)

ア a	イ i	ウ u	エ e	オ o
カ ka	キ ki	ク ku	ケ ke	コ ko
サ sa	シ shi	ス su	セ se	ソ so
タ ta	チ chi	ツ tsu	テ te	ト to
ナ na	ニ ni	ヌ nu	ネ ne	ノ no
ハ ha	ヒ hi	フ fu	ヘ he	ホ ho
マ ma	ミ mi	ム mu	メ me	モ mo
ヤ ya		ユ yu		ヨ yo
ラ ra	リ ri	ル ru	レ re	ロ ro
ワ wa				ヲ wo
				ン n

Figure 1. Katakana Letter

The word "kata" means partial, incomplete, or separated. The word "katakana" means "kana separation", as katakana script which originated from more complicated kenji component (Japan: an Illustrated Encyclopedia, 1993).

In koujien dictionary, the function of katakana (Izuru, 1998) is:

現在では主に外来語や擬音語などの表記に用いる。

"Genzai de wa omo ni gairaigo ya giongo nado no hyouki ni mochiiru " , Which means in the contemporary time, generally *katakana* letter is used for foreign language writing and *onomatope*.

Although *katakana* and *hiragana* letters are in the group of *kana* letter, but the function is different. *Katakana* letter is used in the writing of Japanese words which comes from absorption of foreign languages, foreign countries, animals, foreigner's name, plants, and foreign cities (Ang, 2005).

Katakana also often (not always) used for transcription of Japanese company, e.g: Suzuki becomes スズキ, and Toyota becomes トヨタ. *Katakana* also serves as reinforcement, specially on signs, advertising and billboard. For example, seems common to see *koko* ココ (here), gomi ゴミ (waste), or megane メガネ (glasses). Reinforcement using *katakana* in sentences is sometimes done by writers..

Katakana sometimes serves as *hiragana* replacer or as *furigana* to give utterance a word that is written in Latin letter, or for foreign word, which is written in *Kanji* for meaning, kanji but meant to be read as the origin form.

Katakana sometimes also used to indicate words that speech in foreign accent or unusual, in foreign character, robot, etc. For example, in *manga*, foreign character speech or robot can be represented by *konnichiwa* コンニチワ (hello) not *hiragana* which is more typical こんにちは.

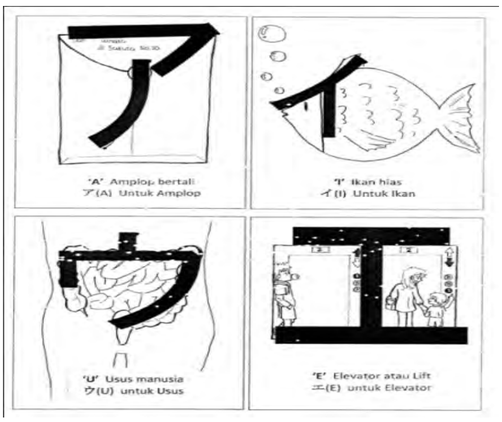


Figure 2. Association Picture Media sample

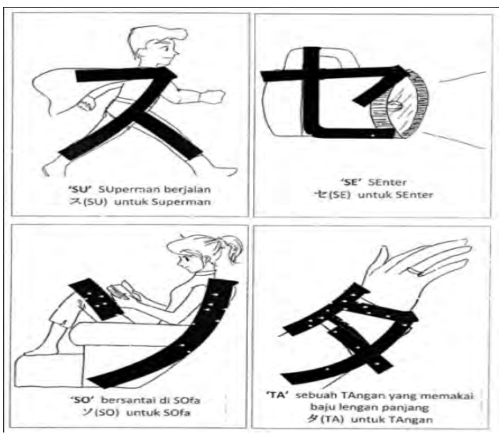


Figure 3. Association Picture Media sample

2.5 Critical Framework

Since it is important to improve *katakana* reading comprehension of grade ten students of SMK, and the benefit of using association picture media is to ease students in memorizing the shape of *katakana* letter in a long term, then it is assumed that association picture media is suitable to be implemented in the learning process of *katakana* letter. So, it can be inferred that association picture media is effective to improve *katakana* letter reading comprehension.

3. METHOD AND SAMPLING

This research applied quantitative approach using Quasi Method. Sampling was treated by the technique of cluster sampling. Two classes of grade ten students of SMK Manajemen Jakarta are determined as experimental and controlled classes. Instruments used for the research consisted of a set of test and questionnaire as additional data. There were three kind of test forms; multiple choice, translation of katakana into latin, and vocabularies matching.

4. FINDING AND DISCUSSION

The research shows the end result of the comprehension of students in reading *katana* letter as follow:

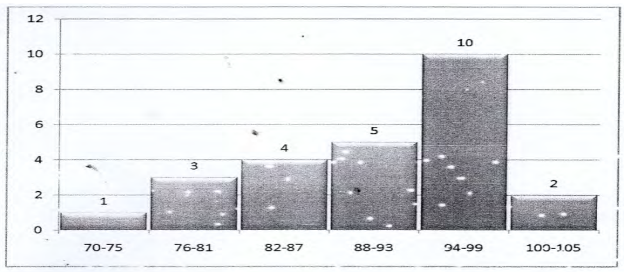


Figure 4 : Experimental Class Score

Figure 4 shows that the score of the comprehension of reading *katana* letter of the Experimental Class , from 10 students the scores are in the span of 94-99 and the highest score is 100, and the average score is 90,5.

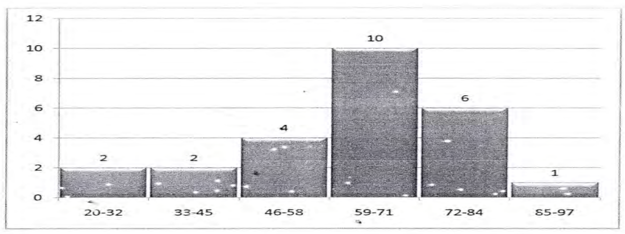


Figure 5: Controlled Class Score

Figure 5 shows that the score of the comprehension of reading *katana* letter of the Controlled Class , from 10 students the scores are in the span of 59-71 and the highest score is 97, and the average score is 61,2.

Looking at the test result the comprehension of reading *katana* letter of experimental class and controlled class, it is revealed that the average score of experiment class is higher than the score of control class. Looking at the data analysis, it is noted that nul hypothesis (Ho) succesfully rejected and research hypothesis (HI) was successfully accepted. This finding was revealed by t-test formulation, tcount was found at 7.64 and t table which came from degree of freedom at 48 with significant level at 0.05 was 1.67. After research, t count was found to be at 7.638. The requirement of H0 rejection and HI acceptance was $t_h > t_t$. Since t count was higher than t table (7.638 > 1.67), so Ho is rejected and HI was accepted which means “the application of association picture media to increase reading comprehension of grade ten sudents of SMK Jakarta Manajemen is more effective compare to the opposite approach”.

After a comparative test conducted using t-test, it showed that tcount = 7.67 ≥ 1.67 = ttable with the level of significant 0,05, it was assumed that the comprehension of reading *katana* letter from students who were exposed with association picture media was better than those who weren’t exposed by the media.

5. CONCLUSION

Based on the research of data analysis evaluation and all data that the writer received for this research, a conclusion reached, that applying educational media in learning process of Japanese will help to reach learning goals, since one of the benefits of using educational media is to make learning easy.

In this research association picture media is concluded to be effective to improve students’ comprehension of reading *katakana* letter. The special feature of association picture media that presents pictures that resembles and inter-relates *katakana* letters eased student to memorize *katakana* letter. Besides that, using association picture media, learning process becomes less boring since the students were exposed with attractive pictures and little verbal elaboration (in written or speech only).

References

All, Muhammmad. (1987). Proses Kependidikan Prosedur Dan Strategi, Bandung: Angkasa.

Ang, Haryono. (2005). Mengenal Aksara Jepang (Hiragana, Katakana, Kanji). (Cetakan pertama). Jakarta: Puspa Swara.

Arsyad, Azhar. (2003). Media pembelajaran. Jakarta: PT Raja Grafindo Persada.

Arsyad, Azhar. (2005). Media Pembelajaran, Jakarta: PT Raja Grafindo Persada.

Danasasmita, Wawan. 2009. Metodologi Pembelajaran Bahasa Jepang (Cetakan Pertama). Bandung: Rizqi Press.

Emzir. (2008). Metode Penelitian Pendidikan: Kuantitatif dan Kualitatif. Jakarta: PT Raja Grafindo Persada.

Frederic, Louis (Translated by Ka The Roth). (2002). Japan Encyclopedia. England: The belkmap Press of Harvard University Press.

Hamalik, Oemar. (1986). Media Pendidikan. Bandung: PT Alumni.

Izuru, Shimura. (1998). /Si^lE ; J^jEJiS. Japan: Iwanami Shoten.

Pusat Pembinaan Bahasa Departemen Pendidikan dan Budaya. (1995). *Kamus Besar Bahasa Indonesia edisi kedua.* Jakarta: Balai Pustaka.

Rohani, Ahmad. (1997).Media Instruksional Edukatif. (Cetakan pertama).Jakarta: PT RINEKA CIPTA.

Roshidi, Abdul Wahab. (2009). Media Pembelajaran Bahasa Arab. Malang: UIN-Malang Press.

Sudjianto dan Ahmad Dahidi. (2007). Pengantar Linguistik Bahasa Jepang. (Cetakan kedua). Bekasi: Kesaint Blanc.

Sadiman, Arief S, dkk. (1990). Media Pendidikan .*Pengertian, Pengembangan dan Pemanfaatannya.* Jakarta: CV Rajawali.

Sadiman, Arif S, dkk. (2009). Media Pendidikan. *Pengertian, Pengembangan dan Pemanfaatannya.* Jakarta: Rajawali Pers.

Sameto, Hudoro. (2003). Kiat Menguasai Bahasa Asing. (Cetakan pertama). Jakarta: Puspa Swara.

Sukardi. (2003). Metcdologi Penelitian Pendidikan. *Kompetensi danPraktiknya.* Jakarta: PT Bumi Aksara

Tim Penyusun Kamus Pusat Bahasa. (2005). Kamus Besar Bahasa Indonesia. Edisi ketiga (Cetakan ketiga). Jakarta: Balai Pustaka

Usman, M. Basyiruddin dkk. (2002). Media Pembelajaran. (Cetakan pertama). Jakarta: Ciputat Pers.

Wiryawan, Sri Anitah, dkk. (1987). Strategi Belajar Mengajar Bahasa Indonesia. (Cetakan Pertama). Jakarta: Karunika.

Yudhi Munadi. (2007). Bahan Ajar Media Pembelajaran. Naskah. Jakarta: PSWUIN Syarif Hidayatullah Jakarta.

Japan; an Illustrated Encyclopedia.(1993). Tokyo: kodansha Ltd. Him. 731

http:/'Aepository.upi.edu/operator/upload/s_prs_981119 chapter2.pdf. accessed on 23 April (2011)

An Online Survey: Studying the Antecedents of Technology Use through the UTAUT Model among Arts and Science Undergraduate Students

Priscilla MOSES[a*], Tiny Chiu Yuen TEY[b], Phaik Kin CHEAH[c], Timothy TEO[d] and Su Luan WONG[e]
[a, b, c] *Universiti Tunku Abdul Rahman, Malaysia*
[d] *University of Macau, ChinaSAR*
[e] *Universiti Putra Malaysia, Malaysia*
*priscilla@utar.edu.my/prismo6@yahoo.com

Abstract: Students need to be well-equipped with the necessary information, understanding, capabilities, skills and awareness to learn a subject and simultaneously to optimize the use of technology. For that reason, this research studied the antecedents of students' technology use through Unified Theory of Acceptance and Use of Technology (UTAUT) model. Besides that, this paper sought to explore whether there is a difference between Arts and Science undergraduate students in terms of technology use. There were 38 Arts and 30 Science undergraduate students who participated in this online survey. Based on the independent-samples t-test, there was no significant difference (t (66) =.558, p =.579) found in terms of technology use among the Arts (M =5.772, SD =.653) and Science (M =5.661, SD =.980) students. The magnitude of the differences obtained was very small. Therefore, the findings of the study suggest that both the Arts and Science students make use of the technology regardless of their major.

Keywords: Antecedents, technology use, UTAUT model, online survey

1. Introduction

One of the key players to successfully integrate technology into the education system is the student. They need to be well-equipped with the necessary information, understanding, capabilities, skills and awareness to learn a subject and simultaneously to optimize the use of technology. Technology is not only a medium to deliver or receive knowledge, but it also acts as a vehicle that helps students to travel along the pathway to prepare them for their future. According to Godin and Goette (2013), students who graduate these days regardless of their major, need to have the capabilities to work in a global marketplace and use whatever technology that is needed to work virtually.

There is also an increasing need for educators to incorporate technology in teaching and learning in universities in Malaysia. Identifying the differences would help the educators to address the challenges faced by Arts and Science students in the teaching pedagogy. A study conducted among undergraduates found that there was no significant difference in the overall scores between undergraduates from the Arts and Science disciplines in an ICT literacy course (Wong & Cheung, 2012). However, Liberal Arts and Business students were found to use less applications in their laptops compared to students who are in the Science disciplines (Percival & Percival, 2009). Despite the greater use of applications, another study reported that there was no significant difference in problem solving skills between Arts and Science students (Williamson, 2011).

Hence, this study sought to study the antecedents that influence the students' technology use through the Unified Theory of Acceptance and Use of Technology (UTAUT) model as a research framework. An online survey was employed to measure six constructs: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Condition (FC), Behavioural Intention (BI), and Use Behaviour (UB). Additionally, this study tested whether field of study (Arts and Science) plays a role among undergraduates' technology use.

2. Literature review

2.1 *The UTAUT Model*

Quite a number of theoretical models have been suggested to facilitate the understanding of factors impacting the user acceptance and usage behaviour of information technology. These models are universally used to predict and explain individuals' behaviours towards technology acceptance (Dulle & Minishi-Majanja, 2011), such as Theory of Reasoned Action (TRA), Theory of Planned Behaviour (TPB), Technology Acceptance Model (TAM), Motivational Model (MM) and so forth were incorporate in the area of perceived ease of use as a determinant of acceptance (Liu & Kostiwa, 2007). Among all the models, Technology Acceptance Model (TAM) is one of the most widely applied and influential models in explaining information technology adoption behaviour (Venkatesh & Davis, 2000).

Venkatesh, Morris, Davis, and Davis (2003) formulated the more recent instrument, Unified Theory of Acceptance and Use of Technology (UTAUT) model in which they included the eight well-known models - Motivational Model (MM), Theory of Planned Behaviour (TPB), Technology Acceptance Model (TAM), Theory of Reasoned Action (TRA), Model of PC Utilization, Innovation Diffusion Theory, Combined TAM-TPB, and Social Cognitive Theory. The UTAUT model does not only describe~~s~~ the main individual-level factors that influence technology acceptance, but the possibilities that would limit and amplify the influence of these factors (Venkatesh & Zhang, 2010). The credibility of the UTAUT model is established in explaining a large portion of variance in the user behaviour intention towards the use of technology (Venkatesh & Zhang, 2010) and it has been validated outside the origin where it was first proposed (Teo & Noyes, 2012). Besides, there are four constructs in UTAUT model which play key roles as direct determinants of user acceptance and usage behaviour - performance expectancy, effort expectancy, social influence and facilitating conditions (Venkatesh et al., 2003).

The UTAUT model presents three direct determinants (see Figure 1) to assess behaviour intention towards the use of technology (performance expectancy, effort expectancy, social influence), two direct determinants of technology use (behaviour intention and facilitating conditions), and four contingencies (age, gender, experience and voluntariness) affecting behaviour and/or intention towards the use of technology (Venkatesh & Zhang, 2010). However, the four contingencies in the UTAUT model were excluded in this research because they are moderating variables which affect the relationship between the determinants and technology use behaviour; while the focus in this research is to examine the direct factors that affect the undergraduates' technology use behaviour (Baron & Kenny, 1986; Brown, Dennis, & Venkatesh, 2010). In the present study, the researchers also sought to explore if field of study plays a role in technology use.

Over the last decade, the UTAUT model has been widely used to examine technology use in educational context, especially in e-learning and mobile learning (Cruz, Boughzala, & Assar, 2014; Lin, Lu, & Liu, 2013; Thomas, Singh, & Gaffar, 2013). According to Cassidy et al. (2014), technology evolution has impacted education as students' exposure to technology has increased dramatically including computer, mobile software, electronic gadgets and social networks. As Cassidy and her colleagues reported, students' technology use for academic purpose, such as the use of e-reader, has doubled in four years. Hence, technology evolution has also contributed to ubiquitous use and access in education. As said by Godin and Goette (2013), future studies should be conducted to examine the virtual learning and technology acceptance with the intention to comprehend better on how to prepare the students to collaborate virtually in a global environment by incorporating these determinants.

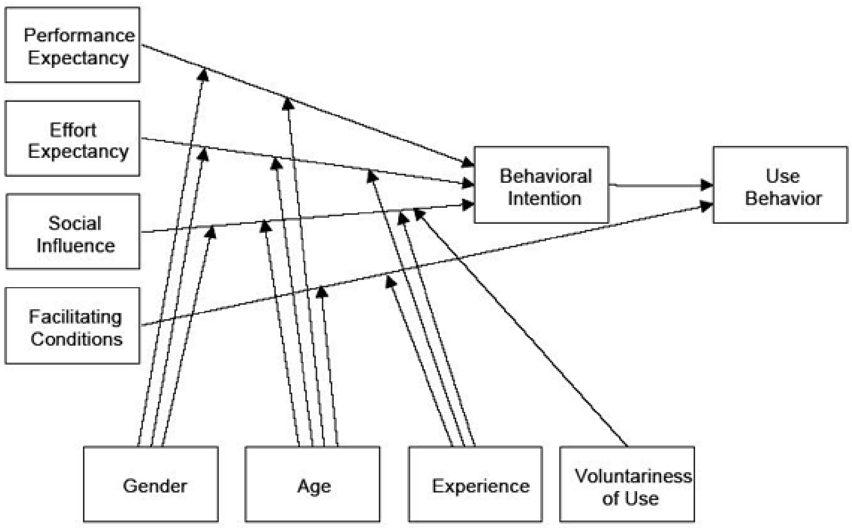


Figure 1. Unified Theory of Acceptance and Use of Technology (UTAUT) Model.

3. Purpose of the study

This study aims to achieve the following objectives: (a) to explore the antecedents that explain the students' technology use through the UTAUT Model, and (b) to test whether there is a significant difference in technology use between Arts and Science undergraduate students.

4. Methodology

4.1 Participants

The sample consists of 68 Arts and Science undergraduates from a private university located in peninsula Malaysia. There were 18 males (26.5%) and 50 females (73.5%) who participated in the online survey. Table 1 provides a summary of the undergraduates' ages. The undergraduates' age ranges from 20 to 27. The mean age of the participants is 22.26 with standard deviation of 1.39.

Table 1: Age

Age	Frequency	Percentage
20	8	11.8
21	9	13.2
22	21	30.9
23	23	33.8
24	5	7.4
27	2	2.9
Total	68	100.0

In addition, Table 2 shows the undergraduates' majors. There were 38 (55.9%) Arts undergraduates from Faculty of Arts and Social Science while the remaining 30 (44.1%) were Science undergraduates from the Faculty of Science.

Table 2: Majors

Major	Frequency (f)	Percentage (%)
Faculty of Arts and Social Science	38	55.9
Faculty of Science	30	44.1
Total	68	100.0

4.2 Research Instrument

The online survey was adapted from the UTAUT model instrument which was developed by Venkatesh, et al. (2013). In this research, the online survey was designed using Google Form. There were two sections in the online questionnaire with a total of 28 items. The respondents filled in their age and major in the first section and clicked on an appropriate option (7-point likert scale from "Strongly Disagree" to "Strongly Agree") for the second section. Subsequently, their responses were recorded and submitted to a Web server, which was used to administrate the online survey.

Besides that, a reliability analysis was executed for the scales using Cronbach's Alpha. As summarised in Table 3, all of the scales tested in the UTAUT constructs were reliable as each computed statistic showed a value above .70 ranging from .70 to .96. The Cronbach's Alpha value of the questionnaire with 28 items was reported to be .95.

Table 3: Instrument Reliability

Scales	Number of Items	Cronbach's Alpha
Performance Expectancy	4	.85
Effort Expectancy	4	.88
Social Influence	4	.79
Facilitating Conditions	5	.88
Behavioural Intention	5	.96
Use Behaviour	6	.70

5. Results and Discussion

A descriptive statistical analysis describing the antecedents of undergraduates' technology use is presented in Tables 4, 5, 6, 7, 8 and 9. Table 4 is a summary of descriptive analysis for the undergraduates' performance expectancy. As demonstrated in Table 4, the statistics suggest that the undergraduates perceive technology as an effective tool that enhances their studies and task accomplishment and productivity. Nevertheless, it appears that the undergraduates tend to be more neutral with respect to the perception that using technology will improve their academic grades.

Table 4: Descriptive Statistics for Performance Expectancy (PE) (n= 68)

Questionnaire Item	1 Strongly Disagree	2 Disagree	3 Slightly Disagree	4 Neither Agree Or Disagree	5 Slightly Agree	6 Agree	7 Strongly Agree	Mean	Std. Dev.
SI1: People who influence my behaviour think that I should use technology.	0 (0%)	3 (4.4%)	14 (20.6%)	21 (30.9%)	18 (26.5%)	8 (11.8%)	4 (5.9%)	4.38	1.23
SI2: People who are important to me think that I should use technology.	2 (2.9%)	5 (7.4%)	8 (11.8%)	21 (30.9%)	15 (22.1%)	13 (19.1%)	4 (5.9%)	4.43	1.43

SI3: The administration of this university has been helpful in the use of technology.	1 (1.5%)	2 (2.9%)	6 (8.8%)	26 (38.2%)	20 (29.4%)	10 (4.7%)	3 (4.4%)	4.53	1.17
SI4: The university has supported the use of technology.	1 (1.5%)	1 (1.5%)	4 (5.9%)	18 (26.5%)	19 (27.9%)	19 (27.9%)	6 (8.8%)	4.97	1.23

Table 5 provides the descriptive analysis for undergraduates' effort expectancy. It shows that the undergraduates are confident in using technology as they believe that learning and operating technology is easy and understandable for them.

Table 5: Descriptive Statistics for Effort Expectancy (EE) (n= 68)

Questionnaire Item	1 Strongly Disagree	2 Disagree	3 Slightly Disagree	4 Neither Agree Or Disagree	5 Slightly Agree	6 Agree	7 Strongly Agree	Mean	Std. Dev.
EE1: My interaction with technology would be understandable.	1 (1.5%)	3 (4.4%)	2 (2.9%)	15 (22.1%)	21 (30.9%)	20 (29.4%)	6 (8.8%)	5.00	1.28
EE2: It would be easy for me to become skilful at using technology .	0 (0%)	1 (1.5%)	4 (5.9%)	13 (19.1%)	19 (27.9%)	19 (27.9%)	12 (17.6%)	5.28	1.22
EE3: I would find technology easy to use.	0 (0%)	2 (2.9%)	4 (5.9%)	16 (23.5%)	18 (26.5%)	17 (25.0%)	11 (16.2%)	5.13	1.28
EE4: Learning to operate technology would be easy for me.	0	3 (4.4%)	4 (5.9%)	13 (19.1%)	21 (30.9%)	17 (25.0%)	10 (14.7%)	5.10	1.30

Table 6 represents the undergraduates' perceptions on social influence towards their technology use. The descriptive statistics suggest that the undergraduates are neutral in terms of their perception that important people around them and the university administration might affect their technology use.

Table 6: Descriptive Statistics for Social Influence (SI) (n= 68)

Questionnaire Item	1 Strongly Disagree	2 Disagree	3 Slightly Disagree	4 Neither Agree Or Disagree	5 Slightly Agree	6 Agree	7 Strongly Agree	Mean	td. Dev.
SI1: People who influence my behaviour think that I should use technology.	0 (0%)	3 (4.4%)	14 (20.6%)	21 (30.9%)	18 (26.5%)	8 (11.8%)	4 (5.9%)	4.38	1.23
SI2: People who are important to me think that I should use technology.	2 (2.9%)	5 (7.4%)	8 (11.8%)	21 (30.9%)	15 (22.1%)	13 (19.1%)	4 (5.9%)	4.43	1.43
SI3: The administration of this university has been helpful in the use of technology.	1 (1.5%)	2 (2.9%)	6 (8.8%)	26 (38.2%)	20 (29.4%)	10 (4.7%)	3 (4.4%)	4.53	1.17
SI4: The university has supported the use of technology.	1 (1.5%)	1 (1.5%)	4 (5.9%)	18 (26.5%)	19 (27.9%)	19 (27.9%)	6 (8.8%)	4.97	1.23

Table 7 is a summary of descriptive analysis for the facilitating conditions in undergraduates' technology use. It suggests that the undergraduates agree they have the necessary resources and knowledge to use technology. However, the undergraduates' perceptions are rather neutral toward the available assistance when they encounter technology use difficulties.

Table 7: Descriptive Statistics for Facilitating Conditions (FC) (n= 68)

Questionnaire Item	1 Strongly Disagree	2 Disagree	3 Slightly Disagree	4 Neither Agree Or Disagree	5 Slightly Agree	6 Agree	7 Strongly Agree	Mean	Std. Dev
FC1: I have the resources necessary to use technology.	0 (0%)	3 (4.4%)	4 (5.9%)	12 (17.6%)	26 (38.2%)	15 (22.1%)	8 (11.8%)	5.03	1.23
FC2: I have the knowledge necessary to use technology.	0 (0%)	1 (1.5%)	8 (11.8%)	10 (14.7%)	17 (25.0%)	22 (32.4%)	10 (14.7%)	5.19	1.28
FC3: When I encounter difficulties in using technology,	0	0	13	18	11	14	7	4.54	1.46

a specific person is available to provide assistance.	(0%)	(0%)	(19.1%)	(26.5%)	(16.2%)	(20.6%)	(10.3%)		
FC4: When I encounter difficulties in using technology, I know where to seek assistance.	0 (0%)	4 (5.9%)	6 (8.8%)	16 (23.5%)	17 (25.0%)	17 (25.0%)	8 (11.8%)	4.90	1.36
FC5: When I encounter difficulties in using technology, I am given immediate assistance.	1 (1.5%)	10 (14.7%)	12 (17.6%)	21 (30.9%)	15 (22.1%)	5 (7.4%)	4 (5.9%)	4.03	1.40

The undergraduates' behavioural intention in technology adoption is statistically described in Table 8. As demonstrated, the undergraduates will use technology in the future. Moreover, they also agree that they have positive intention to use technology often in future or in the next few months.

Table 8: Descriptive Statistics for Behavioural Intention (BI) (n= 68)

Questionnaire Item	1 Strongly Disagree	2 Disagree	3 Slightly Disagree	4 Neither Agree Or Disagree	5 Slightly Agree	6 Agree	7 Strongly Agree	Mean	Std. Dev.
BI1: I intend to use technology in the next few months.	0 (0%)	1 (1.5%)	5 (7.4%)	9 (13.2%)	15 (22.1%)	14 (20.6%)	24 (35.3%)	5.59	1.36
BI2: I predict I would use technology in the next few months.	0 (0%)	2 (2.9%)	3 (4.4%)	7 (10.3%)	19 (27.9%)	13 (19.1%)	24 (35.3%)	5.62	1.34
BI3: I plan to use technology in the next few months.	0 (0%)	1 (1.5%)	5 (7.4%)	8 (11.8%)	12 (17.6%)	19 (27.9%)	23 (33.8%)	5.65	1.34
BI4: I will use technology in the future.	0 (0%)	1 (1.5%)	0 (0%)	5 (7.4%)	11 (16.2%)	20 (29.4%)	31 (45.6%)	6.09	1.08
BI5: I plan to use technology often.	0 (0%)	2 (2.9%)	4 (5.9%)	6 (8.8%)	15 (22.1%)	17 (25.0%)	24 (35.3%)	5.66	1.36

The descriptive statistics in Table 9 entails the undergraduates' use behaviour. The analysis suggests that the undergraduates use technology for leisure, studies and daily communication. Interestingly, according to the statistics shown, the undergraduates' technology use for course-related work overrides the use for other purposes.

Table 9: Descriptive Statistics for Use Behaviour (UB) (n= 68)

Questionnaire Item	1 Strongly Disagree	2 Disagree	3 Slightly Disagree	4 Neither Agree Or Disagree	5 Slightly Agree	6 Agree	7 Strongly Agree	Mean	Std. Dev.
UB1: I check my email.	0 (0%)	6 (8.8%)	5 (7.4%)	3 (4.4%)	13 (19.1)	22 (32.4%)	19 (27.9%)	5.42	1.55
UB2: I communicate via instant messaging.	0 (0%)	0 (0%)	4 (5.9%)	7 (10.3%)	14 (20.6%)	22 (32.4%)	21 (30.9%)	5.72	1.18
UB3: I use the Internet for course-related work.	0 (0%)	1 (1.5%)	0 (0%)	4 (5.9%)	10 (14.7%)	25 (36.8%)	28 (41.2%)	6.08	1.01
UB4: I use the Internet for leisure.	0 (0%)	1 (1.5%)	4 (5.9%)	1 (1.5%)	11 (16.2%)	27 (39.7%)	24 (35.3%)	5.92	1.16
UB5: I talk on the phone.	0 (0%)	3 (4.4%)	7 (10.3%)	8 (11.8%)	21 (30.9%)	12 (17.6%)	17 (25.0%)	5.22	1.44
UB6: I use technology to communicate with others.	0 (0%)	1 (1.5%)	3 (4.4%)	4 (5.9%)	10 (14.7%)	22 (32.4%)	28 (41.2%)	5.95	1.20

Moreover, Table 10 is the summary of descriptive analysis for the UTAUT model. As shown in the table, use behaviour and behavioural intention scored the highest mean value (M =5.72) followed by performance expectancy (M =5.37, SD =1.00) and effort expectancy (M =5.13, SD =1.09). Meanwhile, social influence indicated the lowest mean value (M =4.58, SD =1.00) whereas facilitating conditions demonstrated the second lowest mean value (M= 4.74, SD= 1.12). This shows that most of the students either use or have the intention to use the technology for course-related work, relaxation, and communication regardless of their social influence and facilitating conditions.

Table 10: Descriptive Statistics for UTAUT Model (n= 68)

Variables	Mean	Standard Deviation
Performance Expectancy	5.37	1.00
Effort Expectancy	5.13	1.09
Social Influence	4.58	1.00
Facilitating Conditions	4.74	1.12
Behavioural Intention	5.72	1.20
Use Behaviour	5.72	.81

Findings of this study also reported that there is a strong positive correlation between performance expectancy (r= .695, p< .0005), effort expectancy (r= .635, p< .0005), social influence (r= .544, p< .0005) and behavioural intention to use technology. In addition, there is also a positive correlation between facilitating conditions (r= .538, p< .0005) and use behaviour; and medium positive correlation between behavioural intention (r= .496, p< .0005) and use behaviour. Thus, the credibility of the UTAUT model in investigating the antecedents that influence technology use among Arts and Science undergraduate students is continuously being proven (Venkatesh & Zhang, 2010). This is also consistent with the research done by Venkatesh et al. (2003) a decade ago.

Lastly, an independent-samples t-test was carried out to compare the technology use between the Arts and Science undergraduate students. There was no significant difference (t (66) =.558, p =.579) found in terms of technology use for Arts (M =5.772, SD =.653) and Science (M =5.661, SD =.980) students which found to be concurrent with Williamson's study (2011). The magnitude of the differences in the mean values (mean difference = .111) was very small (eta squared =.005).

6. Conclusion

Integrating technology in teaching and learning is to some extent an expectation in tertiary education. Technology is no longer regarded as novelty but a standard feature in the delivery of a course in tertiary institutions in Malaysia. However, there is a need to understand and identify the antecedents of technology use among graduates to help educators and education managers address the challenges and concerns experienced by them.
However, this study was conducted with a modest sample size from two faculties within a university in Malaysia. Therefore it is not representative of the scenario in Malaysia. However, future research could explore the possibility of expanding the sample size or comparing Arts and Science undergraduates from different universities located in different countries.

Acknowledgements

We would like to thank Universiti Tunku Abdul Rahman, Malaysia for providing us the fund to conduct this project. This paper is also a part of the abstract entitled "Predictors of Behavioral Intention to Use Technology among Undergraduate Students" submitted and presented in the 2nd International Conference on Behavioral & Social Science Research 2014 under the same funding.

References

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51(6), 1173-1182. doi: 0022-3514/86/$00.75

Brown, S. a., Dennis, A. R., & Venkatesh, V. (2010). Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research. Journal of Management Information Systems, 27(2), 9–54. doi:10.2753/MIS0742-1222270201

Cassidy, E. D., Colmenares, A., Jones, G., Manolovitz, T., Shen, L., & Vieira, S. (2014). Higher education and emerging technologies: Shifting trends in student usage. *The Journal of Academic Librarianship*, *40* (2), 124–133. doi:10.1016/j.acalib.2014.02.003

Cruz, Y., Boughzala, I., & Assar, S. (2014). Technology acceptance and actual use with mobile learning: First stage for studying the influence of learning styles on the behavioral intention. *Twenty Second European Conference of Information Systems, Tel Aiv,* 1–16.

Dulle, W. F., & Minishi-Majanja, M. K. (2011). The suitability of the Unified Theory of Acceptance and Use of Technology (UTAUT) model in open access adoption studies. *Information Development*, *27* (1).

Godin, J., & Goette, T. (2013). A review of models: Virtual teamwork training model and UTAUT. *Proceedings of the Southern Association for Information Systems Conference*, Savannah, GA, USA, 41-46.

Lin, P. -C., Lu, H. -K., & Liu, S. -C. (2013). Towards an education behavioural intention model for e-learning systems: An extension of UTAUT. *Journal of Theoretical and Applied Information Technology*, *47*(3), 1120-1127.

Liu, J. T., & Kostiwa, K. (2007). An application of the utaut model for understanding students perceptions using course management software. *Communications of the IIMA, 7*(2).

Percival, J., & Percival, N. (2009). A case of a laptop learning campus: how do technology choices affect perceptions?. *ALT-J: Research in Learning Technology, 17*(3), 173-186.

Teo, T., & Noyes, J. (2012). Explaining the intention to use technology among pre-service teachers: A multi-group analysis of the Unified Theory of Acceptance and Use of Technology. *Interactive Learning Environments, iFirst Article*, 1–16.

Thomas, T. D., Singh, L., & Gaffar, K. (2013). The utility of the model UTAUT model in explaining mobile learning adoption in higher education in Guyana. *International Journal of Education and Development Using Information and Communication Technology (IJEDICT), 9* (3), 71-85.

Venkatesh, V., & Zhang, X. J. (2010). Unified Theory of Acceptance and the Use of Technology: US vs China. *Journal of Global Information Technology Management*, *13* (1), 5–27.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003, September). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science, 46* (2), 186-204.

Williamson, P. K. (2011). The creative problem solving skills of arts and science students--the two cultures debate revisited. *Thinking Skills and Creativity, 6* (1), 31-43.

Wong, K. K., & Cheung, W. K. (2012). Evaluating ict literacy differences in arts and science students. *Journal of Next Generation Information Technology, 3*(1), 1. doi:10.4156/jnit.vol3.issue1.1

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Equipping High School Students with the Abilities of Evaluating Evidence and Formulating Evidence for an On-line Decision-making Task

Shu-Sheng LIN[a*], Ying-Shao HSU[b]
[a]Graduate Institute of Mathematics and Science Education, National Chiayi University, Taiwan
[b]Graduate Institute of Science Education, National Taiwan Normal University, Taiwan
*lin-s-s@mail.ncyu.edu.tw

Abstract: Preparing students with appropriate abilities for taking some of on-line tasks is necessary. It will help students to build self-confidence and increase the probability to reach the success of learning. The purpose of this study aimed at equipping high school students with the skills of evaluating evidence and formulating evidence for taking an on-line task, in which students had to make a decision on choosing a location to build a reservoir. The developed instructional activity was a 4-hour unit, which provided a socioscientific context for students to understand the concept about evidence and discuss the reliability and validity of evidence to support if the global warming has been accelerating. The participant consisted of one earth science teacher and forty students. Two questionnaires and individual interviews were used to collect data. The results showed that the students had significant improvement in evaluating evidence, formulating evidence and justifying their arguments. Their understandings of criteria used for evaluating evidence became more clearly after the teaching. Most students reported that the instruction was beneficial for them to complete the decision-making task.

Keywords: Making decision, evidence evaluation, evidence formulation

1. Introduction

If a decision maker wants to make an evidence-based decision, he/she have to find evidence, evaluate evidence, use evidence, and justify his or her decision with evidence. These abilities are important but have been paid little attention in science curriculum and instruction. Therefore, Gott and Duggan (2007) advocated that science education has to get procedural and declarative understanding of evidence involved in science instruction. Prior studies have revealed that many of the students, from elementary students to university students, have difficulties in using evidence to support their arguments (Sandoval, Sodian, Koerber, & Wong, 2014), coordinating the claim or conclusion with the evidence (Zimmerman, 2007), and evaluating evidence (Nicolaidou, 2011). These abilities regarded as higher cognitive thinking skills are vital for students to deal with some issues, especially they are asked to make a decision in a socioscientific context. The socioscientific issue is one kind of scientific and social issues, which embed with some problems, controversies and dilemma caused by the application and development of science and technology (Zeidler et al, 2005), such as genetically modified organisms, radiation of mobile phones, or building a reservoir etc. People argue the solutions of the problem in an issue without reaching a conclusion. Zeidler and Nichols (2009) suggested that science teachers can select an appropriate socioscientific issue (SSI) for instruction, lead students to examine and discuss the arguments the stakeholders have, evaluate evidence of each argument with criteria, practice making a decision and justifying their decision with evidence. It means before making a decision on a SSI, there are many steps regarding understanding of and using of evidence involved in the process that are difficult for most of students. Hug and McNeil (2008), Schalk, Van der Schee and Boersma (2013) suggested that it is helpful for students to make a deliberate decision after they experience the instruction planned to improve their understanding of evidence and skills of evaluating and using evidence. In this study, we attempt to equip high school students with the skills of evaluating evidence and formulating evidence before they take an on-line decision-making task.

2. Method

2.1 Participants

The participants consisted of one experienced teacher and forty grade eleven students (27 girls and 13 boys). The earth science teacher has 20-year teaching experiences. She joined the workshop held by the researcher to learn the declarative and procedural knowledge about evidence, the operation and the contents of software, and to discuss the teaching materials and methods with the researcher. The students whose age was 16-17 years old did not have the formal experiences to evaluate evidence, formulate evidence and justify arguments with evidence before.

2.2 The Instruction

The instruction was to enhance the students' understanding of the concept of evidence, and improve their abilities to evaluate evidence, formulate evidence and justify arguments with evidence. The instructional unit included four hours. At the first two hours, the teacher led the students to discuss the importance of evaluating and using evidence in everyday life. After the criteria used to evaluate the reliability and validity of evidence were developed through group discussion, the students assessed the criteria each group formulated for their appropriateness through the whole class discussion. During the last two hour, one text provided for the students to read includes five stakeholders and their arguments to the question – "Global warming becomes more serious than before. Is it man-made? Each argument had at least one piece of evidence to support that it is caused by man-made. The students were led to discuss the reliability of evidence, the relevance between the evidence and the claim, and how to make much stronger evidence to support or rebut the claim and justify their arguments.

2.3 The software and the task

The students were asked in a software environment to complete an evidence-based decision-making task, in which they had to choose an appropriate location to build a reservoir within a limited time. During the trial-and-error process of making decision, they had to use the abilities of formulating criteria, selecting and evaluating evidence, and justify their choice with evidence. Figure 1 shows one of the interfaces of "Constructing Reservoir" software.

Fig.1 The interface of "Constructing Reservoir" software

2.4 Instruments

The questionnaire "Evidence evaluation and use" included a scientific research context and three questions. The context described three animal studies on exploring the function of onions, the results and the consistent conclusions. Three questions were used to assess the students' abilities to generate an argument, evaluate the reliability of the evidence, and formulate evidence to support the arguments. The questions are: (1) Do you agree or disagree with the conclusion of three animal studies? Why? (2) Do you think the evidence the author described is reliable to support the conclusion? Explain why in detail. (3) If the scientists can make more evidence to support the conclusion, do you think what it is? Explain your reasons. The pre- and posttest administered to the students before and after the instructional intervention were the same. The other questionnaire is "Reflection on Learning" used to collect the students' feedback about the instruction. It was related to their attitude towards the teaching contents, methods and their reflection on learning for completing the decision-making task.

3. The Results

A series of t-test were run to examine the improvement of the students' abilities after instructional intervention (Table 1). The findings showed that the students had statistically significant improvement in scores for making warrants, evaluating the reliability of the evidence, and formulating evidence to support the arguments (p<.01).

Table 1 Summary of the paired t-test for the scores of making warrants, evidence evaluation and formulation

Questions	Mean(S.D.)		*t* value
	Pretest N=40	Posttest N=40	
Making warrants	1.40(0.81)	2.15(1.19)	4.39(0.00**)
Evaluating evidence	1.35(0.92)	2.13(0.97)	5.69(0.00**)
Formulating evidence	1.43(0.68)	2.17(0.96)	4.13(0.00**)
Total Scores	4.18(1.89)	6.45(2.09)	12.23(0.00**)

Figure 2 and 3 respectively revealed that the criteria the students used for evaluating evidence were a few different before and after the instruction. The criteria of "time" used in the posttest instead of in the pretest. The criteria the students used for formulating evidence were the same in the pretest and posttest. The number of criteria appeared in posttest had been significantly increased than in the pretest in both abilities.

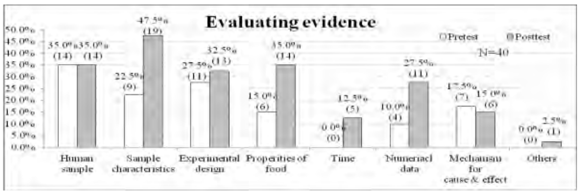


Fig2. The criteria the students used for evaluating evidence in pretest and posttest

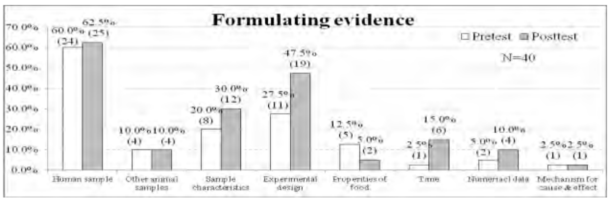


Fig3. The criteria the students used for formulating evidence in pretest and posttest

Meanwhile, according to the feedbacks the students expressed in the questionnaire, most of the students pointed that they benefited a lot from the instruction. For example, the student S03 said that "if without the instruction before I took the decision-making task, I nearly did not have the idea to use criteria to evaluate and select evidence to help myself to complete it."

4. Conclusions

The instructional design in this study supports the students to develop the abilities in making arguments, evaluating evidence and formulating evidence to support their arguments. Based on the students' feedback and individual interviews, it is really helpful for the students to apply these abilities to take the decision-making task in the software environment.

However, it is not enough to proof that all of the students transfer these abilities well for taking the task in this pilot study. Therefore, the researcher further plans to adopt " two-group pretest-posttest experimental design". The experimental group receives the instruction this study showed. The control group receives the instruction without emphasizing on learning to develop criteria for evidence evaluation. Through the comparison we can confirm the effect of instructional intervention. Moreover, the researchers will examine what strategies the students adopt during the period of taking the task. It will reveal the abilities the students apply in the task.

References

Gott, R. & Duggan, S. (2007). A framework for practical work in science and scientific literacy through argumentation, *Research in Science and Technological education, 25*(3), 271-291.

Hug, B. & McNeil, K. L. (2008). Use of first-hand and second-hand data in science: Dose data type influence classroom conversations? *International Journal of Science Education, 30*(13), 1725-1751.

Nicolaidou, I., Kyza, E.A., Terzian, F., Hadjichambis, A., & Kafouris, D. (2011). A framework for scaffolding students' assessment of the credibility of evidence. *Journal Research in Science Teaching, 48*(7), 711-744.

Sandoval, W. A., Sodian, B., Koerber, S., & Wong, J. (2014). Developing children's early competencies to engage with science. *Educational Psychologist, 49*(2), 139-152.

Schalk, H. H., van der Schee, J. A., & Boersma, K. T. (2013). The development of understanding of evidence in pre-university biology education in the Netherlands. *Research in Science Education, 43*, 551-578.

Zeidler, D., & Nichols, B. H. (2009). Socioscientific issues: Theory and practice. *Journal of Elementary Science Education, 21*(2), 49-58.

Zeidler, D., Sadler, T. D., Simon, M. L., Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education, 89*(3), 357-377.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27*(2), 172-223.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

An Investigation of Relationships between Biology Attitudes and Perceptions toward Instructional Technology in Analogy-based Simulation on Light Reaction

Sarunya PINATUWONG[a] & Niwat SRISAWASDI[b*]
[a] *Science and Technology Education Program, Faculty of Education, Khon Kaen University, Thailand*
[b] *Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Thailand*
*niwsri@kku.ac.th

Abstract: Computer-simulated scientific phenomena have become an indispensable tool in modern scientific investigations and in contemporary science class. In context of science-based education, computer simulation or visualization are now commonly used to promote student's meaningful understanding in science concepts, and motivation to learn science. This study presents the development of an analogy-based simulation, as a novel pedagogy-oriented simulation, for biology learning of light reaction phenomenon. 46 eleventh-grade students were recruited to participate interacting with the analogy-based simulation. The 25-item Likert-scale questionnaire on attitude toward biology lesson scale was administered to the student before interacting the simulation, and another 21-item Likert-scale questionnaire on perception toward instructional technology was completed by the student after the interacting. Overall finding of this study suggests that students' attitudes towards biology did not determine their perceived learning, flow, enjoyment, perceived ease-of-use, perceived usefulness, and satisfaction, delivered by the analogy-based simulation. As such, this study concludes that the analogy-based simulation could be used in biology class disregarding students' attitudes toward biology. The findings could be implied for designing effective computer simulation to facilitate biology teaching and learning in school science.

Keywords: Computer simulation, analogy, perception, attitude, biology education

1. Introduction

Currently, the prospect of research on the use of ICT in education in general, or even in the specific case of science education is widespread, especially studies on the use of computer simulations. Nowadays, computer simulation has become increasingly powerful and available to teachers in the past decade (de Jong, Linn and Zacharia, 2013; Trundle and Bell, 2010). For instructional context, computer-simulated technology has been used to facilitate teaching and learning by visualizing objects, processes, and interact dynamics models of natural phenomenon, that are normally beyond the user' control in the natural world (de Jong, Linn and Zacharia, 2013; Perkins et al., 2006; Wieman, Perkins and Adams, 2008). These technology offer idealized, dynamic and visual representations of invisible phenomena and experiments which would be dangerous, costly or otherwise not possible in school laboratories (Hennessy, 2007). Since, simplified versions of the natural world were showed by computer simulation, students' attention can be more focus directly on the desired phenomenon (de Jong and van Joolingen, 1998; Perkins et al., 2006; Wieman, Perkins and Adams, 2008).

As such, successful concepts of simulation-based teaching and learning have been reported by means of discovery learning (de Jong, Linn and Zacharia, 2013) and inquiry-based learning (Perkins et al., 2006; Wieman, Perkins and Adams, 2008). The researchers have interested to science learning with analogy, which is an effective pedagogy to assist students learning about an unfamiliar concept, system, or process, called *target* by means of its relationship to a familiar concept, system, or process, called *analog*. Based on literature reviewing, very few study (i.e. Ashe and Yaron, 2013; Ünlü and Dökme, 2011) has investigated effects of analogy-based simulation that may influence students' conceptual development in science. Especially, considering to the nature of biology knowledge, lot of invisible processes and biological mechanism that occurred in organisms was presented to student and they encountered with learning difficulties about biology. Teaching biology through analogy is an instructional idea with aim to help student learn biology meaningfully. Schiff (1970) stated that perception and attitudes were related together in which attitudes affect perception, perception affects attitudes, and cognition plays a role in both of them. An important consequence of instruction is the student's attitude toward the subject and the previous study found that there is usually positive correlation between attitudes and achievement, but researcher cannot assume a positive attitude on the basis of achievement alone (Russell and Hollander, 1975). It is possible for a student with low achievement to have developed a very positive attitude toward the subject matter, but it is also possible that a student who indicated on post-test that they have learned the subject matter well may also have learned to dislike or, worse yet, hate the content (Russell and Hollander, 1975). From these reasons, the researchers have created a computer simulation emphasized analogy approach on biology concept of light reaction and an evaluation of students' perception towards the simulation regarded their existing biology attitudes.

2. Analogy-based Simulation

Analogy-based simulation is computer-simulated visualization that does not directly depict atoms or molecules; rather, they simulate a different physical system that relates to a chemical system or concept by analogy. For examples, Ashe and Yaron (2013) created computer simulation visualized chemistry phenomena by making analogies between chemical systems and familiar objects from students' everyday experiences, e.g. boxes, steps, and balls. Because students are not typically experienced in reasoning qualitatively about chemical systems, the analogy-based simulation aims to leverage students' experience with the world around them to help them better understanding of scientific concepts (Ünlü and Dökme, 2011). As the way it was created and students are interactive and dynamic by their nature, analogy-based simulations are likely to be more engaging to students than static, non-interactive presentations of analogies. Moreover, since students cannot benefit from an analogy if they do not engage with it, simulations may offer an advantage here. There are dual-situated events in which analogy could be used to incorporate into computer simulation. Firstly, analogies were used to simulate concepts which are not possible to simulate directly with portrayal of atoms or molecules, or which could be simulated more clearly by analogy. Another, analogy could be used to address students' understanding for important knowledge abstractions (Ashe and Yaron, 2013).

3. Methods

3.1 Participants

A total of 46 student-respondents in their eleventh grade, age ranging from 16 to 17 years in a local public school at the Northeastern region of Thailand participated in this study. They have no experience yet using analogy-based simulation. This implied that they are heterogeneous before interacting with the simulation.

3.2 Learning Materials

Figure 1 illustrates examples of analogy-based simulation on light reaction phenomenon of plant photosynthesis. As aforementioned, target and analog are important concepts for analogy approach, and both were created into the computer simulation. In this analogy-based simulation, the target event was electron transfer during photosystem I (PS-I) of light reaction and the analog event was coal transfer on a pirate ship. In an addition, there were two parameters which student could vary for their investigation; wavelength of light and light intensity.

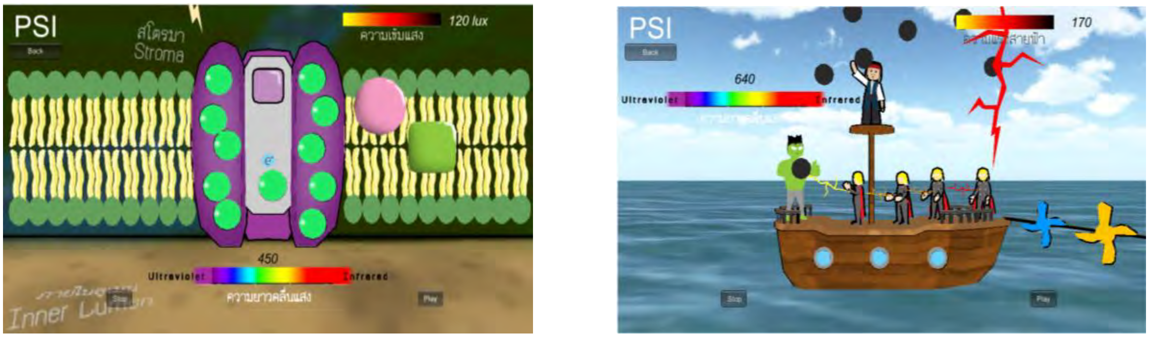


Figure 1. Illustrative example of target (Left) and analog (Right) event in analogy-based simulation.

3.3 Procedures

Before exposing to the analogy-based simulation, the students took a 25-item five-point Likert-scale questionnaire on attitude toward biology lesson scale for 10 minutes. A 25-item Likert-scale questionnaire was developed to use in this study for examining students' interest in biology lessons (IBL), understanding and learning biology (ULB), importance of biology in real-life (IBR), biology and occupational choice (BOC) toward biology learning. There were six items of IBL, 10 items of ULB, five items of IBR and four items of BOC (Ayyıldız and Tarhan, 2013). To develop a Thai version of the questionnaire, the original English version was translated identically in Thai language. One expert was recruited to identify communication validity of the items. On each item, respondents were assigned to rate how the respondent agree with into five scale, from 1-strongly disagree to 5-strongly agree. The reliability for Thai version on IBL, ULB, IBR, BOC was 0.79, 0.70, 0.70, 0.65 respectively, and the overall was 0.76. To experience with the analogy-based simulation, they were assigned to interact with the simulation independently for 30 minutes. After that, the students took 21-item five-point Likert-scale questionnaire measured perception toward instructional technology for 10 minutes. The questionnaire was developed to examine students' perceived learning (PL), flow (FL), enjoyment (EJ), perceived ease of use (PEOU), perceived usefulness (PU), and satisfaction (ST). There were four items of PL, five items of FL, three items for EJ obtained from Barzilai and Blau (2014), and three item of PEOU, three items of PU, three items of ST obtained from Cheng (2014). The reliability for Thai version on PL, FL, EJ, PEOU, PU, ST was 0.80,0.82, 0.75, 0.74, 0.84, 0.77 respectively and the overall was 0.95.To examine correlation between students' attitude and perception toward instructional technology, Pearson's correlation was performed in SPSS 21.0.

4. Results and Discussion

Table 2 shows Pearson's correlations among the variables. Regarding the Pearson's correlation analysis of attitudes towards biology, the results indicated that correlation among PL, FL, EJ, PEOU, PU, and ST, reveals a significant positive correlation (*p*-value < 0.01), as well as the correlation among IBL, ULB, IBR, and BOC in perceptions towards instructional technology (*p*-value < 0.01). However, there was no significant correlation between variables of attitudes towards biology and variables of perceptions towards instructional technology, except the product-moment correlation between perceived ease of use (PU) and interest in biology lessons (IBL), r=0.294, *p*-value < 0.05. For the overall summarization, these findings suggest that students' attitudes towards biology did not determine their perceptions towards instructional technology delivered by the analogy-based simulation. This means the use of analogy-based simulation for biology learning could support students' perceptions even individual student who have a positive or negative biology attitudes. This evidence consistent with previous research finding that students perceived learning with simulation offer instructional advantages and they can gain benefits for their experiences (Baillie and Curzio, 2009; Prokop, Prokop and Tunnicliffe, 2007). Furthermore, the use of instructional technology could be a challenge in teaching methods for providing interactive classroom activities and stimulate students' engagements and attitudes toward science (Hansen and Birol, 2010).

Table 1: Descriptive and Pearson product-moment correlation matrix of attitudes towards biology lesson and perceptions towards instructional technology.

variable	PL	FL	EJ	PEOU	PU	ST	IBL	ULB	IBR	BOC
PL	-									
FL	.773**	-								
EJ	.783**	.764**	-							
PEOU	.715**	.753**	.732**	-						
PU	.773**	.758**	.781**	.779**	-					
ST	.711**	.611**	.701**	.720**	.732**					
IBL	.188	.273	.163	.097	.294*	.147	-			
ULB	.041	.167	.238	-.009	.117	.100	.501**	-		
IBR	.058	.104	.005	.028	.125	.144	.647**	.454**	-	
BOC	.068	.261	.138	.062	.262	.109	.866**	.487**	.616**	-
Mean	16.61	19.72	12.70	12.02	12.26	12.59	21.11	31.80	17.80	13.74
SD	2.55	3.13	1.90	2.04	2.08	1.81	3.70	5.59	2.76	2.49

** *p*-value < 0.01; * *p*-value < 0.05

5. Conclusion

To facilitate students' learning in biology, an analogy-based simulation was created to assist students learning about an unfamiliar concept, system, or process. The result indicated that the use of analogy-based simulation for biology learning to students could be benefits on their perceptions even holding a positive and negative attitude towards biology lessons. Thus, it is crucial to design computer-simulated visualization by the use of analogy approach to scaffold students’ perceptions. To this end, researchers should pay more attention on principle of analogy when designing computer simulation to enrich the learning content for biology learning in school science.

Acknowledgements

This work was financially supported by Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Ashe, C.A., & Yaron, D.J. (2013). Designing analogy-based simulations to teach abstractions. In J.P. Suits, & M.J. Sanger (Eds.), *Pedagogic roles of animations and simulations in chemistry courses* (pp. 367-388). Washington, DC: American Chemical Society.

Baillie, L., & Curzio, J. (2009). Students’ and facilitators’ perceptions of simulation in practice learning. *Nurse Education in Practice, 9*, 297-306.

de Jong, T., Linn, M.C., & Zacharia, Z.C. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305-308.

Hennessy, S., Wishart, J., Whitelock, D., Deaney, R., Brawn, R., et al. (2007). Pedagogical approaches for technology-integrated science teaching. *Computers & Education, 48*(1), 137-152.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., et al. (2006). PhET: Interactive simulations for teaching and learning physics. *The Physics Teacher, 44*(1), 18-23.

Prokop, P., Prokop, M., & Tunnicliffe, S.D. (2007). Is biology boring? Student attitudes toward biology. *Journal of Biological Education, 42*(1), 36-39.

Russell, J., & Hollander, S. (1975). A biology attitude scale. *The American Biology Teacher, 37*(5), 270-273.

Schiff, M. (1970). *Some theoretical aspects of attitudes and perception*. Toronto: Natural Hazard Research.

Trundle, K.C., & Bell, R.L. (2010). The use of a computer simulation to promote conceptual change: a quasi-experimental study. *Computer & Education, 54*(4), 1078-1088.

Ünlü, Z.K., & Dökme, I. (2011). The effect of combining analogy-based simulation and laboratory activities on Turkish elementary school students’ understanding of simple electric circuits. *The Turkish Online Journal of Educational Technology, 10*(4), 320-329.

Wieman, C.E., Perkins, K., & Adams, W.K. (2008). Oersted Medal Lecture 2007: Interactive simulations for teaching physics: What works, what doesn’t, and why. *American Journal of Physics, 76(*4), 393-399.

Visual Behavior and Cognitive Load in Augmented Reality Learning Environment

Wei-Yan LIN *, Meng-Jung TSAI & An-Hsuan WU
Graduate Institute of Digital Learning and Education
National Taiwan University of Science and Technology, Taiwan
*M10211004@mail.ntust.edu.tw

Abstract: This pilot study was attempted to explore the relationships among cognitive load, visual behavior, and reading comprehension in augmented reality (AR) learning environment. An ASL MobileEye-XG eye-tracking system was used to record the participants' visual behaviors while they read Taiwan traditional folk belief information demonstrated by AR technologies. A survey was used to examine participants' cognitive loads due to different cognitive elements (videos, texts and pictures) designed in the AR contents. Currently, the results showed that participants with higher reading comprehensions seemed to have higher cognitive loads from videos. Although no significant relationships were found between visual behaviors and reading comprehensions, some significant correlations were found between visual behaviors and cognitive loads, and also between cognitive loads and reading comprehensions. More results will be discussed in the conference.

Keywords: Cognitive load, visual behavior, reading comprehension, augmented reality

1. Introduction

Augmented reality (AR) technologies, which utilize image recognition technology to link computer-generated virtual elements with real-world elements and display two types of elements on the same platform at the same time, are gradually promoted to be used in educational settings (Chen & Tsai, 2012; Wu, Lee, Chang, & Liang, 2013) in different learning domains. This study was attempted to use augmented reality technology in Taiwanese folk belief exhibition to explore its possibility of use in multicultural education, especially for the cultural inheritance of Taiwan traditional folk beliefs.

Eye movement studies have been applied to learning for decades and have caused increasing attentions in educational research recently (Lai et al., 2013). Visual attentions recorded by eye-tracking systems have been found to be significantly related to learning performance in environments where pictures and texts were displayed simultaneously (Yang, Chang, Chien, Chien, & Tseng, 2013; Ho, Tsai, Wang, & Tsai, 2014). However, little studies explored the situations that videos were added simultaneously as used in AR learning environments. Thus, this study added video into learning environment along with picture and text. Besides, eye tracking measures seemed to be found relating to cognitive load, but didn't analyze which type of cognitive load related to participants' eye movements (Wang, Yang, Liu, Cao, & Ma, 2014). According to the literature, cognitive load can be classified into three different types, i.e., intrinsic cognitive load, extraneous cognitive load, and germane cognitive load (Sweller, van Merriënboer, & Paas, 1998). Therefore, this pilot study aimed to explore the relationships among different types of cognitive loads, visual behaviors, and reading comprehensions in a Taiwanese folk belief exhibition which includes texts, pictures, and videos linked by AR.

2. Method

2.1 Experiment

Seven university students, four males and three females, in north Taiwan were volunteered to participate in a pilot eye-tracking experiment. The average age of the participants was 23.71(SD =

0.756). All of them had little knowledge regarding Taiwanese traditional folk beliefs and little experience in participating related traditional cultural events. Therefore, the participants had about the same level of prior knowledge or experience regarding the learning contents.

The learning material of this study was set as a real Taiwanese folk belief exhibition. There were eight ceremonies' introductions shown as posters in this exhibition. Every ceremony was introduced by one picture with a short caption below, one paragraph of text aside the picture that describes how the ceremony held, and one video showing the procedure of real ceremony situation linked by the augmented reality technology. The videos were designed to be activated by markers on posters and shown in a smart phone.

2.2 Instrument

In order to investigate different types of cognitive load from different types of materials (videos, texts, pictures), nine questions were designed to examine the types of cognitive loads. A self-reported 0-10 scale was used to evaluate each of the questions. The larger the number, the higher the cognitive load was perceived. Besides, an ASL MobileEye-XG system (60 Hz) was used to record the participants' eye movement data during the participants viewing the AR exhibition. After the experiment, a video analysis was used to analyze each participant's visual behavior data. Total reading time on each type of learning materials, total entered count on each type of learning materials, and average entered duration of each type of learning materials were calculated based on the scan path video of each participant. There were total twelve visual behavior measures coded in current analysis. Finally, the posttest in this study was assessed by twenty multiple-choice questions concerning the concepts demonstrated in the AR exhibition.

2.3 Data Collection and Analyses

The procedure of this study's experiment can be divided into four parts. First, each of the participants was taught how to link the video by AR with a smart phone. Second, an eye-tracking calibration was then conducted. Third, the participant started to view the learning materials for twenty minutes. At last, the participant rated the cognitive load questionnaire and received the posttest with a comprehension test of twenty questions. As for data analyses, Mann–Whitney U test was used to analyze the differences of cognitive loads and visual behaviors between high performance group and low performance group. Furthermore, Spearman's correlation analyses were used to analyze the relationships among the participants' cognitive loads, visual behaviors, and reading comprehension.

3. Result

According to the result of Mann-Whitney U test, it was found that there was significant difference between high performance group and low performance group toward extraneous cognitive load from video ($z = -1.999$, $p = .046$). The participants in high performance group had higher extraneous cognitive load from video (Mean = 5.250, SD = 2.062) than the participants in low performance group (Mean = 1.667. SD = 1.155). Moreover, it was found that there was no significant difference between high performance group and low performance group on visual behaviors.

In addition, regarding the result of Spearman's correlations analysis between the participants' cognitive load and visual behaviors, the participants' extraneous cognitive load from picture had significantly negative correlation with total entered count of picture ($r = -.805$, $p < .05$). This finding reveals that the participants who had higher extraneous cognitive load from picture, they had lower counts of looking pictures. It was also found that the participants' extraneous cognitive load from text had significantly positive correlation with average entered duration of pictures' short explanation. This finding reveals that the participants who had higher extraneous cognitive load from text, they had longer average entered durations on the short explanations of pictures.

Finally, based on the result of Spearman's correlations analysis between the participants' visual behaviors and reading comprehension, it was found that there was no significant relationship between the participants' visual behaviors and reading comprehension. However, according to the result of

Spearman's correlations analysis between the participants' cognitive load and reading comprehension, it was found that the participants' extraneous cognitive load from video had significantly positive correlation with the participants' reading comprehension ($r = .972$, $p < .001$). This finding reveals that the participants had higher extraneous cognitive load from video might get higher reading comprehension.

4. Conclusion and discussion

According to the results, the participants in high performance group had higher extraneous cognitive load from video than the participants in low performance group. In addition, the participants had higher extraneous cognitive load from video might get higher reading comprehension. In opposition to general viewpoint, which suggests that extraneous cognitive load should be reduced for better learning (Sweller & Chandler, 1994), it seemed that extraneous cognitive load might actually help the participants to learn better.

In addition, the results showed that the participants who had higher extraneous cognitive load from picture had lower counts of looking pictures. It seems to mean that learners who think that displaying learning contents with pictures is not good for learning might not take pictures as main learning materials; on the contrary, learners who think that displaying learning contents with pictures is good for learning might take pictures as helpful learning materials. Furthermore, the participants who had higher extraneous cognitive load from text tended to have longer average entered durations on the short captions of pictures. It indicates that learners who think that displaying learning contents with texts is not good for learning might need to take more time on realizing what short captions of pictures are talking about; on the contrary, learners who think that displaying learning contents with texts is good for learning might need to take less time on realizing what captions of pictures are talking about. It is interesting to notice that extraneous cognitive load might be the main element that influences the participants' viewing behaviors in AR learning environments.

Furthermore, opposite to the previous study (Yang et al., 2013), it was found that there was no significant relationship between the participants' visual behaviors and reading comprehension in this study. It might be caused by the small number of participants that made the statistical number too hard to be significant. It is suggested that the relationship between visual behaviors and reading comprehension should be paid more attention. More details of the results will be presented and discussed in the conference.

References

Chen, C. M. & Tsai, Y. N. (2012). Interactive augmented reality system for enhancing library instruction in elementary schools. Computers & Education, 59(2), 638-652.

Ho, H. N. J., Tsai, M. J., Wang, C. Y., & Tsai, C. C. (2014). Prior knowledge and online inquiry-based science reading: evidence from eye tracking. International Journal of Science and Mathematics Education, 12(3), 525-554.

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., …Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educational Research Review, 10(0), 90-115.

Sweller, J. & Chandler, P. (1994). Why Some Material Is Difficult to Learn. Cognition and Instruction 12(3), 185-233.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10, 251–296.

Wang, Q., Yang, S., Liu, M., Cao, Z., & Ma, Q. (2014). An eye-tracking study of website complexity from cognitive load perspective. Decision Support Systems, 62(0), 1-10.

Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. Computers & Education, 62, 41-49.

Yang, F. Y., Chang, C. Y., Chien, W. R., Chien, Y. T., & Tseng, Y. H. (2013). Tracking learners' visual attention during a multimedia presentation in a real classroom. Computers & Education, 62, 208-220.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Visual Behavior and Cognitive Load on E-book Vocabulary Learning

An-Hsuan WU*, Po-Fen HSU, Hui-Jou CHIU & Meng-Jung TSAI
Graduate Institute of Digital Learning and Education,
National Taiwan University of Science and Technology, Taiwan
*M10111022@mail.ntust.edu.tw

Abstract: This pilot study examined learning English vocabulary by an E-book from cognitive load aspect, and adopted eye-tracking data as a reference to investigate participants' attention distribution among learning targets, pictures and background graphics. An ASL Mobile Eye-XG eye-tracking system was used to record participants' visual behaviors in this study. A Wilcoxon test was conducted to compare the performance before and after the learning process. Correlation tests were conducted to examine the relationships among posttests, visual behaviors and cognitive loads. Also an ANOVA analysis was conducted to compare the total reading time spent on texts, pictures and background graphics. The preliminary results indicated that participants demonstrated a significant progression through reading the E-book. Besides, it was found that the participants spent more time on reading texts than pictures and background graphics. More details of the results will be discussed in the conference.

Keywords: E-book, cognitive load, eye-tracking, English vocabulary learning, visual behavior.

1. Introduction

Owing to the convenience and motility, the rapid growth of mobile devices generated. Accompany with this state, it springs up varied kinds of products that fit into the mobility mobile devices. E-Book is one of them, and plays an important role in it. Moreover, it pushes the relative studies forward. Mayer's Multimedia Learning Theory became a reference for multimedia learning materials design and examining in many studies. Mayer's (2009) Spatial Contiguity Principle and Temporal Contiguity Principle (2009) claimed that to present integrated texts and graphics on the same page at the same time will be a support during cognitive processing. Furthermore, it may lead to a better learning performance since learners have distributed their cognitive resource to deal with connecting the relation between texts and graphics rather than searching the position of target texts and the corresponding graphics. In addition, the Segmenting Principle, which was also proposed by Mayer, noticed that the key point to carry out this principle is to design a self-paced learning material. On the other side, the cognitive load theory proposed by Sweller, van Merrienboer, and Pass (1998) distinguished cognitive load into intrinsic cognitive load, extraneous cognitive load, and Germance cognitive load (Sweller, et al., 1998; Paas, Tuovinen, Tabbers, and Van Gerven, 2003). Further, van Gog & Scheiter (2010) hold the view that eye tracking could be a direct measurement to help explain the cognitive load during learning process. According to the reviewed paper of Lai et al (2013), studies were prompt in adopting eye tracking from 2009, and also mentioned that many studies suggested using eye-tracking measure as index to combine with interpreting cognitive or meta-cognitive concerned learning. Overall, this study aims to inquire into the relation between eye -tracking measure and cognitive load throughout learning English vocabulary by E-book and also examine the interaction among texts, pictures, and background graphics (hereinafter referred to as background) while reading an English vocabulary E-book.

2. Method

2.1 Participants

Six participants (three males and three females, aged twenty and over) from an university of north Taiwan got involved in this study. Their English ability was under the standardized test – GETP intermediate level (equal to CEFR B1 threshold).

2.2 Instrument

An eighteen English stationary and kitchenware vocabularies E-book is the leaning material of this study. All the participants read the same eighteen vocabularies, yet these vocabulary cards' orders were randomized for individuals. In addition, this study adopted ASL Mobile Eye-XG eye-tracking system, with a sapling rate of 60 Hz, which recorded sixty eye movements information per second. It allowed participants to read the E-book comfortable that close to real reading state without limiting their movement. Meanwhile, the difficulty of each vocabulary that the participants perceived when they were answering the pretest sheet was also asked following by the meaning of each vocabulary. And participants rated a self-reported number lines cognitive load questionnaire after the experiment. The cognitive questionnaire consisting of nine questions, participants rated from one to ten, which indicated that, the vocabulary difficulty from very easy to very difficult, and the degree of assistance and interference about pictures and backgrounds from very low to very high. Finally, the pretest sheet contains all the vocabularies of E-book. Participants were asked to fill in the blank with corresponding words of vocabulary in Chinese and rated the difficulty of each vocabulary at the same time. For the posttest, they were also asked to fill in the blank with corresponding Chinese words, but the order was different from the pretest sheet by randomizing.

2.3 Data Collection and Analyses

Participants were asked to fill in pretest sheet at first. Then, they went through on 5-point eye calibration. After calibration, they began to read the English vocabulary E-book with a time limit of ten minutes. Last, they were asked to fill out a posttest sheet and a cognitive load questionnaire. As for the data analyses, this study conducted Wilcoxon test to compare the pretest and posttest. Besides, correlation analysis was used to examine the relationship between visual behavior and cognitive load, and visual behavior and posttest. In addition, an ANOVA analysis was conducted to compare the total reading time and total entered count of background, pictures, and texts. Besides, recording videos of vocabularies learning behavior were analyzed using Interact 9 software. Statistics of total reading time, entered count and duration of backgrounds, texts, and pictures were coded.

3. Result

3.1 Wilcoxon test of pretest and posttest

As shown in Table 1, through a Wilcoxon test, pretest (M=50.50, SD= 10.436) has significant different from posttest (M=97.71, SD= 2.704) (z=-2.207，p=0.027, p <0.05) , and further comparing the mean score of each, we can learn that the posttest is higher than pretest. That is, it has enhanced the English vocabulary learning performance through the designed E-book in this study.

Table 1: Wilcoxon test of pretest and posttest

	N	Mean difference	SD	z	p
pretest-posttest	6	-47.208	9.592	-2.207	0.027

* p < 0.1, ** p < 0.05, *** p < 0.01

3.2 Correlation between Posttest, Visual behavior, and Cognitive load.

As shown in Table 2, Background Assistance had significantly negative correlations with Posttest (r=-.939，*p*=.005，*p*< .05). This finding reveals that participants who had better performance on learning English vocabulary perceived less assistance from backgrounds. On the contrary, participants who had worse performance perceived more assistance from backgrounds.

Table 2: Correlation between Posttest and Cognitive load

	PiA	PrA	BA	D	PiI	PrI	BI
Posttest	-0.485	-0.127	-0.939**	0.000	-0.211	0.018	-0.949**

* $p < 0.1$, ** $p < 0.05$

Note: PiA=Picture Assistance, PrA= Pronunciations Assistance, BA=Background Assistance, D=Difficulty of vocabularies, PiI= Pictures Interference, PrI= Pronunciation Interference, BI=Background Interference

3.3 Correlation between Visual behavior and Cognitive load

As shown in Table 3, Total Entered Count of Background had significantly negative correlations with Pronunciation Interference (r=-.845，*p*=.034，*p*< .05). This finding reveals that during the reading process, with paying more attention to backgrounds, participants perceived less pronunciation interference, and with paying less attention to backgrounds, participants perceived more pronunciation interference relatively.

Table 3: Correlation between Visual behavior and Cognitive load

	PiA	PrA	BA	D	PiI	PrI	BI
TEC of Background	0.696	-0.441	-0.232	0.696	-0.683	-0.845**	0.098
TET of Background	-0.116	0.383	-0.348	-0.058	-0.293	-0.845**	0.098
TEC of Text	-0.116	0.647	0.145	-0.493	-0.488	-0.507	0.293
TET of Text	-0.319	0.765*	-0.029	-0.493	-0.293	-0.676	0.293
TEC of Picture	-0.162	0.493	-0.338	-0.441	-0.396	-0.429	-0.198
TET of Picture	0.145	0.088	-0.464	-0.029	-0.488	-0.507	-0.293

* $p < 0.1$, ** $p < 0.05$

Note: PiA= Pictures Assistance, PrA= Pronunciations Assistance, BA=Background Assistance, D=Difficulty, PiI= Pictures Interference, PrI= Pronunciations Interference, BI=Background Interference

TEC of Background= Total Entered Count of Background, TET of Background= Total Entered Time of Background, TEC of Text= Total Entered Count of text, TET of Text= Total Entered Time of Text, TEC of Picture= Total Entered Count of Picture, TET of Picture= Total Entered Time of Picture

3.4 Correlation between Visual behavior and posttest

As shown in Table 4, there are no significant correlation among posttest and eye-tracking measures.

Table 4: Correlation between Visual behavior and posttest

	TRT	TRP	TEC of Background	TET of Background	TEC of Text	TET of Text	TEC of Picture	TET of Picture
Posttest	.031	.309	-.062	.123	-.370	-.185	.172	.278

* $p < 0.1$, ** $p < 0.05$

TRT= Total Reading Time, TRP= Total Reading Pages, TEC of Background= Total Entered Count of Background, TET of Background= Total Entered Time of Background, TEC of Text= Total Entered Count of text, TET of Text= Total Entered Time of Text, TEC of Picture= Total Entered Count of Picture, TET of Picture= Total Entered Time of Picture

3.5 ANOVA results

In order to compare the total reading time and total entered count on backgrounds, texts, graphics, a factorial ANOVA was conducted. The results indicate that the all participants spent longer reading texts than looking at backgrounds. On the other hand, the frequency of reading texts and looking at graphics are both more than looking at backgrounds.

Table 5: Total Reading Time and Total Entered Count on background, text and graphic

	(1)Background		(2)Text		(3)Graphic		F	Scheffe
index	Mean	SD	Mean	SD	Mean	SD		
TRT	22.512	9.347	81.870	55.067	66.845	37.566	3.784**	(2)>(1)
TEC	47.000	17.251	120.167	46.154	121.167	54.024	5.734**	(2)>(1),(3)>(1)

* $p < 0.1$, ** $p < 0.05$

Note: TRT= Total Reading Time, TEC= Total Entered Count

4. Conclusion and discussion

This study aims to investigate the correlation among visual behavior, cognitive load, and performance during the process of reading an English vocabulary E-book. In sum, according to the results, it is an effective English vocabulary E-book to help learners to learn English vocabulary by themselves. As for cognitive of background shows negative correlation with posttest, which means that to some extent, it seems that background has interaction with learners' cognitive load of background and performance. The further study can probe the interaction between the background design and the performance and relative reading behavior.

In another aspect, the significant negative correlation among frequency and duration looking at backgrounds with the cognitive load of pronunciation interference and the total reading time of texts shows significant positive correlation with pronunciation assistance that the participants perceived. For this result, we speculate that participants thought that pronunciation function was helpful while they integrated elements, which are more relative to each other (texts and pronunciations). However, while integrated elements that are less relative to each other (backgrounds and pronunciations) they would take it as interference. Last, with the result of total reading time and total entered count on reading texts and looking at pictures and backgrounds, we can know that learners would primly focus on texts and pictures, while background plays a subordinate role.

Further study can be conducted to involve more participants to infer to larger population. On the other hand, the collected eye-tracking data can be further interpreted by deeper analysis. More details of results will be presented and discussed in the conference if the paper is accepted by the conference.

References

Lai, M. L., Tsai, M.-J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., ... & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10, 90-115.

Mayer, R. E. (2009). Multimedia learning (2ed ed.). New York: Cambridge University Press. *Journal of Computer Assisted Learning*, 20 (4), 257-268.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63-72.

Sweller, J., van Merrienboer, J. J. G., & Paas, F.(1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296.

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20(2), 95-99.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Incorporating augmented reality into learning practical skills for medical surgery

Ying-Shao Hsu[a], Yuan-Hsiang Lin[b], Beender Yang[c], Shih-Fan[b] Yang, Ya-Yen Chan[b], Zi-Hao Lin[b], and Yi-Sheng Chan[b]

[a]*Graduate Institute of Science Education, National Taiwan Normal University, Taiwan*
[b]*Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taiwan*
[c]*Saturn Imaging Inc., Taiwan*
*yshsu@ntnu.edu.tw

Abstract: In this paper, we demonstrate how to incorporate augmented reality (AR) into the learning of practical skills for medical surgery. We embedded AR in authentic inquiry activities so that students could experience when and how to carry out a certain medical surgical procedure. Two learning modules related to medical surgery were developed, "laparoscopic surgery" and "cardiac catheterization". Thirty-two senior high school students participated in this study and their perceptions of AR were examined. A survey of student perceptions included the three constructs of authenticity, engagement and motivation. The results showed that the students had positive perceptions (overall mean = 4.1) of AR after completing the two modules. However, AR authenticity was the concept perceived as having the lowest ranking (mean = 3.7). In contrast, both the motivation triggered by AR and engagement reached 4.3. This article provides a possible solution for the alignment among instructional approaches (authentic inquiry), technology design (AR) and learning experience.

Keywords: Augmented reality, Authentic inquiry, Practical skills

1. Introduction

Augmented reality offers a new form of interactivity between the physical and virtual worlds, and enhances users' perceptions of the real world (Kesima & Ozarslan, 2012). According to Wu et al.'s review (2013), AR allows students to develop important practices, and has become one of the key emerging technologies in education. Although AR may present opportunities for teaching and learning, do students perceive an adequate level of realism when they are immersed in such a learning environment? How can researchers and educators work together to advance learning by aligning instructional approaches, technology design and learning experience? This study aims to propose a possibility to embed AR in authentic inquiry activities for contextualizing student exploration of medical surgery.

2. The AR learning modules

Authentic inquiry refers to performing the complex process which scientists actually carry out (Chinn & Malhotra, 2001). Instead of the simple inquiry tasks seen in most science textbooks, authentic inquiry tasks allow students to interact with computer-simulated experiments or equipment so as to develop their inquiry skills. AR is a promising way to combine authentic contexts and simulated experiments for student exploration. Therefore, we incorporated AR with authentic inquiry to engage students in two surgical procedures, laparoscopic surgery and cardiac catheterization. The authentic inquiry activities were designed to facilitate the students' experience of the diagnosis of symptoms and to operate the laparoscopic surgery and cardiac catheterization simulators (Figures 1 and 2).

Originally, the simulated experiments in these two modules were designed for medical majors. For an outreach purpose, we invited senior high school teachers to design authentic inquiry activities for introducing metical practical skills to senior high school and helping them explore the possibility of their future career. A total of 32 senior high school students were grouped into eight groups, each of which was provided with an Android tablet computer to interact with the simulators using the scanning function on the tablet. The tablet computer delivered an authentic context of a patient's symptoms, and the students were required to diagnose the possible disease using the data such as X-ray images and electrocardiograms on the tablet. Then, the students used the simulator to help the patient recover from the disease. We applied the role-play technique in these activities whereby the students were told that they were surgery interns who were required to learn practical surgical skills. The students spent 2.5 hours working in a group to complete these two learning modules, laparoscopic surgery and cardiac catheterization.

A survey was conducted right after the modules to elicit the students' perceptions of AR including the three constructs of realism, engagement, and motivation. Each construct consisted of six 5-point Likert-scale items modified from the study of Change, Lee, Wang and Chen (2010).

Figure 1. Laparoscopic surgery simulator.

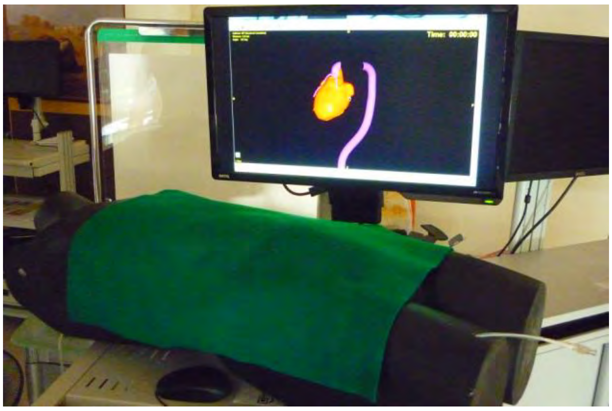


Figure 2. Cardiac catheterization simulator.

2.1 The laparoscopic surgery module

The students were told that they were to role-play surgery interns. Each group of students worked with a tablet and the simulator. First, a video clip on the tablet showed a patient describing his symptoms, and the students were required to diagnose the possible disease after checking the x-ray images. Second, some surgery options were shown to the students from which they were required to select the most appropriate procedure for the disease. Third, the students used the scanning function to activate the simulator, and operated the laparoscopic surgery simulator which displays 3D images for promoting practical skills (Figure 1). Finally, the students needed to reflect on the pros and cons of laparoscopic surgery and to offer their opinions regarding improvements to such surgery.

2.2 The cardiac catheterization module

The cardiac catheterization module was developed following a similar procedure to that of the laparoscopic surgery module. The students role-played surgery interns to diagnose a patient's disease from the electrocardiogram on the tablet. Then, the students used the scanning function to activate the simulator and carried out angioplasty by operating a catheter in the simulator. A 3D heart image was displayed synchronously on the computer to indicate the location of the catheter when the students moved it in the simulator. For some critical points, the computer provided a doctor's advice to help the students overcome their difficulties operating the catheter. At the end, the students needed to answer some questions related to cardiac catheterization.

3. Findings and Conclusions

As Table 1 shows, the overall average score of student perceptions of AR was 4.1. The average scores of student perceptions of realism, engagement and motivation are separately 3.7, 4.3, and 4.3. We found evidence that embedding AR in authentic inquiry promotes students' engagement and motivation in developing the practical skills for medical surgery. Although we used the same equipment as medical surgery, the students still perceived that the AR developed in this study was not as realistic as it might have been. More improvements can be made to increase the realism of AR such as connecting surgical equipment to a human model which can react to students' operation synchronously through sensors. Future studies need to investigate other important aspects such as examining the benefits of AR for learning practical skills in authentic inquiry activities.

Table 1: Summary of student perceptions of AR.

Constructs	Mean	S.D.
Realism	3.7	1.1
Engagement	4.3	1
Motivation	4.3	0.7
Total	4.1	1

Acknowledgement

This research is partially supported by the"Aim for the Top University Project" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the "International Research-Intensive Center of Excellence Program" of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301.

References

Chang, C. W., Lee, J. H., Wang, C. Y., & Chen, G. D. (2010). Improving the authentic learning experience by integrating robots into the mixed-reality environment. *Computer & Education, 55*(4), 1572-1578.

Chinn, C. A., & Malhotra, B. (2001). Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Science Education, 86*(2), 175-218.

Kesima, M., & Ozarslan, Y. (2012). Augmented reality in education: current technologies and the potential for education. *Procedia - Social and Behavioral Sciences 47*, 297 – 302.

Wu, H.-K., Lee, S W.-Y., Chang. H.-Y., and Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41-49.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Path Analyses of How Students Develop Conceptual Knowledge and Inquiry Skills in a Simulation-Based Inquiry Environment

Hsin-Yi CHANG[a*], Ying-Shao HSU[b], Hsin-Kai WU[b], & Chih-Ming CHEN[c]
[a]*Graduate Institute of Science Education and Environmental Education, National Kaohsiung Normal University, Taiwan*
[b] *Graduate Institute of Science Education, National Taiwan Normal University, Taiwan*
[c] *Graduate Institute of Library, Information and Archival Studies, National Cheng-Chi University, Taiwan*
*hsinyichang@nknucc.nknu.edu.tw

Abstract: We implemented a simulation-based learning environment in a summer camp program in which 45 ninth-grade students conducted virtual experiments using computer-based simulations to learn the concepts of buoyancy. We collected data including all students' responses to the pretests, assessments embedded during class in the learning environment, and posttests. All the tests measured the students' conceptual knowledge of buoyancy and inquiry skills. We conducted path analyses to investigate how the students developed their conceptual knowledge and inquiry skills during and after learning within the learning environment. A common significant path is identified among students' prior knowledge and skills and students' developed knowledge and skills.

Keywords: Computer-based simulation, virtual experiment, inquiry, science, path analysis

1. Introduction

Computer-based simulations enable human-computer interactions that can allow students to conduct virtual experiments using simulations. Compared to hands-on experiments, virtual experiments are equal to or more effective for enhancing students' understanding of scientific concepts due to their technological efficiency and their ability to provide ideal experiment conditions (Klahr, Triona and Williams, 2007; Zacharia, Olympiou and Papaevripidou, 2008). However, it is unclear how students develop their conceptual knowledge in a simulation-based virtual experiment environment as they conduct inquiries. Some research indicates that the development of conceptual knowledge and inquiry skills are often interwoven (Bao, Gotwals, Songer and Mislevy, 2006). Despite this, it is possible to measure conceptual knowledge and inquiry skills separately (e.g., P.-H. Wu, H.-K. Wu and Hsu, 2013), and to discern how conceptual knowledge and inquiry skills are interwoven at multiple time-points such as before, during, and after class. Specifically, we are interested in, in a simulation-based inquiry learning environment, how students' prior knowledge and prior inquiry skills affect their knowledge and inquiry performance during class, and in turn after class in the posttests.

2. The Simulation-Based Inquiry Environment

The inquiry learning environment incorporates a newly developed simulation for students to conduct virtual experiments to learn factors related to an object sinking or floating in a given fluid. The students can change the values of four variables in the simulation to conduct their experiments (Figure 1): (1) the material of the object (duck): brick, wood, ice, styrofoam, or steel; (2) the size of the object (duck): large, medium or small; (3) the composition of the duck: solid or hollow; and (4) the type of fluid: water, saline water, gasoline, or mercury. In addition, on the upper right hand side of Figure 1, four virtual probes are provided so that the students can use them to measure the volume, mass, density or buoyant force of the object, and the density of the liquid. A small pop-up window appears on the left side to indicate the volume of the fluid displaced after the duck is placed in the fluid. The students were asked to synthesize from their experiments, and to reason the variables directly related to the phenomenon of sinking and floating. In addition, we embedded a function in the simulation for the students to create a worksheet of their experiments recording the properties of the ducks and fluids they experimented with (Figure 2). Using the worksheets, the students were guided to reason that for a floating object, the buoyant force is equal to the weight of the object, and that for a sunken object (and a floating object), the buoyant force of the object in a fluid equals the density of the fluid times the volume of the object immersed in the fluid. Moreover, they were guided to draw their visualizations of the buoyant forces acting upon floating and sunken objects.

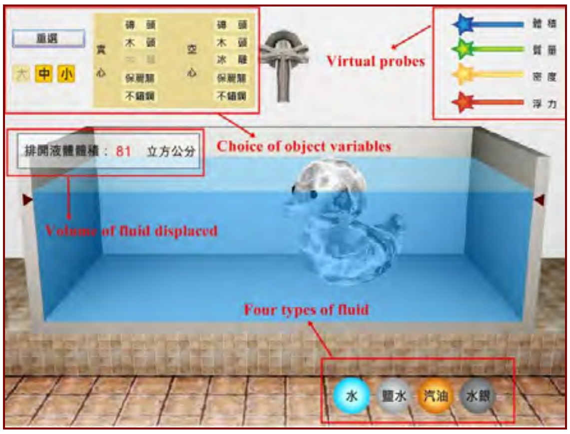


Figure 1. A screenshot of the *bath duck* simulation that allows students to conduct virtual experiments

實驗次數	ID	步驟	物體	實心/空心	是否為沉體	物體所受浮力(gw)	物體在液面下體積(cm^3)	液體與其密度(g/cm^3) 油:0.7 水:1 鹽水:1.1 水銀:13.6	物體重量(gw)
1	fl10230325	2.3	小木鴨	實心	浮體	15	15	水:1	15
2	fl10230325	2.3	中冰鴨	實心	浮體	54	54	水:1	54
3	fl10230325	2.3	大磚鴨	空心	沉體	135	90	水:1	90
4	fl10230325	2.3	小保麗龍鴨	空心	浮體	2.1	2.1	水:1	2.1
5	fl10230325	2.3	大鋼鴨	實心	沉體	270	90	鹽水:1.1	99
6	fl10230325	2.3	中木鴨	空心	浮體	18	25.7	油:0.7	18
7	fl10230325	2.3	大冰鴨	實心	浮體	81	6	水銀:13.6	81
8	fl10230325	2.3	中磚鴨	空心	浮體	90	6.6	水銀:13.6	90
9	fl10230325	2.3	大鋼鴨	空心	浮體	45	45	水:1	45
10	fl10230325	2.3	小木鴨	實心	浮體	15	15	水:1	15
11	fl10230325	2.3	大冰鴨	空心	浮體	45	45	水:1	45
12	fl10230325	2.3	中鋼鴨	實心	沉體	180	60	水:1	60
13	fl10230325	2.3	大保麗龍鴨	空心	浮體	6.3	6.3	水:1	6.3
14	fl10230325	2.3	小磚鴨	空心	沉體	45	30	水:1	30
15	fl10230325	2.3	中磚鴨	實心	沉體	120	60	水:1	60
16	fl10230325	2.3	中保麗龍鴨	實心	浮體	12	12	水:1	12
17	fl10230325	2.3	小冰鴨	空心	浮體	15	15	水:1	15
18	fl10230325	2.3	大木鴨	實心	浮體	45	40.9	鹽水:1.1	45
19	fl10230325	2.3	小磚鴨	空心	沉體	45	30	鹽水:1.1	33
20	fl10230325	2.3	中鋼鴨	實心	沉體	180	60	油:0.7	42
21	fl10230325	2.3	中冰鴨	空心	浮體	30	2.2	水銀:13.6	30
22	fl10230325	2.3	小木鴨	實心	浮體	15	1.1	水銀:13.6	15
23	fl10230325	2.3	大保麗龍鴨	實心	浮體	18	1.3	水銀:13.6	18
24	fl10230325	2.3	小木鴨	空心	浮體	9	0.700005	水銀:13.6	9
25	fl10230325	2.3	中磚鴨	空心	浮體	90	6.6	水銀:13.6	90

Figure 2. A computer-generated worksheet recording the student's experiments

3. Methods

The study involved 45 ninth-grade students who volunteered to participate in a summer science camp program at a public high school in North Taiwan. These students demonstrated high interest and motivation in learning science. They had not learned buoyancy prior to this study. The students spent 4 hours to complete their learning in the environment with the guidance of a science teacher. Each individual student took a pretest before and a posttest after. The pretest and posttest were identical and included two parts. The first contained 8 items to measure the students' conceptual knowledge of buoyancy. The second part contained 15 items to measure the students' inquiry skills including planning experiments, identifying variables, conducting reasoning, using evidence and evaluating explanations. The items went through several rounds of revision by science educators to ensure their content and construct validity. In addition to the pretest and posttest, another 8 conceptual items and 11 inquiry items were embedded in the learning environment to measure the students' conceptual knowledge and inquiry performance demonstrated during the students' learning in the environment.

We developed detailed scoring rubrics to score the students' responses. In general, for the conceptual items, one point was given for an appropriate response and zero for an inappropriate one. For the inquiry items, two points were given for a high quality response, one point for a moderate quality response, and zero for a low quality one. Two independent raters coded all the tests, and the inter-coder agreement reached 95%. Inconsistent codes were discussed and resolved. Each individual student had six scores, namely, pretest knowledge, pretest inquiry, embedded assessment knowledge, embedded assessment inquiry, posttest knowledge, and posttest inquiry. We employed multiple regressions to conduct path analyses to test the relationships among these variables (Foster, Barkus and Yavorsky, 2006).

4. Results

We summarized the path analysis results in Figure 3 for the posttest knowledge scores and in Figure 4 for the posttest inquiry scores. Figure 3 shows a significant path from the students' prior inquiry skills to their knowledge demonstrated during class, then to the inquiry skills demonstrated during class, and finally to their knowledge demonstrated in the posttests. The path shows direct significant effects from one variable to another among the four variables. The students' prior knowledge did not have any significant effect. The students' prior inquiry skills had an indirect effect on their inquiry skills during class, and the knowledge learned during class also had an indirect effect on the knowledge demonstrated after class. Figure 4 shows a similar path pattern for the inquiry skills demonstrated in the posttests. The only difference is that the students' prior inquiry skills had a direct effect on their inquiry performance on the posttest.

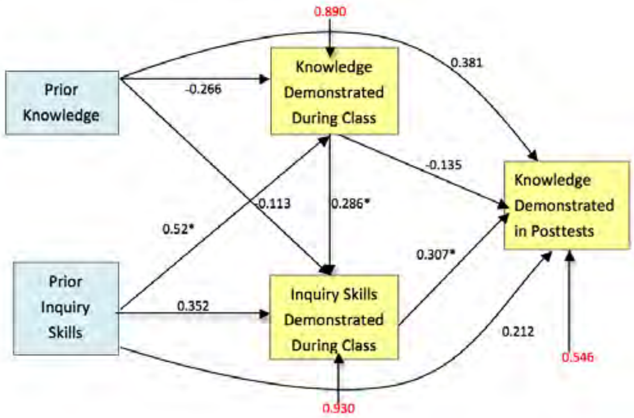


Figure 3. A path diagram for conceptual knowledge demonstrated in the posttests

Table 1: Total significant effects for the path shown in Figure 3

	Prior Knowledge	Prior Inquiry	During Class Knowledge	During Class Inquiry
Posttest Knowledge	0	0.045	0.087	0.307

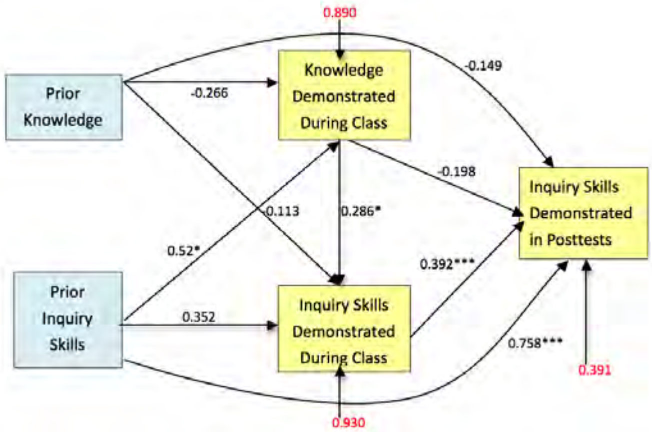


Figure 4. A path diagram for inquiry skills demonstrated in the posttests

Table 2: Total significant effects for the path shown in Figure 4

	Prior Knowledge	Prior Inquiry	During Class Knowledge	During Class Inquiry
Posttest Inquiry	0	0.816	0.112	0.392

We summarized the total effects on the posttest knowledge performance in Table 1, and on the posttest inquiry performance in Table 2. Table 1 indicates that the students' inquiry during class had the strongest effect on their posttest knowledge performance, with their knowledge during class having an effect in between, and their prior inquiry skills having the least effect. For Table 2, the inquiry skills that the students possessed before class had the strongest effect on their inquiry performance on the posttest, with the middle effect of their inquiry developed during class, and least effect on their knowledge gained during class.

The red numbers in Figures 3 and 4 are the calculated error values for the dependent variables. These error values indicate variances not explained by the model. For both the knowledge and inquiry skills demonstrated during class, the error values are high, indicating that there are other variables not included in the model that might be better than the students' prior knowledge and prior inquiry skills at accounting for their knowledge and inquiry demonstrated during class. We conjecture that the other variables include guidance and scaffolding from the teacher and learning environment. Qualitative analyses can verify this conjecture, but this is beyond the scope of this paper.

5. Concluding Remarks

Compared to SEM, multiple regression models for path analyses are suitable for studies with smaller sample sizes to explore possible relationships for further study. Through such techniques we found that compared to their prior knowledge, students' inquiry skills are more important in terms of predicting posttest performance on both the conceptual and inquiry measures. Nevertheless, knowledge, especially the knowledge learned during class, still has effects on students conducting adequate inquiries. From a broader viewpoint, the development of knowledge and inquiry skills may be interwoven (Bao et al., 2006). When examining on a finer scale, the paths in the two diagrams follow a "Z" shape, starting from personal inquiry skills. Such a result has implications for the call for learning environments that emphasize the progressive development of inquiry skills. However, the results reported here are specific to the context of this study, including the simulation-based inquiry environment in which semi-structured curricular scaffolds and teacher guidance are provided to support student inquiry with the simulations (for details of the curriculum design, see Hsu, Chang, Fang and Wu, in press). Future studies can explore whether there are different paths when engaging students in other inquiry learning environments with less or more scaffolding or guidance.

Acknowledgements

This material is based upon work supported by the Ministry of Science and Technology, Taiwan, under grants MOST101-2511-S-003-055-MY3, and MOST102-2628-S-017-001-MY3. We would like to thank all the people who were involved in the collection and coding of the data.

References

Bao, H., Gotwals, A. W., Songer, N. B., & Mislevy, R. J. (2006). Using structured item response theory models to analyze content and inquiry reasoning skills in BioKIDS. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education*. Maple Grove, MN: Journal of Applied Measurement Press.

Foster, J. J., Barkus, E., & Yavorsky, C. (2006). *Understanding and using advanced statistics*. London, UK: SAGE Publications Ltd.

Hsu, Y.-S., Chang, H.-Y., Fang, S.-C., & Wu, H.-K. (in press). Developing technology-infused inquiry learning environment to promote science learning in Taiwan. In M. S. Khine (Ed.), *Science education in East Asia: Pedagogical innovations and best practices*.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching, 44*, 183-203.

Wu, P.-H., Wu, H.-K., & Hsu, Y.-S. (2014). Establishing the criterion-related, construct, and content validities of a simulation-based assessment of inquiry abilities. *International Journal of Science Education, 36*(10), 1630-1650.

Zacharia, Z. C., Olympiou, G., & Papaevripidou, M. (2008). Effects of experimenting with physical and virtual manipulatives on students' conceptual understanding in heat and temperature. *Journal of Research in Science Teaching, 45*, 1021-1035.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Understanding middle and high school students' views of model evaluation and model change

Silvia Wen-Yu Lee[a*], Hsin-Kai Wu[b] & Hsin-Yi Chang[c]
[a]*National Changhua University of Education*
[b]*National Taiwan Normal University*
[c]*National Kaohsiung Normal University*
*silviawyl@cc.ncue.edu.tw

Abstract: This study aimed to understand students' views of the nature of model evaluation and the nature of change of models in different context. A total of 102 eighth graders and 87 eleventh graders were surveyed. Two cases, the SARS and dinosaur extinction, were presented to prompt students' ideas about different models proposed by scientists. The statistical results showed different context of the model influenced how the students viewed model evaluation and model change. The students' answers also showed significantly differences between the high school level and the middle school level for their views of model change. The common reasons behind students' choice were related to students' understanding of the changeable nature of model and the science process. The students who chose that "one model is better than another" tended to justify their response by their understanding of the content. Interestingly, some students' views of the dinosaur extinction model were guided by their beliefs that information about the dinosaurs is unfathomable. The findings suggest that researchers should be aware that the models chosen for teaching and for assessment can interact with other factors, such as their familiarity of the content, their level of education and understanding of the nature of science. The results from written responses were further used to develop a multiple-choice survey and validated in the follow-up study.

Keywords: views of modeling, computer-based survey, model evaluation, change of models

1. Introduction

Researchers found that students held little understanding of the concepts of models (Carey & Smith, 1993; Grosslight, Unger, Jay, & Smith, 1991; Saari & Viiri, 2003); even with formal training in modeling, students still encountered difficulties in fully understanding the nature of models (Harrison & Treagust, 2000; Schwarz & White, 2005). Researchers stated the needs of tapping into the interaction between epistemic beliefs and the contexts in which the epistemic beliefs being measured and being developed (Franco, Muis, Kendeou, Ranellucci, & Sampasivam, 2012). Earlier studies used interviews or paper-and-pencil questionnaires to understand students' general beliefs of models and modeling. However, these studies provided students with little referential information to models. Thus the purpose of the study is to gain insight into the potential interaction of the views of models and the given context and the students' justification to their views. This study focused on two of the major aspects of views of model, that is, the nature of model evaluation and the change of models.

In sum, we posted the following research questions: 1) What are the students' view of model evaluation and model change in the two different context? 2) To which extent do the high school students' views differ from the middle school students' views in the given context? 3) How do the students justify their views of models in relation to the context?

2. Methodology

In this study, we surveyed 102 eighth graders and 87 eleventh graders. Two cases were presented to prompt students' ideas about different models proposed by scientists. The first case involved two routes of infection for SARS (Severe Acute Respiratory Syndrome) virus, and the second case included different explanations of the dinosaur extinction. Students were asked to answer the multiple-choice questions and then provide a written response to justify their answers. We also asked students to rate that to which extent they were familiar with the two content of the two cases.

We conducted a series of Chi-square analyses including independent tests and goodness to fit tests for understanding the differences within the same educational level or between educational levels. We also used McNemar tests and McNemar-Bowker tests (Elliott & Woodward, 2006) for examining the consistency of students' answers across different items. Opened coding methods were first applied to students' written justification to their choice of answers. Then a list of coding schemes were tested on the data and modified until the coding schemes were saturated.

3. Findings

3.1 Model Evaluation

In terms of model evaluation, after reading the two cases, students had to make a choice among three options: (1) one model is better than another; (2) cannot know which model is better unless new evidence supports one of them; (3) both explanations can be valuable; there is no need to decide which model is better. Results show that nearly one fifth of the middle and high school students believed that one model is better than another. However, in the SARS case, nearly 70% of the middle school students and 57.47% of the high school students thought that both explanations can be valuable. The majority of high school and middle school students chose this answer for the SARS case (middle school $\chi2(2) = 62.00$, $p < .001$; high school $\chi2(2) = 23.24$, $p < .001$). For the dinosaur extinction question, the most chosen answer for high school students was "cannot know which model is better unless new evidence supports one of them (45.98%)" and "both explanations can be valuable; there is no need to decide which model is better (44.55%)" for middle school students. However, the results of chi-square analysis showed no statistical significant relationships between students' educational levels and their views of model evaluation.

Further analysis with McNemar-Bowker tests also confirmed that the context of the item influenced students' views of whether a model is better than another ($p < .001$ for middle school students; $p = .004$ for high school students). Only 50.4% of the middle school students chosen the same answers between the two questions; even less percentage (40.2%) of the high school students had consistent answers between the two contexts. A high percentage of students who answered "both explanations can be valuable" for the SARS question shifted their views to "cannot know which model is better" when it came to the dinosaur extinction question. An interesting finding was revealed in students' self-reported levels of understanding of the two topics. For middle school students, students reported similar level of understanding; however, for high school students, they reported significant higher level of understanding of the SARS topic than the dinosaur topic (SARS mean = 2.64; dinosaur mean = 2.38; $p<.001$). This could be a possible explanation of why the high school students seemed to shift their answers between the two contexts and believed that they could not know which of the models of dinosaur appeared to be better.

Overall, the high school students were more likely than middle school students to provide meaningful justification to their choices in both cases. Students who chose "one model is better than another" mainly focused their explanations on the science content of the cases (e.g., "if SARS were air-borne, then everyone should be infected by now"; "I think climate changes sound like the cause [for dinosaur extinction]"). Their justification to the answer of "no need to decide which model is better" focused on the changeable nature of models.

Examples of students' responses included "models can change when the new one is better," and "there could be more than one explanation or possibility." Compared with students' responses to the SARS case, the percentages of choosing the second option (cannot know which model is better unless new evidence supports one of the two) were higher in the dinosaur extinction case. Interestingly, one special set of responses to the dinosaur questions was unforeseen in the responses to the SARS case. Because "the dinosaur extinction happened long time ago; no one really knows" and "dinosaurs are dead", many students believed that there is no way to know which model is better. We found that 43% of high school students who chose this option because that "dinosaur do not exist anymore" and only 18% of students who chose this answer really thought about the importance of finding new evidence. In the SARS case, students who chose the second option tended to justify their answers based on understanding of the scientific process (e.g., "if an error is found, scientists should correct it immediately"), science content (e.g., "I think it is air-borne"), or changeable nature of models (e.g., "if necessary, a model should change to respond to a new question").

3.2 Change of Models

Based on the same SARS and dinosaur extinction context, we also asked students whether a model changes often. In the same context of the SARS and dinosaur extinction cases, students were asked whether models need to change often. Students could choose among the three options: (1) need to change often; (2) no need to change often; (3) it depends. For the SARS case, the most chosen answer was "it depends (44.55% for middle school; 55.17% for high school)". A large percentage of students also chose "need to change often". However, only 3.45% of the high school students chose "no need to change often" while nearly 15% of the middle school students preferred this option. There was significant relationship between students' views of change of models and the two educational levels ($\chi2 = 7.42$, $p = .024$) regarding the SARS question.

When answering the same question in the context of dinosaur extinction, nearly 52% of the high school students and about 41% of the middle school students believed that it depends. Less than 30% of the students chose either "need to change often" or "no need to change often" for both groups. The majority of students were unsure about whether models about dinosaur extinction need to change often. The results of McNemar tests showed that the context of the two questions had an impact on students' views of whether models need to change, but only for high school students ($p = .002$). About 54% of the high school students held the same views of models between SARS and dinosaur questions. The percentage of the high school students who chose "no need to change" increased in the dinosaur case.

For the SARS question, students who chose "it depends" or "need to change often" tended to provide reasons related the changeable nature of models (e.g., "there must be more than one pathway to spread the viruses" or "virus is always mutating") and then scientific process (e.g., "it change when new evidence is found"). Students also gave similar explanations for the dinosaur question, but more students provided reasons related to the scientific process than changeable nature of model. A large percentage of the students to chose "no need to change often" did not provide a meaninful explanation. As for the students who provided justfication, some of them stated "no need to change unless there is new evidence (coded as "science process")" or "no need to the change the current model because multiple models can co-exist (coded as " changeable nature of model").

4. Discussion

Students' views of model evaluation can be interpreted from a personal epistemological point of view (Justi & Gilbert, 2002; National Research Council, 2007). Based on different levels of personal epistemology (Yang & Tsai, 2010), the answer of "one model is better than another" is close to an absolutist perspective; the answer of "both explanations can be valuable; there is no need to decide which model is better" is closer to a multiplist perspective; and the answer of "cannot know which model is better unless new evidence supports one of them" is similar to an evaluatist perspective. One interesting observation is that according to our data, students who chose an absolutist view of model seemed to focus on factual reasons. This can be interpreted as a way to support their judgment by evidence. Students who took a multiplist or evaluatist perspective, tended to think in terms of the nature of science (e.g., changeable nature of model).

In summary, we found that the different context of the model influenced how the students answered the questions of model evaluation and model change. The students' answers also showed significantly differences between the high school level and the middle school level for their views of model change in the SARS case. The common reasons behind students' choice were related to students' understanding of the changeable nature of model or the science process. For the students who chose that one model is better than another tended to justify their response by their understanding of the content. Interestingly, some students' responses to the dinosaur extinction case were guided by the beliefs that further information about the dinosaurs is unfathomable. These findings confirmed that students' development of personal epistemology can be dynamic and somehow unstable influenced by factors such as the context, affection, or cognitive ability (Bendixen & Rule, 2004). The findings suggest that researchers should be aware that the models chosen for teaching and for assessment can interact with other factors, such as their familiarity of the content, their level of education and understanding of the nature of science.

5. References

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, 39(1), 69-80.

Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist,* 28, 235-251.

Elliott, A. C., & Woodward, W. A. (2006). Statistical analysis quick reference guidebook: With SPSS examples. Thousand Oaks, CA: Sage Publications, Inc.

Franco, G. M., Muis, K. R., Kendeou, P., Ranellucci, J., & Sampasivam, L. (2012). Examing the influences of epistemic beliefs and knowledge representations on cognitive processing and conceptual change when learning physics. *Learning and Instruction,* 22, 62-77.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching,* 28(9), 799-822.

Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education,* 22(9), 1011-1026.

Justi, R. S., & Gilbert, J. K. (2002). Science teachers' knowledge about and attitude towards the use of models and modelling in learning science. *International Journal of Science Education*, 24(12), 1273-1292.

National Research Council. (2007). Understanding how scientific knowledge is contructed. In R. A. Duschl, H. A. Schweingruber & A. W. Shouse (Eds.), Taking Science to School: Learning and Teaching Science in Grades K-8.

Saari, H., & Viiri, F. (2003). A research-based teaching sequence for teaching the concept of modelling to seventh-grade students. *International Journal of Science Education,* 25(11), 1333-1352.

Schwarz, C. V., & White, B. Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, 23(2), 165-205.

Yang, F.-Y., & Tsai, C.-C. (2010). Reasoning about science-related uncertain issues and epistemological perspective among children. *Instructional Science,* 38, 325-354.

Bio Detective: Student science learning, immersion experience, and problem-solving patterns

Mei-En HSU[a] & Meng-Tzu CHENG[a*]
[a]*Department of Biology, National Changhua University of Education, Taiwan*
*mtcheng@cc.ncue.edu.tw

Abstract: Many studies have shown that there is a positive impact of serious educational games (SEGs) on student learning. Because of the game graphics, animations, sounds, and narratives, the learners can immerse in the virtual surroundings. Once they immersed, they might try their best to solve the in-game tasks. Therefore, the purpose of this study was to develop a SEG–*Bio Detective* and to evaluate its impact on student science learning outcomes. Moreover, we further investigated the relationships between students' game immersion experience and their science learning and problem solving. The obtained results showed that student science learning can be significantly improved through *Bio Detective* play, but there were no significant correlations between game immersion and learning outcomes. Comparing the problem-solving patterns and problem-solving abilities between students with high- and low immersion experience, we found that students with high immersion experience had a more complete problem-solving pattern and a better problem-solving performance than students with low game immersion experience.

Keywords: serious educational games, problem solving, immersion

1. Introduction

With the rapid improvement of technology, video game play has become popular entertainment. Most of today's children have the experience of playing video games. Due to the features of video games, such as excellent interaction and attractive entertainment, which engage players so much, researchers and educators commenced considering the probabilities of using video games in education since 2002 (Griffiths, 2002; Squire, 2008). Up to now, serious educational games (SEGs) have gradually become a term indicating any video games which are used for teaching and learning purposes in k-20 educational settings (Annetta, 2008). Mouaheb, Fahli, Moussetad and Eljamali (2012) suggested that playing SEGs is actually the process of learning and many studies have shown that SEGs did have a positive impact on various aspects of learning, such as learning achievement, cognitive development, learning motivation, and learning interests (Hwang, Yang, and Wang, 2013; Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012). SEGs have been regarded as a potential vehicle which can enhances student learning because SEGs not only contain the entertaining features of games, but also combine those game features with learning content, which make learners engage in the game and further learn in the virtual surrounding (Van Eck, 2006; Cheng, Su, Huang, & Chen, 2013).It is generally argued that SEGs can provide learners with a vivid world which bridges virtual reality into reality in where players can experience so-called situated learning.

Brown and Cairns (2004) and Cheng, She, and Annetta (2014) contended that games can provide players with game immersion experience. The animations, sounds, and sophisticated graphics of games offer players an immersive environment in where they might ignore changes and forget everything at their surroundings; moreover, they might feel like they are the leading role of the game, and therefore put their whole concentration, thoughts, and even emotions into the game. While players experiencing game immersion, their intrinsic motivation and intrinsic rewards may increase, which subsequently make them actively play the game again and again. If the experience of immersion happened in the learning situation that allows learners to involve in the surrounding without external interference, then they will be willing to invest time and effort to learn. In terms of SEGs, the learners will try to learn constantly because of the characteristics of SEGs of appropriately combining game features with learning content. Students will be getting familiar with the embedded concepts because of continuing practices while immersing in the game and better learning outcomes will subsequently reached (Liu, Cheng, & Huang, 2011).

On the other hand, the players have to learn the rules and plan and figure out ways to solve tasks and problems in the virtual environment in order to succeed in the game. Garris, Ahlers, & Driskell (2002) proposed the concept of input-process-outcome game model, claiming that SEGs offer an immersive environment which allows players to experience the game cycle of judgment–behavior–feedback and finally debriefing learning outcomes. The game cycle is actually the mechanism that provides the players with the complete problem-solving pathway. Therefore, it is argued that the problem-solving ability of players can be improved if they enjoy and immerse in the game (The Economist, 2005).

Although theoretical claims have been proposed by many researchers to explain why SEGs can improve student learning, there is still a lack of evidence that empirically investigates the impact of SEGs on student learning outcomes and how game immersion experience influences student learning outcomes and problem-solving patterns. Therefore, this study has a dual purpose. One is to develop an SEG, *Bio Detective*, to improve student science learning, and the other is to figure out the relationships between student science learning, game immersion, and problem solving through SEG playing.

2. Materials and methods

2.1 Serious educational game design

2.1.1 Bio Detective

In this study, the developed SEG, entitled *Bio Detective*, is an adventure/role-playing game. The storyline of *Bio Detective* is that there is a murder happened and the player should play as a detective's assistant in the game. What he/she has to do is to collect clues, to conduct experiments to do blood-type and glucose tests, and to interpret data to find out whom the murderer actually is. Therefore, the scientific concepts embedded include the inheritance of blood-types (the inheritance of multiple alleles), blood-type test (antigen-antibody reaction) and glucose test.

Figure 1 is the laboratory scene in the *Bio Detective*. After collecting all the clues, the players are able to enter the lab to conduct experiments. They can drag any equipment they need to the table. If they drag the right equipment to the table, the button "Beginning of the experiment" will show up.

Figure 1. Laboratory Scene.

Figure 2 is the library scene. When the learners encounter problems and they don't know how to solve them, they can go to the library to find information which might be useful in helping them solve the problems.

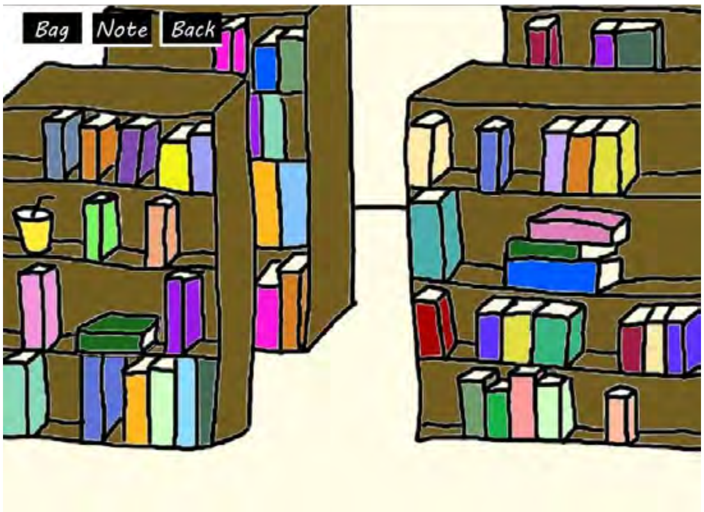


Figure 2. Library Scene.

Bio Detective is created in an attempt to provide learners with a virtual environment as well as a vivid context in where they can actively learn the relative biology knowledge which is required for completing the game. It is hoped that students' willingness and interests of learning can be improved because of their immersion in an atmosphere of suspense and excitement.

2.1.2 Learning objectives

According to the scientific concepts and science process skills embedded in the game, several learning objectives were addressed as below:

- Cognitive domain (the learner is able to)
  1. Comprehend principles and procedure of the glucose test
  2. Understand the principles and procedure of the blood-type test
  3. Interpret the experimental results and perceive the meaning
  4. Reinforce the scientific concepts of genes and blood types
- Psychomotor domain
  1. Promoting the learner's ability to conduct the experiment independently
  2. Improving the learner's ability to solve problems
- Affective domain
  1. Fostering the learner's biological literacy
  2. Inspiring the learner's enthusiasm and interest in exploring the field of biology

2.2 Participants

The participants were two classes of seventh grade students. One class consisted of 22 students and the other consisted of 30 students, resulting in a total of 52 students took part in the study.

2.3 Instrumentation

The instrumentation in this study includes the learning achievement assessment, the game immersion questionnaire (GIQ), a semi-structured interview guide and the students' gaming performance.

In order to evaluate students' learning outcomes, the embedded scientific concepts, the inheritance of blood-types and the principles and procedure of blood-type test and glucose test, are included in the learning achievement assessment. The assessment consists of fifteen multiple-choice items and five non-multiple-choice items (two are fill-in-the-blank questions and three are well-structured questions), and its total score is 100. The KR_{20} values of pretest and posttest were 0.75 and 0.78, respectively.

The GIQ was developed by Cheng, She, & Annetta (2014). It consists of 24 items categorized into three dimensions, engagement (9 items), engrossment (7 items) and total immersion (8 items), with a five-point Likert scale ranging from 1(strongly disagree) to 5(strongly agree). The Cronbach's alpha value of the total questionnaire and three dimensions were 0.94, 0.87, 0.87 and 0.95, respectively.

The semi-structured interview guide is designed to explore the students' problem solving patterns while playing *Bio Detective*. It is divided into three parts. The first part includes questions regarding students' perceptions of the in-game tasks (question 1 to 4). The second part consists of stimulated-recall questions, asking students what were they thinking and doing in the game using the stimulated recall method by viewing the recorded videos of their game play simultaneously (question 5 to 9). The last part is to investigate students' feelings and suggestions regarding *Bio Detective* (question 10 to 13).

Finally, students' gaming performance was recorded in the database, such as the number of times and duration students played the game and whether they have successfully completed the game or not. The gaming performance represents as the efficiency and accuracy of student problem-solving process.

2.4 Procedure

Before playing *Bio Detective,* students received a 45-min lecture about human ABO blood type first. After the lecture, the pre-test of learning achievement assessment was administrated.

Then, students were allowed to spend two sessions (90 min) playing *Bio Detective*. After finishing the play, an identical posttest of learning outcomes with different order of items was distributed to students no matter they succeed or fail the game. Moreover, the post-experience GIQ was also applied to investigate student game immersion after playing *Bio Detective*.

According to the GIQ scores, we selected four students with the lowest GIQ scores as low immersion group and the top four students as the high immersion group to conduct stimulated-recall interviews.

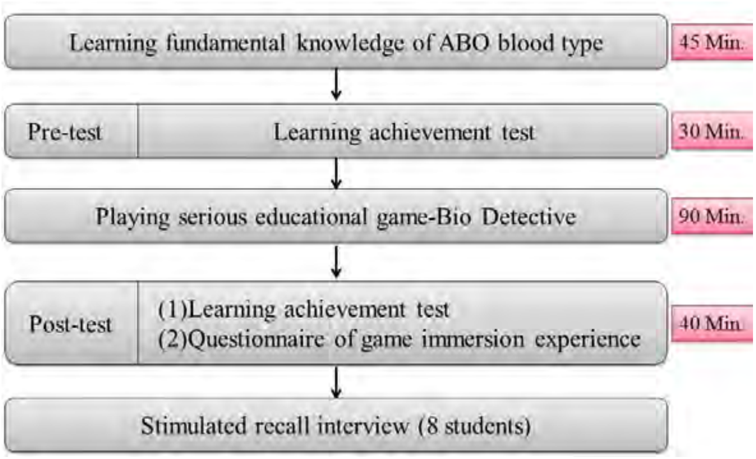


Figure 3. Experiment procedure.

3. Result

3.1 Knowledge assessment

Table 1 shows the results of paired-sample t-test. It is demonstrated that there was no significant difference between student performance on pre- and posttest of multiple-choice questions (t=-0.15, $p>0.05$), but the performance on the non-multiple-choice test had significantly improved (t=-4.57, $p<0.01$). Overall, the students' learning achievement were significantly improved after playing *Bio Detective* (t=-2.95, $p<0.01$).

Table 1: Result of the t test showing the per-test and post-test score.

	Pre-test		Post-test		t value (pre-post)
	Mean	SD	Mean	SD	
Multiple-choice test	37.92	12.95	38.08	13.54	-0.15
Non-multiple-choice test	30.69	7.87	34.77	6.30	-4.57**
Total score	68.60	19.09	73.08	18.31	-2.95**

** p<0.01

3.2 Impact of immersion on learning achievement

Table 2 shows the results of Pearson correlations between game immersion experience and students' learning achievement. It is showed that the three scales, engagement, engrossment and total immersion were highly inter-correlated. However, there was no significant correlation between game immersion experience and learning achievement.

Table 2: Result of Pearson's Correlations between Game Immersion Experience and students' learning achievement.

	Mean	SD	Engagement	Engrossment	Total immersion
Engagement	32.06	7.36			
Engrossment	20.79	7.16	0.46**		
Total immersion	23.25	8.78	0.56**	0.71**	
Pre-test	68.60	19.09	0.23	0.00	-0.15
Post-test	73.08	18.31	0.25	0.02	-0.00

** p<0.01

3.3 Impact of immersion on problem solving pathway

Figure 4 reveals the high immersion group students' problem solving pattern, inferring from the stimulated-recall interview. This shows that the high immersion group students knew what happened in the game. These students show that they had already planned how to find out the murderer before looking for the clues. They also demonstrated the ability to use the resources provided to search for information from the library or discuss with classmates when facing problems in the game. Most of them found the information provided in the library were useful. All of them completed the *Bio Detective* successfully.

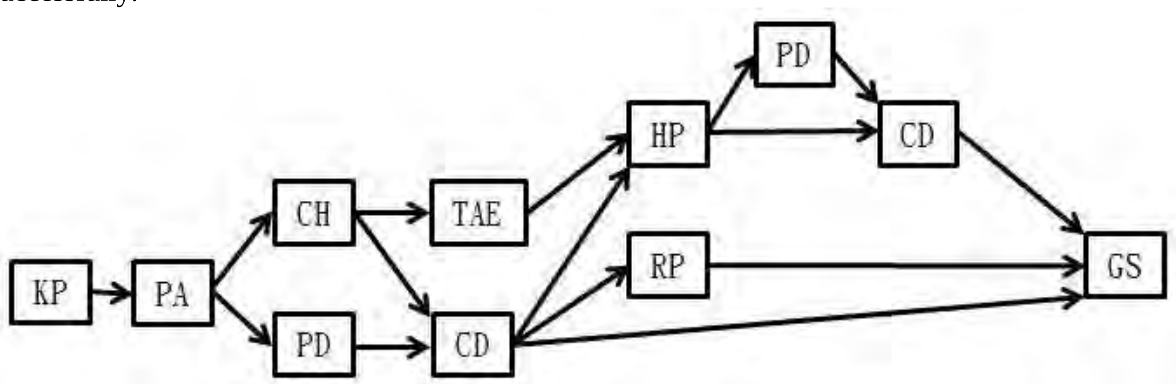


Figure 4. High immersion group students' problem solving pattern.
Note. KP= knowing the problem; PA=Plan in advance; CH=Collect hints; PD=Peer discussion; TAE=Trial and error; CD=Collect data; HP=Have problem; RP=Re-plan; GS=Games success.

Table 3 shows the results of the high immersion group students' accuracy and efficiency of problem-solving ability. Student H1, H3, H4 completed the game in 40 to 45 min. Student H2 spend much more time than others. But his learning progress was better than any others in the high immersion group (pre-test score=46; post-test score=80).

Table 3: High immersion group students' accuracy and efficiency of problem-solving ability.

Student	Game times	Playtime	Accuracy
H1	2	41 Min.	success
H2	4	55 Min.	success
H3	2	45 Min.	success
H4	2	42 Min.	success

Figure 5 depicts the low immersion group students' problem-solving pattern. It shows that the low immersion students might not know what happened in the SEG. These students required teacher's guidance in order to understand the directions of the game. In the low immersion group, not all the students had a plan to find out the murderer. They tend to use trial-and-error strategy to solve problems. Some of them might ask teachers or peers for help when they cannot find the answer.

There was one student who couldn't solve the problem, so she didn't complete the game. While another student used guessing method to identify the suspects.

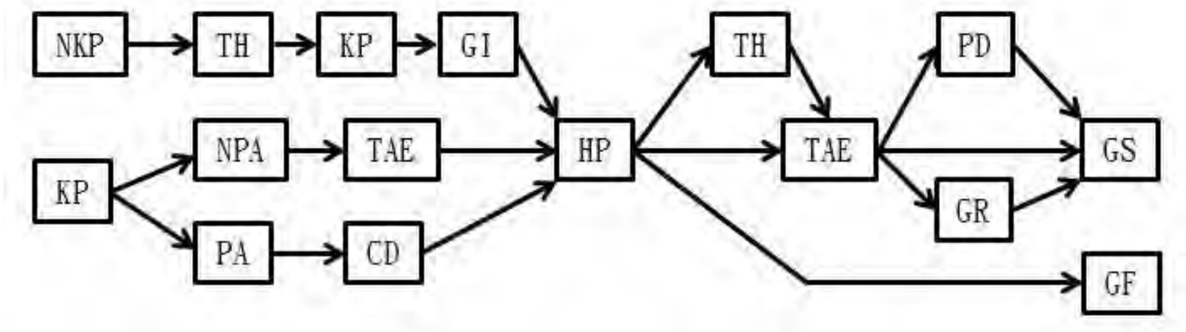


Figure 5. Low immersion group students' problem solving pattern.
Note. KP=Knowing the problem; NKP=Not knowing the problem; TH=Teacher hint; PA=Plan in advance; NPA=Not knowing the problem; GI=Games instructions; PD=Peer discussion; TAE=Trial and error; CD=Collect data; HP=Have problem; GR=Guess result; GS=Games success; GF=Games failed.

Table 4 shows the results of the students' accuracy and efficiency of problem-solving ability low immersion group. Student L3 cost less time than any other. However, she used trial and error method to complete the game. Like other students in the low immersion group, she also didn't have any plan before looking for clues, instead, she find those clues one by one.

Student L1, L3, L4 spends 55 to 65 min to finish the SEG. There was one student who failed to complete the game.

Table 4: Low immersion group students' accuracy and efficiency of problem-solving ability.

Student	Game times	Playtime	Accuracy
L1	2	59 Min.	success
L2	4	65 Min.	failure
L3	1	28 Min.	success
L4	3	55 Min.	success

Overall, the high immersion group students demonstrated higher level of problem-solving ability than students in the low immersion group. In the high immersion group, all of them completed the SEG and spend less time compare to the low immersion group.

4. Discussion and conclusions

In recent years, many studies have shown that SEGs can provide students with meaningful learning experience, as the design of SEGs attempts to combine learning content with game format which increases the opportunity of motivating and engaging students in the learning activities embedded in the game (Federation of American Scientists, 2006). In other words, proper game design promotes *immersive* experience of students and increases their willingness to learn the concepts and materials in the game (Cheng, She, & Annetta, 2014). Today, SEGs have been considered one of the potential methods for students to learn and construct knowledge (Pivec, 2007). In this study, a SEGs, *Bio Detective* was developed and its effectiveness was investigated. It is found that the participants did learn the scientific concepts embedded in *Bio Detective* as their performance on the knowledge assessment significantly improved after playing the game. The results are in alignment with previous studies illustrating that learning through playing SEGs can be effective for student biology learning (Cheng, Annetta, Folta, & Holmes, 2011; Cheng, &Annetta, 2012; Cheng, Su, Huang, & Chen, 2013; Cheng, She, &Annetta, 2014).

What is interesting in the study is that, students performed significantly better on non-multiple-choice questions rather than multiple-choice ones. As we might know that multiple-choice questions generally involve tasks related to recognition, and students all have to do is to choose or recognize the knowledge that has been learned. On the contrary, non-multiple-choice questions require students to recall relative information from their memory. Compared to recognition, it is a higher-level information processing because students have to recall specific knowledge and concepts without any clues or hints (Sternberg, 2009). Only when the knowledge has been elaborately processed by students, can they been easily and accurately recalled. As a result, it is more difficult for students to answer non-multiple-choice questions if they are not very familiar with and never elaborately process the concepts. However, our study revealed that the use of *Bio Detective* is much more helpful for students to understand what they have learned and even learned better because of their better performance on the recall tasks.

This study also explored whether the different degree of student game immersion can impact student science learning outcomes through SEG play. The results of Pearson's Correlations indicate that the three dimensions, engagement, engrossment and total immersion were highly inter-correlated. However, there was no significant correlation between game immersion experience and science learning outcomes. The results are pretty much in accordance with the previous study conducted by Cheng, She, & Annetta (2014). Researchers have suggested that cognitive load might be a key that should be considered while learning through SEG play (Cheng et al., 2014; Cheng et al., 2013). What should be taken into account is, in which aspect do the students really invest their mental effort? The game itself such as storylines and narratives or the educational aspect such as the embedded concepts and knowledge required for completing the game? Obviously, more mental efforts students spend in the game, more engaged they are. However, they might ignore the educational materials they should learn in the SEGs because of over immersing in the game.

Furthermore, immersion experience engages and even absorbs learners in a situation. While experiencing game immersion, learners invest much time and efforts in solving the tasks because of the enhancement of their internal motivation. Hence, the ability of problem solving should be considered an important element of learning outcomes of SEG play either. Only assessing the effectiveness of SEGs from the aspect of concept acquisition might underestimate the impact of using SEGs on students' science learning. The obtained results of the study additionally demonstrate that although there were no correlations between game immersion and concept learning, differences in student problem solving patterns did exist between students with high and low game immersion experience. Sternberg (2009) proposed that people who have better problem solving ability will prefer to spend time on planning how to solve the problem. On the contrary, people with poor problem solving ability will cost lots of time to trial and error, without any strategy. According to the playtime, game accuracy and problem solving pattern, we found that students with high immersion performed better than the students with low immersion in the game. The high immersion group would plan how to complete the task before finding the clues and they spent less time completing *Bio Detective* than students in the low immersion group did. Namely, in the high immersion group, students have better problem solving performance. When the students immersed in the SEG, the immersive experience might enhance learning. Students would change their perspective and engage in the game (Dede, 2009) as if they were the real detective. That is why high immersion group had better problem solving performance.

Problem solving is an important aspect of learning. Students learn at school and the teachers provide lots of knowledge for students to learn. It is hoped that students will utilize what they have learned in the classroom to solve daily life problems instead of rote learning. As a pilot study, we found that students with high immersion experience perform better problem-solving strategies in the SEG –

Bio Detective. Although SEGs provide learning experience, the problem solving ability which is affected by SEGs need to be further investigated.

Acknowledgements

The authors would like to thank the National Science Council, Taiwan, R.O.C.
for financially supporting this research under Contract No. NSC 101-2511-S-018 -004 -MY3 and NSC 102-2815-C-018-018-S.

References

Annetta, L. A. (2008). *Serious educational games*. The Netherlands: Sense Publishers.

Breeding evil? The real impact of video games. (2005, August 6-12). The Economist. Retrieved February 14, 2013, from www.economist.com/node/4247084

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. Paper presented at the CHI '04 extended abstracts on Human factors in computing systems. ACM Press. pp.1297-1300.

Cheng, M.-T., Annetta, L. A., Folta, E., & Holmes, S. Y. (2011). Drugs and the Brain: Learning the impact of methamphetamine abuse on the brain through virtual brain exhibit in the museum. *International Journal of Science Education, 33*(2), 299-319.

Cheng, M.-T., & Annetta, L. A. (2012). Students' learning outcomes and learning experiences through playing a Serious Educational Game. *Journal of Biological Education, 46*(4), 203-213.

Cheng, M.-T., Su, T., Huang, W.-Y., & Chen, J.-H. (2013). An educational game for learning human immunology: What do students learn and how do they perceive? *British Journal of Educational Technology*. DOI: 10.1111/bjet.12098

Cheng, M. -T., She, H. C., & Annetta, L. A. (2014). Game immersion experience: Its hierarchical structure and impact on game-based science learning. *Journal of Computer Assisted Learning*.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. Computers and Education, 59(2), 661-686.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*, 66-68

Federation of American Scientists. (2006). Harnessing the power of video games for learning. Technical report. Federation of American Scientists. Retrieved March 25, 2014, from http://www.fas.org/gamesummit/Resources/Summit%20on%20Educational%20Games.pdf

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming*, 33(4), 441-467. Doi: 10.1177/1046878102238607

Griffiths, M. (2002). The educational benefits of videogames. *Education and Health*, 20(3), 47-51.

Hektner, J. M. & Csikszentmihalyi M. (1996). A Longitudinal Exploration of Flow and Intrinsic Motivation in Adolescent. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY. Retrieved February 19, 2013, from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED395261&ERICExtSearch_SearchType_0=no&accno=ED395261

Hou, H.-T., Li, M.-C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior*, *30,* 29-38.

Hwang, G.-J., Yang, L.-H.,& Wang, S.-Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science coursed. Computer & Education, 69, 121-130.

Liu, C.-C., Cheng, Y.-B., & Huang, C.-W. (2011). The effect of simulation games on the learning of computational problem solving. *Computers & Education*, 57(3), 1907-1918.

Mouaheb, H., Fahli, A., Moussetad, M., & Eljamali, S. (2012). The serious game: what educational benefits?. *Procedia – Social and Behavioral Sciences*, 46, 5502-5508.

Pivec, M. (2007). Editorial: play and learn: potentials of game-based learning. *British Journal of Educational Technology, 38*(3), 387-393. Doi: 10.1111/j.1467-8535.2007.00722.x

Sternberg, R. J. (2009). *Cognitive psychology*, *5e*. Wadsworth, Cengage Learning.

Squire, K. D. (2008). Video games and education: Designing learning systems for an interactive age. *Educational Technology Magazine: The Magazine for Managers of Change in Education*, 48(2), 17-26.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. EDUCAUSE, 41(2), 16-30

The effect of students' effectiveness and attitude in heterogeneous and free grouping cooperative learning applied in sixth-grade students' Scratch program teaching

Lei Chen[a], Xiuyu Yang[a] , Xintong Wang[a], Feng−kuang Chiang[ab*]

[a]*School of Educational Technology, Faculty of Education, Beijing Normal University, China.*
Collaborative & Innovative Center for Educational Technology, Beijing Normal University, China.
*fkchiang@bnu.edu.cn

Abstract: Although the method of cooperative learning has been used popularly, but researchers paid less attention to different grouping way in the past, especially heterogeneous and free grouping. In this paper, researchers analyzed students' effectiveness and attitude in heterogeneous and free grouping cooperative learning in sixth-grade students' Scratch program teaching. 52 students participated in the study and the Scratch animation works assessment scale was prepared and revised by some experts, students and Scratch teaching teachers. Many interesting things have found in the study, such as heterogeneous grouping brings better students' works, and more efficiently arouses students' curiosity. In addition, free grouping stimulates more positive attitudes, while heterogeneous grouping leads to more learning pressure, and boys have less pressure in Scratch learning than girls. This can provide some constructive suggestions for cooperative learning.

Keywords: heterogeneous grouping, free grouping, Scratch, program teaching

1. Introduction

Nowadays, with the rapid development of modern information technology, technology plays a critical role in the education reform，and the ability of ICT has been regarded as essential skills for primary and middle school students. Since Ministry of Education published the curriculum guidelines of primary and middle school in 2003, education of information technology has been a major issue. And at once setting forth from the third level, students deliver a single lesson of information technology every week at least, thus as to improve the ability of ICT. Though students in primary can learn many things from lessons of information technology, programming lessons are abandoned or delayed until middle or senior high school.

Scratch is a new programming environment developed by MIT Media Lab. It is suitable for children aged over eight years old to make their own small programs creatively, such as interactive stories, animation, games and art (Chen, 2009). He (2013) also said that students could make a program easily by dragging blocks, which is suitable for young children. Some teachers have used the method of cooperative learning, but they don't give enough attention to it, and whether different grouping has some close relationship with students' effectiveness is still unknown.

2. Literature review

2.1 Study of Scratch programmed instruction methods

Xie (2013) puts forward a teaching method "creating situation—analyzing cases—interacting between teachers and students—design (personalized works) —sharing and communicating" by exploring the teaching model and the method of Scratch. Xu & Huang (2013) found that collaborative project-based on learning could enhance students' effectiveness and attitudes toward Scratch. With the didactic and inquiry teaching methods respectively, Yang (2010) tried teaching the Scratch program design, and the results showed that with these two kinds of teaching methods on the fifth grade primary school students, there were not significantly different, and the teaching method and the mathematics learning achievement didn't have close relationship. Ke (2013) explored the potentials of computer-aided and mathematics game making activity based on Scratch on promoting students' mathematics learning. Research showed, participants who involved in making computational game were more positive to mathematics learning. And the experience and design of game could connect to daily mathematical experience.

In summary, students usually make a Scratch work in a group, whose aim is to express their own feelings by engaging in developing comprehensive abilities, cultivating creative thinking, and learning how to cooperate with others (Brennan & Karen, et al, 2010). So the method of cooperative learning is suitable for Scratch learning. However, there are few studies about Scratch program cooperative learning of pupils. So researchers tend to focus on this area.

2.2 Studies of free grouping and heterogeneous grouping

Zhang (2006) made a study based on heterogeneous grouping, and after a semester's study the experimental group whose members were divided into groups by their basic ability got higher scores than the control group. The research showed that the heterogeneous grouping could improve different aspects of ability. A research (Deng & Huo, 2010) also found the basketball skills of heterogeneous grouping had been improved much more than that of free grouping. Teaching activities may be carried out in accordance with students' aptitude.

All in all, there are no definite conclusions about whether different grouping in programmed learning will bring disparate imparts on students' effectiveness and attitude. Similarly heterogeneous and free grouping are often used during daily teaching activities, but studies about students' effectiveness based on heterogeneous and free grouping are few.

3. Method

3.1 Participants

52 students in two classes in the 6th grade of Elementary School in Beijing are included in the research. The classes were assigned as the experimental group (22 students，Class 1) and a control group (30 students，Class 3). For 8 weeks, the same Scratch program teaching was conducted in both two groups. Only one researcher taught them Scratch.

3.2 Research design

Despite the lack of pre-test, two classes are normal classes, which aren't selected and changed. Given that students had no significant differences in effectiveness and attitude at first, One class (class 1) is free grouped, and the other one (class 3) is heterogeneous grouped. One group has two students. Then through evaluation of the works and an attitude questionnaire survey of cooperative learning groups, researchers compare the Scratch programming learning effectiveness and attitudes of students grouped by two different ways.

According to the objectives of the study and relevant literature information, the framework of this research was established as follows.

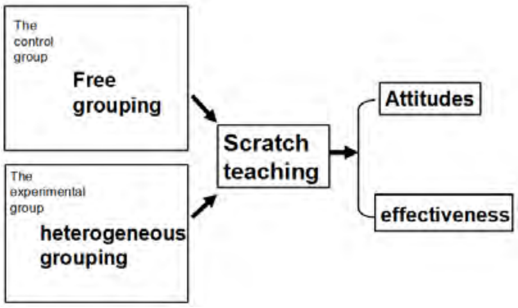


Figure 1. Framework of Research

3.3 Course design

According to the researchers' teaching practice and the general process of cooperative learning in the programming previously submitted by Ju （2007）, the teaching model during the research is shown in the following figure:

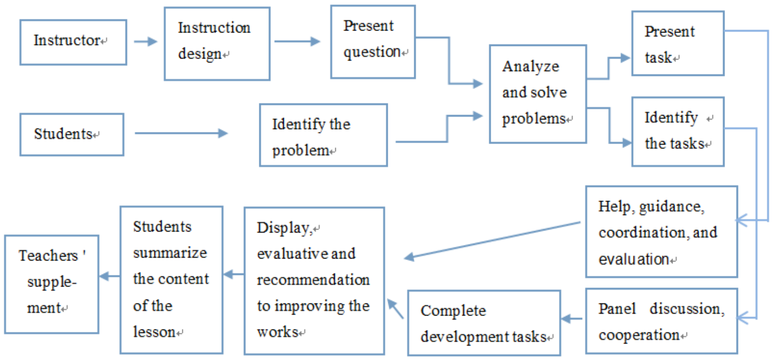


Figure 2. Programming cooperation model (Source: revised from chart of Ju, 2007)

During the eight-week research with one lesson a week each class, designed by the author, the teaching content was divided into four sections: broadcast, asking-answering, logical operators, and making Scratch works. Instructor, also a researcher, taught Scratch in two pilot classes in the same way.

As shown in the figure, cooperative learning in each class has the following main procedures: Based on the pupils ' previous knowledge, the teacher designs, reasonable and practicable teaching and learning activities, presents problems of cooperative learning and tasks after solving the knowledge problems. In task, with learned knowledge, students design Scratch programming theme, analyze problems and determine the division after group discussion, and solve problems by group cooperation. Then the teacher and students, show group works, evaluate performance of another group and present improvements together. At last students' summary what they have learnt in this class.

3.4 Research tools

3.4.1 Scratch animation works assessment scale and learning attitudes questionnaire of Scratch

The assessment scale was developed by the researcher and revised by two experts, two masters and five Scratch teaching teachers across the country, consisting of four dimensions: design, aesthetic, technology and innovation. The scores of Cronbach's Alpha are both more than 0.8.

The questionnaire consists of five dimensions, respectively, motivation, self-efficacy, usability, usefulness, autonomous learning and learning pressure. It was developed by researchers referring to Professor Huang Guozhen's questionnaire of learning attitude and scientific accept mode (Hwang，Tsai & Tseng, 2010). Score of Cronbach's Alpha in one class is 0.895, and another is 0.933.

4. Results

4.1 Analysis of students' animation works

At the end of the term, works of each group were collected, but some group did not hand in their works finally. There were twenty-two animation works in total, nine from class 1 and thirteen from class 3. According to Scratch animation works assessment scale, three Scratch researchers (undergraduate students), gave a score for each animation seriously. The average score of three undergraduate students was the final score of animation works. Because each class had less than 30 students, and some students did not hand in their works, so samples might be not enough. When analyzing the samples, the method of nonparametric analysis was used to analyze works between class 1 and class 3 comparatively.

4.1.1. Heterogeneous grouping might bring better students' works

By nonparametric analysis of animation works between class1 and class 3, it indicated that these two kinds of grouping has no significant differences in scores of works. But According to descriptive analysis, it indicated that the average score of students' works in class 3 was about 79.0 points and the average score of students' works in class1 was about 83.7, four point seven points higher than works in class 3.

Table 1: Descriptive Data of students 'works in different dimensions

category	class	N	Mean	Std. Deviation	Std. Error Mean
total	1	9	83.72	6.45	2.15
	3	13	78.97	6.14	1.70
design	1	9	21.20	1.47	0.49
	3	13	19.66	2.11	0.59
aesthetics	1	9	20.39	2.00	0.67
	3	13	20.25	1.57	0.43

technology	1	9	21.48	2.03	0.68
	3	13	19.60	1.34	0.37
innovation	1	9	20.69	1.60	0.53
	3	13	19.44	1.69	0.47

Table 2: The results of Nonparametric analysis of students' works in different classes in different dimensions

Null Hypothesis	Sig.
The distribution of total is the same across categories of class	.03*
The distribution of design is the same across categories of class	.014
The distribution of aesthetics is the same across categories of class	.65
The distribution of technology is the same across categories of class	.03*
The distribution of innovation is the same across categories of class	.13

Then researchers analyzed the scores in four dimensions of Scratch animation works assessment scale, found that only the dimension of technology had significant differences between the two classes. And the average score of technology in class 1 was about 21.5 points, 1.9 points higher than that in class 3.

By nonparametric analysis on items of the dimension of technology, it indicated that score of the item "Scratch can arouse my curiosity" had significant differences between the two classes. And The score in class 1 was 0.5 points higher than that in class 3.

4.2 Analysis of students' attitudes towards Scratch

4.2.1 Free grouping brought out more positive attitudes.

All participants completed the attitude questionnaire towards Scratch at the end of term, and forty effective questionnaires were collected finally, 27 from class 3 and 13 from class 1. By nonparametric analysis of attitudes between class1 and class 3, it indicated that these two kinds of grouping had significant differences about attitudes towards Scratch. Descriptive statistics showed that the score of attitudes in control class(class 3) was 16.5 points higher than that in experimental class（class1）, and the score in each dimension of attitudes in class 3 was higher than class 1. So students in free grouping had more positive attitudes towards learning Scratch than students in heterogeneous grouping.

Table 3: Descriptive Data of students' attitudes in different class in different dimensions

category	class	N	Mean	Std. Deviation	Std. Error Mean
total	1	13	40.77	19.76	5.48
	3	27	57.26	20.73	3.99
motivation	1	13	6.77	3.52	0.98
	3	27	9.30	5.10	0.98

self-efficiency	1	13	8.23	5.20	1.44
	3	27	10.26	4.79	0.92
usability	1	13	8.00	4.53	1.26
	3	27	10.78	4.94	0.95
usefulness	1	13	6.62	2.96	0.82
	3	27	9.11	4.30	0.83
autonomous learning	1	13	5.31	3.52	0.98
	3	27	6.70	2.88	0.55
pressure	1	13	5.85	3.02	0.84
	3	27	11.11	5.32	1.02

4.2.2 Heterogeneous grouping might lead to more learning pressure.

The scores in the six dimensions of learning attitude questionnaire of Scratch were analyzed, and the result shows that the dimensions of usability, usefulness and pressure had significant differences between the two classes, especially learning pressure between two classes had abnormally significant differences (sig=. 00**).

Table 4: The results of Nonparametric analysis of students ‘works in different classes in different dimensions

Null Hypothesis	Sig.
The distribution of total is the same across categories of class	.01*
The distribution of motivation is the same across categories of class	.011
The distribution of self-efficiency is the same across categories of class	.014
The distribution of usability is the same across categories of class	.04*
The distribution of usefulness is the same across categories of class	.03*
The distribution of autonomous is the same across categories of class	.11
The distribution of pressure is the same across categories of class	.00**

Then researchers analyzed the five items of learning pressure detailedly and found that three of them had significant differences between the two classes. These items included” Usually I can’t concentrate on learning Scratch.”, “I'm stressed when using Scratch.”, and “It takes me much time to master the use of Scratch.”

A descriptive analysis was carried out, and reverse scoring was used in the dimension of learning pressure when analyzing data in SPSS, so it means the more score of learning pressure is, the less pressure students have. From Table 3, it indicated that the score in each item of pressure in class 3 was higher than that in class 1, so students in class 3 have less learning pressure than that in class 1. In other words, students in heterogeneous grouping feel more stressed than students in free grouping. Table 3 shows the score in each item of learning pressure in class 3 was respectively higher than that in class 1. So heterogeneous grouping makes students feel more stressed. They couldn’t concentrate on learning Scratch, felt stressed when using Scratch and needed to spend more time to master it.

4.3 Analysis of gender impact on student attitudes

The scores in six dimensions of learning attitude questionnaire of Scratch were analyzed and it shows there were no significant differences totally between boys and girls in two classes. However, boys and girls had abnormally significant differences in pressure, and boys had less pressure than girls in Scratch learning. The results of nonparametric analysis of students' attitudes between boys and girls in different dimensions are shown in table 5.

Table 5: The results of nonparametric analysis of students' attitudes between boys and girls in different dimensions

Null Hypothesis	Sig.
The distribution of total is the same across categories of class	.09
The distribution of motivation is the same across categories of gender	.78
The distribution of self-efficiency is the same across categories of gender	.91
The distribution of usability is the same across categories of gender	.20
The distribution of usefulness is the same across categories of gender	.19
The distribution of autonomous is the same across categories of gender	.17
The distribution of pressure is the same across categories of gender	.00**

5. Conclusions and discussion

5.1 Analysis of students' works

Students in heterogeneous grouping got more scores than students in free grouping, which means heterogeneous grouping may lead to more positive effect in effectiveness than free grouping. Also, students in class 1 had significant differences in technology, especially in the item of "Scratch can arouse my curiosity". So Scratch teaching in heterogeneous grouping could more efficiently arouse students' curiosity and stimulate a strong desire to learn.

5.2 Analysis of students' attitudes

According to the nonparametric analysis of attitudes between class1 and class 3, it indicated that different grouping made significant differences about attitudes towards Scratch. Descriptive statistics show that the score of attitudes in free grouping was 16.5 points higher than that in heterogeneous grouping, and the score in each dimension of attitudes in free grouping was higher than heterogeneous grouping. Free grouping brings out more positive attitudes.

Dimensions of usability, usefulness and pressure had significant differences between the two classes, especially learning pressure (sig=.00). From Table 3, it indicated that students in heterogeneous grouping feel more pressure than students in free grouping. In other words, Heterogeneous grouping may lead to more learning pressure, particularly in these aspects: concentrating on learning Scratch, using Scratch and costing more time to master it.

Boys and girls have abnormally significant differences in the dimension of pressure both in two classes, and boys have less pressure than girls in Scratch learning. Maybe girls have difficulty in programming, and some studies have confirmed this inference.

5.3 Research Limitations

This study analyzed students' effectiveness and attitude in heterogeneous and free grouping cooperative learning in sixth-grade students' Scratch program teaching, and some distinguished results have been shown above. However, without pre-test, it may have some influence on the later analysis of the results. The research design will be improved and more differences in heterogeneous and free grouping are going to be found in the future.

References

Adams-Byers J, Whitsell S S, Moon S M (2004). Gifted students' perceptions of the academic and social/emotional effects of homogeneous and heterogeneous grouping. *Gifted Child Quarterly*, 48(1): 7-20.

Brennan, K., Monroy‐Hernández, A., & Resnick, M. (2010). Making projects, making friends: Online community as catalyst for interactive media creation. *New directions for youth development*, 2010(128), 75-83.

Jie Chen. (2009). Introduction of Scratch and application. *Computer knowledge and technology*, 5 (26).

Ferrer-Mico, T., Prats-Fernàndez, M. À., & Redo-Sanchez, A. (2012). Impact of Scratch Programming on Students' Understanding of Their Own Learning Process. *Procedia-Social and Behavioral Sciences*, 46, 1219-1223.

Huimin Ju, Wuzi Cui, & Lihong Sun (2007).Design and practice of cooperative learning in program design course. *The journal of Beijing United University*, (03): 85-88.

Hwang,G.J.,Tsai,C.C.,& Tseng, Judy C.R. (2010).A two-tier test approach to tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education*,55(4),1618-1627.

Thanh, P. T. H.,& Gillies, R. (2010). Group composition of cooperative learning: Does heterogeneous grouping work in Asian classrooms?. *International Education Studies*, 3(3).

Ke, F. (2014). An implementation of design-based learning through creating educational computer games: A case study on mathematics learning during design and computing. *Computers & Education*, 73, 26-39.Teresa

Xianxiao Xie. (2013). Research on Scratch teaching model – based on "Infinite lines". *China Information echnology Education*, (3), 56-58.

Xiaotong Xu, & Yaping Huang. (2013). Study on Scratch collaborative learning achievement of different students.

Xiuhai Zhang. (2006). The teaching and research of cooperative learning group strategy. Journal of Liaoning Teachers College: *Social Science Edition*, (2), 75-76.

Yinting He. (2013). Analysis on the effects of Scratch program teaching based on guided learning.

Zhiyou Deng & Zhongyang Huo. (2010). The research of heterogeneous group teaching based on basketball class of primary school. *Sport*, (9), 64-65.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Designing Mobile Application for STEM: Building Individual Interest and Supporting Creative and Innovative Thinking Skills

Ilker YENGIN*
*Institute of High Performance Computing, A*STAR, Singapore*
*yengini@ihpc.a-star.edu.sg

Abstract: This paper explains the design of a mobile application that is aimed to support the development of creative and innovative skills using stories as the external factors to create and maintain students' long-term individual interests in STEM area. In this paper, first, we discussed the connection between individual interests and creative – innovative thinking process. Then, we described research proved pedagogical methods, which based on using stories to build and maintain students' individual interests. Following that we illustrated the design of mobile application that applies the pedagogical methods into practice. In the discussion part, we discussed the potential use of the application as well as future follow-up research studies.

Keywords: STEM, creativity, creative thinking, innovative thinking, mobile learning, interest building, individual interest, mobile application design.

1. Introduction

Creative and innovative thinking are couple of the key thinking skills that would be core to science, technology, engineering and mathematics (STEM) education (Atkinson and Mayo, 2010). Developing students' creative thinking skills have significant role in building modern economies.

Investments on creativity and innovation had a dramatic positive impact on the global competitiveness of countries (MacLeod, etl.al., 2007). Similarly, research on highly creative and innovative countries (e.g. Finland, Norway etc.) showed that supporting creative and innovative thinking clearly facilitates the transition from the postindustrial economy, towards an emergent knowledge-based creative economy (Duell, Wright and Roxburgh, 2014). To remain competitive in a challenging global environment, supporting creative and innovative thinking in work and learning plays a crucial role (Ferrari, Cachia and Punie, 2009).

Specially designed work and learning spaces encourage us to think in creative and innovative ways (Oksanen and Ståhle, 2013). According to a meta-analysis of 32 research studies on creativity, enabling social interaction is identified as one of the significant category of the key characteristics of conditions promoting creativity and innovation (Davies et.al, 2013). Similarly, it is possible to provide same characteristics into STEM education to support creativity and innovation (Cooper and Heaverlo, 2013). For instance, using technology to provide social understanding may have different positive influences on creative and innovative thinking skills in STEM (Kärkkäinen and V.Lancrin, 2013). However, solely providing technological tools are not sufficient to develop creative and innovative thinking skills by itself (Selwyn, 2011; Jackson et.al., 2012). It is suggested that technology should often be perceived as a catalyst for change in the application of pedagogical methods in order to have successful results (Watson, 2011).

One of the pedagogical methods that help to have successful results is to helping students to develop an individual interest (Arnone et.al., 2011). For instance, increasing individual interests on a particular subject/object enhance creativity (Shalley, Zhou and Oldham, 2004; Grant and Berry, 2011). Developing students' individual interests towards the STEM related learning activities help them to increase their chance to develop creative and innovative thinking skills (Bairaktarova and Evangelou, 2012).

It has been suggested that students with individual interest for STEM subjects are likely to be motivated to pursue the STEM careers that require creative and innovative thinking (Tyler-Wood, Knezek and Christensen, 2010). Developing interests in STEM requires carefully planned interactive activities in school environment. Informing students on the topic is not a sole component of developing interest (Hidi, Renninger and Krapp, 2004). There should be activities supporting both positive feelings and opportunities for gathering knowledge for social understanding in order to be able to shift students' interests on STEM areas (Renninger and Shumar, 2002). Instead of simply supporting having fun and participation in science, institutions should also provide activities for productive social understanding by addressing individual interests (Stocklmayer and Gilbert, 2002). Research of U.S. National Research Council (2006) clearly indicated that students shows more productive participations in STEM activities when they developed individual interests by finding meaning thorough social understanding in the learning strategies. Social understanding develops through the process of observation, introspection, and imagination (Carpendale and Lewis, 2004).

Student's individual interests in STEM may not be influenced all the time by the school related activities (Hidi, 1990). First of all identifying students' individual interests would need a great effort. Second, these interests may not find a place to be pointed in school subjects. Thus, it is important to understand what types of activity may address interests of different range of the student population. Interest on STEM subject might be an outcome of interactive activity or a content in which the activity and the environment may have an effect on the dynamics of the interests (Krapp, 2005). Other external motivators such as role models (e.g. scientists, teachers) and peers may contribute to an increased emerging individual interest (Krapp and Lewalter, 2001).

After individual interest is developed well, students usually have a tendency to pursue further understanding independently and work on developing deep understanding in STEM fields (Renninger, 2000). It is expected that external factors allowing the well-developed individual interest make it possible for students to maintain positive feelings and interests for STEM related activities in future. Thus it is very important to create a learning environment that provides set of social activities within the inclusion of knowledge from the external individuals in order to create interests in STEM. Also engaging social activities with an established individual interest may extend creative and innovative thinking as suggested in several studies (Stocklmayer and Gilbert; Tyler-Wood, Knezek and Christensen, 2010).

As mentioned above paragraphs, it is expected that individual interest support creative and innovative thinking. Also external motivators could increase it. Thus building individual interests for students becomes a very important factor in success of students in STEM. Addressing the needs, this study discusses the design of a mobile learning application using a pedagogical model that implements core dynamics of individual interest building using the external motivators in order to develop creative and innovative thinking in STEM field.

The pedagogical methods that are using external motivators to increase individual interest in STEM are discussed in the first part of the paper. Second part of the paper illustrates the design of the mobile learning application, which is conceptualized according to the pedagogical methods in the first part. Finally a discussion part explains the possible use case scenarios of designed application in STEM education.

This paper would be useful for the instructional designers, teachers and school administrators who like to learn more about developing creative and innovative thinking skills and individual interests in STEM field as well as an example design of mobile application for this purpose.

2. Pedagogical Methods

2.1 Increasing Individual Interest

Students with high interest in STEM subjects may possibly have gained more knowledge than their peers due to the fact that they may put more effort on the STEM activities (Tobias, 1992). This implies that students with prior knowledge on the STEM area may have increase individual interest. The argument is also supported by the research. Research suggested that as students become more familiar with an area, their interest toward the areas is expected to increase (Tobias, 1994).

The external sources of information could be integrated in STEM education to increase the prior knowledge of the students. Especially, external institutions (e.g. science centers, research companies, universities and museum etc.) and individuals (e.g. scientists, innovators , artists and researchers etc.) may be a great external resources of information. As the research shows, connecting and having a regular relationship with these institutions and individuals in an out of school may develop a personal interests slowly over time with a tendency of having a long term effect on students' knowledge and values on the subject (Ainley and Ainley 2011; Dabney et.al., 2012; Renninger, Hidi and Krapp 2014). Thus, following the research based suggestions; it is assumed that when students start connecting to institutions and individuals in and out of school, they would have greater prior knowledge on the STEM related areas, which would also increase individual long-lasting interests on STEM area.

Based on this, it is possible to have institutions and individuals as external factors to increase the individual interests. As a pedagogical method, we could seek for possibilities to establish connection with external institutions and individuals. In the application design, we implement elements to apply this practice to increase students' individual interests.

2.2 Using Stories

Stories create engaging and pedagogically effective learning experiences (Mcquiggan et.al., 2008; Kuyvenhoven, 2005; Dyer and Wilkins, 1991). Stories are very effective in learning and theory building (Badreddine and Buty, 2011). Using stories increase students' interests on a particular subject (Fulmer and Frijters, 2011). Similarly, stories are used as pedagogic practices in social learning to support science education (Shelby and Ernst, 2011). Stories that are interesting motivate students to continuously to keep their attention on the subject. For instance, students reading stories with elements such as novelty, character identification, life themes, and activity level may have a continuous interest to get more knowledge about the topic (Hidi, 2001; Hidi and Baird, 1986) including science (Sandoval, 1995). Also stories are powerful applications with a potential to creating an emotional connection to the individual interests that results extending with some degree of continuity on interaction (Tan, 2013). Also stories enable students to make connections between ideas in the ongoing STEM issues which help them to have a personal meaning and interests towards the science concepts (Scott, Mortimer and Ametller, 2011). Stories are also allowing teachers and students to have a platform to share the experiences of the individuals in STEM area to help students to grasp the related concepts, accommodate with diverse perspectives and realties of the experts in the field (McDrury and Alterio, 2003).

It is an accepted idea that enabling students to write stories are commonly used to support their development of the creativity (Hennessey and Amabile, 1988). Similar to writing stories, it is suggested that reading stories are also supporting students to have some initial ideas to be able to have some creative associations (Smogorzewska, 2014). Thus, the system design is expected to have an impact on students' creativity with a certain degree. Reading stories approach is expected to grab and maintain the students' attention. Especially some interesting elements such as life themes of scientist and how they react to the challenges in daily life, their solutions and services to the important problems of global scene etc. is expected to arouse the students' curiosity which will also expected to turn into the individual interests with a long term interactions. Thus, carefully selected stories of external individuals in STEM area may help students to understand the tasks involved in the professional STEM related fields such as decision making process in daily tasks, theories, tools and work environments. As the second pedagogical method, we implement the idea and key elements of using stories to create individual interest in STEM related areas

3. System Design

Implementing the pedagogical methods described in Section 2, we designed a learning environment where students have a chance to use a platform to connect to the science and read stories about the problems in science and the scientists' lives. As research studies suggested, by connecting students to science using the stories as a meeting point, it is expected to maintain students long-term individual interests on STEM areas. Also as an effect of continuous long-term individual interest, it is expected that students may develop their creative and innovative thinking skills, which would be essential in STEM education. To enable the pedagogical methods, we designed the system architecture prototype as in Figure 1. Accordingly, story content collector engine collects the STEM related stories form different online resources using RSS then sort and tags into database. Distributor engine sends out the stories to student based on the rules defined by individual interest profiler. Individual interest profiler creates a dynamic profile of individual interest for each student in the system and feeds the distributer engine with rules. Interactive interface displays the content received from the distributor engine with a set of interactive events.

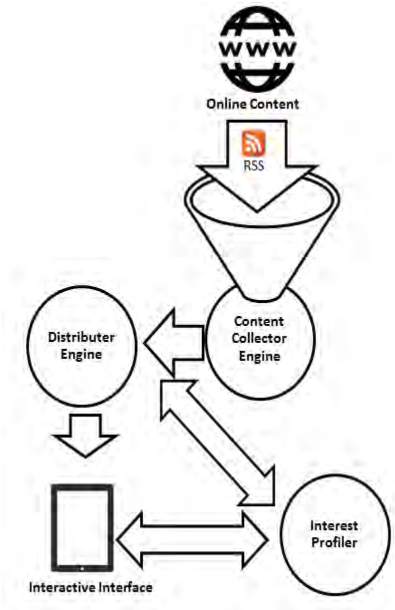


Figure 1. Mobile Application System Architecture

3.1 Story Collection and Distribution

Story contents are fed into the system in forms of newsletters and short articles forms, which are collected from the online resources. The content collections are categorized in a variety of topics. Themes would be challenges in science, new scientific and technical developments as well as professional lives of scientists.

The content collection is centralized and processed by the story content collector engine using “Really Simple Syndication” (RSS). The content collector engine signs up automatically to the various online resources related to STEM subjects and collects the content to centralize them before the distribution. Once the system administrator set up the resources’ addresses for the content collection the content collector will continue to function for as long as the system runs. System administrators have no need to check the online resources to find out what is going on and whether there is a new update available. All the updates and stories come automatically to the content collector engine and sorted and tagged according to the topics.

The distributor engine sends out the stories to students according to a rule set defined by the individual interest profiler. When students log in to the system, they also start building their profile in the system. According to their profile, students would regularly receive stories related to STEM fields, which are selected and personalized according to each student’s individual interests in the topic. The selection of the stories according to individual interest analysis is described in the next section.

3.2 Individual Interest Profiler

The individual interest profiler identifies students’ individual interests. The individual interests and related mechanism is dynamically calculated according to students’ previous selections of the topic as well as asking feedback from the students. When students registered into the system they are asked to fill a registration form. In the registration form, students are asked to respond to a questionnaire developed according to the instruments that measure the students' views on nature of science (Osborne, Simon and Collins, 2003; Chen, 2006). Furthermore, students are also asked to indicate which topics and subjects they feel they like to read in STEM areas. Moreover, in addition to the initial data collection, after each time reading a story, students are asked to give feedback on their individual interest in the stories. Students indicate their self-interest using a 5 points rating scale. Our system does not rely on self-rating data, it also have its own tracking and measurement system embedded in individual interest profiler. Accordingly, the system tracks and measures students' interactions by activity analysis, clicks and patterns of behavior in relation to other interface and content.

3.3 Interactive Interface

Interactive interface displays the story content. This interface is where students engage in to the stories. It is designed for mobile devices including tablets, which allow students to benefit from the application in and out of the school and to follow the content according to their own phases.
The interactive interface is designed as a scientist's diary application (e.g. design of look and feel of the application, activities etc.) where students would get their STEM stories regularly.
In the application students have a library where he could select stories that are updated according to his/her individual interests regularly (Figure 2).

Figure 2. Mobile Application Interactive Interface for reading STEM related stories

The interface has basically functions as e-readers application. However, to have a look specific to STEM education, the reading interface provides additional design functions to create experiences for the students to help them feel like they are engaging in scientific diary reading experiences. This special design makes the interface design different than the traditional e-readers. More than just embedding the content from the online source and delivering it on a tablet device without adding any value to the content, the interface have interactive features such as scrolling up and down in a blocks of text and using some artwork to resemble a science related reading activity. In the application, the reading interaction is scrolling based in which students need to scroll up and down to see the contents and swipe it right and left to turn the pages. In that sense, the interface captures the essence of reading science diary rather than reading a book. Adding the interactivity, the interface creates a meaningful science reading experience for the students. Interface also has standard functions such as search, bookmark, sharing, printing annotation and taking notes.

4. Discussion and Future Studies

The mobile application intended to build students' self interest, which has a direct connection in supporting the creative and innovative thinking process. Using the mobile application to deliver science related stories as the external resources for motivational process; it is aimed to increase the individual interest in STEM areas. Relying on the research studies that show the connection between the interest and creative – innovative thinking, we expect that the mobile application will be helpful to support the development of that mentioned skills of creative and innovative thinking. The design of mobile learning application was based on the implementations of pedagogical methods that we picked based on the suggestions of several reach studies as explained with details in the previous chapters.
This mobile application can be used to increase students' long-term interaction in STEM subjects. It could be used as a part of class project where students may read regular science stories in the classroom. As the application is good for tablet devices, students may use it in and out of the classroom.

As a future study, we expect to test the mobile application in terms of design effects on helping students in developing motivation and the creative – innovative skills within experimental studies.

5. Conclusion

In this study, using the stories in STEM areas as the external motivators, we designed a mobile application that helps building and maintaining students' individual interest in STEM area. As previous studies suggested it is expected increasing interest would directly support students' creative and innovative thinking. For addressing this issue, applying the pedagogical models using stories as the external motivators, we provided a design for a mobile application. The mobile application is expected to build and maintain students' individual interests, which also result in development of their creative and innovative skills in STEM area. The pedagogical methods and the design of the mobile learning application are discussed in the course of the paper. Finally in the discussion part, the possible use case scenarios and future studies are mentioned shortly.

This paper would be useful for the instructional designers, teachers and school administrators who like to learn more about developing creative and innovative thinking skills and individual interests as well as an example design of mobile applications for this purpose.

References

Ainley, M., & Ainley, J. (2011). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology*, *36*(1), 4-12.

Arnone, M. P., Small, R. V., Chauncey, S. A., & McKenna, H. P. (2011). Curiosity, interest and engagement in technology-pervasive learning environments: a new research agenda. *Educational Technology Research and Development*, *59*(2), 181-198.

Atkinson, R. D., & Mayo, M. J. (2010). Refueling the US innovation economy: Fresh approaches to science, technology, engineering and mathematics (STEM) education. *The Information Technology & Innovation Foundation, Forthcoming*.

Badreddine, Z., & Buty, C. (2011). Discursive reconstruction of the scientific story in a teaching sequence. *International Journal of Science Education*, *33*(6), 773-795.

Bairaktarova, D., & Evangelou, D. (2012). Creativity and science, technology, engineering, and mathematics (STEM) in early childhood education. Contemporary perspectives on research in creativity in early childhood education, 377-396.

Betcher, C., & Lee, M. (2009). *The interactive whiteboard revolution: Teaching with IWBs*. Aust Council for Ed Research.

Carpendale, J. I., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. Behavioral and Brain Sciences, 27(01), 79-96.

Catala, A., Jaen, J., van Dijk, B., & Jordà, S. (2012). Exploring tabletops as an effective tool to foster creativity traits. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* (pp. 143-150). ACM.

Chen, S. (2006). Development of an instrument to assess views on nature of science and attitudes toward teaching science. *Science Education*, *90*(5), 803-819.

Cooper, R., & Heaverlo, C. (2013). Problem Solving And Creativity And Design: What Influence Do They Have On Girls' Interest In STEM Subject Areas? . *American Journal of Engineering Education (AJEE)*, *4*(1), 27-38.

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012). Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B*, *2*(1), 63-79.

Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education—A systematic literature review. *Thinking Skills and Creativity*, *8*, 80-91.

Duell, C., Wright, N., & Roxburgh, J. (2014). Developing 'Design Minds' for the 21st Century Through a Public Sector Initiated Online Design Education Platform. *Design and Technology Education: an International Journal*, *19*(1).

Dyer, W. G. J., & Wilkins, A. L. (1991). Better stories, not better constructs, to generate better theory: A rejoinder to Eisenhardt. Academy of Management Review, 16(3), 613–619.

Evans, M. A., & Rick, J. (2014). Supporting Learning with Interactive Surfaces and Spaces. In *Handbook of Research on Educational Communications and Technology* (pp. 689-701). Springer New York.

Ferrari, A., Cachia, R., & Punie, Y. (2009). Innovation and creativity in education and training in the EU member states: Fostering creative learning and supporting innovative teaching. *JRC Technical Note*, *52374*.

Friess, M.R, Kleinhans, M., Forster, F., Echtler, F., & Groh, G. (2010). A Tabletop Interface for Generic Creativity Techniques. In *Proc. MCCSIS '10*. IADIS , 203-210.

Fulmer, S. M., & Frijters, J. C. (2011). Motivation during an excessively challenging reading task: The buffering role of relative topic interest. *The Journal of Experimental Education*, *79*(2), 185-208.

Geyer, F., Klinkhammer, D., & Reiterer, H. (2010, November). Supporting creativity workshops with interactive tabletops and digital pen and paper. In *ACM International Conference on Interactive Tabletops and Surfaces* (pp. 261-262). ACM.

Grant, A. M., & Berry, J. W. (2011). The necessity of others is the mother of invention: Intrinsic and prosocial motivations, perspective taking, and creativity. *Academy of Management Journal*, *54*(1), 73-96.

Hämäläinen, R., & Vähäsantanen, K. (2011). Theoretical and pedagogical perspectives on orchestrating creativity and collaborative learning. *Educational Research Review*, *6*(3), 169-184.

Hennessey, B. A., & Amabile, T. M. (1988). Story telling: A method for assessing children's creativity. The Journal of Creative Behavior, 22(4), 235-246.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational research*, *60*(4), 549-571.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review*, *13*(3), 191-209.

Hidi, S., & Baird, W. (1986). Interestingness—A neglected variable in discourse processing. *Cognitive Science*, *10*(2), 179-194.

Hidi, S., & Baird, W. (1988). Strategies for increasing text-based interest and students' recall of expository texts. *Reading Research Quarterly, 23,* 465-483.

Hidi, S., Renninger, K. A., &Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.), Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development (pp. 89–115). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Jackson, L. A., Witt, E. A., Games, A. I., Fitzgerald, H. E., von Eye, A., & Zhao, Y. (2012). Information technology use and creativity: Findings from the Children and Technology Project. *Computers in human behavior*, *28*(2), 370-376.

Kärkkäinen, K., & Vincent-Lancrin, S. (2013). Sparking Innovation in STEM Education with Technology and Collaboration.

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. Learning and Instruction, 15, 381–395.

Krapp, A., & Lewalter, D. (2001). Development of interests and interest based motivational orientations: A longitudinal study in vocational school and work settings. In S. Volet & S. Järvelä (Eds.), Motivation in learning contexts: Theoretical advances and methodological implications (pp. 201–232). London: Elsevier.

Kuyvenhoven, J. C. (2005). In the presence of each other: A pedagogy of storytelling.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*.

MacLeod, Douglas, Larissa Muller, David Covo and Richard Levy. 2007. *Design as an instrument of public policy in Singapore and South Korea*. Vancouver, BC: Asia Pacific Foundation of Canada.

McDrury, J., & Alterio, M. (2003). *Learning through storytelling in higher education*, Kogan Sterling, VA: Page Limited.
Mcquiggan, S. W., Rowe, J. P., Lee, S., & Lester, J. C. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In Intelligent Tutoring Systems (pp. 530-539). Springer Berlin Heidelberg.
Mercer, N., Warwick, P., Kershner, R., & Staarman, J. (2010). Can the interactive whiteboard help to provide 'dialogic space' for children's collaborative activity? Language and Education, 24(5), 367–384.
Mercier, E., & Higgins, S. (2014). Creating joint representations of collaborative problem solving with multi‐touch technology. *Journal of Computer Assisted Learning*.
National Research Council (NRC). (2006). *Taking Science to School: Learning and Teaching Science in Grades K-8*. Washington, D.C.: National Academics Press.
Oksanen, K., & Ståhle, P. (2013). Physical environment as a source for innovation: investigating the attributes of innovative space. *Journal of Knowledge Management*, *17*(6), 815-827.
Oppl, S., & Stary, C. (2014). Facilitating shared understanding of work situations using a tangible tabletop interface. *Behaviour & Information Technology*, *33*(6), 619-635.
Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International journal of science education*, *25*(9), 1049-1079.
Piper, A. M., & Hollan, J. D. (2009, April). Tabletop displays for small group study: affordances of paper and digital materials. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1227-1236). ACM.
Renninger, A., Hidi, S., & Krapp, A. (Eds.). (2014). *The role of interest in learning and development*. Psychology Press.
Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J.M. Harackiewicz (Eds.), Intrinsic and extrinsic motivation: The search for optimal motivation and performance (pp. 375–407). New York: Academic.
Renninger, K. A., & Shumar, W. (2002). Community building with and for teachers: TheMath Forum as a resource for teacher professional development.In K. A. Renninger&W. Shumar (Eds.), Building virtual communities: Learning and change in cyberspace (pp. 60–95). New York: Cambridge University Press.
Rojas-Drummond, S., Mazón, N., Fernández, M., & Wegerif, R. (2006). Explicit reasoning, creativity and co-construction in primary school children's collaborative activities. *Thinking Skills and Creativity*, *1*(2), 84-94.
Sandoval, J. (1995). Teaching in subject matter areas: Science. Annual Review of Psychology, 46(1), 355-374.
Scott, P., Mortimer, E., & Ametller, J. (2011). Pedagogical link-making: a fundamental aspect of teaching and learning scientific conceptual knowledge. *Studies in Science Education*, *47*(1), 3-36.
Selwyn, N. (2011). *Education and technology: Key issues and debates*. Bloomsbury Publishing.
Shalley, C. E., Zhou, J., & Oldham, G. R. 2004. The effects of personal and contextual characteristics on creativity: Where should we go from here? *Journal of Management,* 30: 933–958.
Shelby, A., & Ernst, K. (2013). Story and science: How providers and parents can utilize storytelling to combat anti-vaccine misinformation. Human vaccines & immunotherapeutics, 9(8), 1795-1801.
Smogorzewska, J. (2014). Developing children's language creativity through telling stories–An experimental study. Thinking Skills and Creativity, 13, 20-31.
Stocklmayer, S., & Gilbert, J. K. (2002). New experiences and old knowledge: Towards a model for the personal awareness of science and technology. *International Journal of Science Education, 24*(8), 835-858.
Sundholm, H., Artman, H., & Ramberg, R. (2004). Backdoor Creativity: Collaborative Creativity in Technology Supported Teams. In *COOP* (pp. 99-114).
Tan, E. S. (2013). *Emotion and the structure of narrative film: Film as an emotion machine*. Routledge.
Tobias, S. (1992) . *Interest and metacognition in mathematics.* Paper presented at a symposium on "Interest, Attention, and Knowledge Acquisition" at the Annual Meeting of the American Educational Research Association, San Francisco.
Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, *64*(1), 37-54.
Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education*, *18*(2), 345-368.
Watson, D. M. (2001). Pedagogy before technology: Re-thinking the relationship between ICT and teaching. *Education and Information technologies*, *6*(4), 251-266.
Wood, R. & Ashfield, J. (2008). The use of the interactive whiteboard for creative teaching and learning in literacy and mathematics: a case study, British Journal of Educational Technology39(1): 84–96.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Development and Evaluation of an Educational Game- *Shimmer*© with Computer Visualization for Optics Learning

Huei-Tse HOU[a*], Shu-Ming WANG[b], Hsin-Hung YU[a] & Shie-He LIU[a]
[a]*Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan*
[b]*Department of Information Management, Chinese Culture University, Taiwan*
[*]hthou@mail.ntust.edu.tw

Abstract: Learning physics involves the reasoning of abstract, dynamic, and complex that cannot be mapped to learners' daily experience or observed with naked eyes. with the advancement of computer technologies, nowadays, computer technologies, such as simulation and scientific visualization, are able to depict the scientific phenomena with vivid visualization and rich interaction. Recently, in addition to provide simulation and scientific visualization, many educational games were developed to support learning of various subjects in order to promote the level of students' engagement in learning. Educational games with clear goals, immediate feedback, and adequate scaffolding were considered as a powerful means of improving students' learning motivation and subsequent learning performance. Though game-based learning has been a trend in recent years, researchers suggested that research on the application of game-based learning to STEM is still relatively limited. To fill the literature gap, this study developed a game – *Shimmer*© to support the learning of optical concepts. This study conducted an experiment with 50 seventh-grade students. An analysis of pretest and posttest was conducted to examine the effects of the game on learning. In addition, the participants' evaluation of game and their gaming experience were assessed via using flow constructs. Furthermore, individual differences were addressed as well. Results showed that the participants generally better capture the concepts of light reflection and refraction in visual way. Besides, further analyses of individual differences indicated that boys and girls as well as students with high and low prior knowledge evaluated the game differently. Implications for the results of this study are to be used as guidance for subsequent game development and design of instructional strategies.

Keywords: Game-based learning, STEM, Optics, flow

1. Introduction

Scientific learning involves numerous abstract concepts, which cannot be directly observed or mapped to learners' day-to-day experience (Anderson & Barnett, 2013). For example, the concepts of optics and wave cannot be directly observed without adequate instruments (Hamam, 2003). When learning these concepts, it would be difficult for learners to capture and integrate the abstract, complex, and discrete ideas to form correct conception of a particular phenomenon. Sometimes, learners would even form misconceptions if improperly guided (Anderson & Barnett, 2013; Masson, Bub, & Lalonde, 2011). Nonetheless, with the advancement of computer technologies, nowadays, these abstract, dynamic, and complex phenomena can be depicted by using computer simulation or scientific visualization techniques to improve learners' understanding (Chen, Pan, Sung, & Chang, 2013; Van Dam, Forsberg, Laidlaw, LaViola, & Simpson, 2000). Previous studies have adopted technologies, e.g., virtual reality or interactive computer simulation, to support the teaching of abstract concepts, such as electrostatics (Anderson & Barnett, 2013), object motion (Masson et al., 2011), and optics (Mzoughi, Herring, Foley, Morris, & Gilbert, 2007).

As to learning optics, it could be difficult for learners to catch the concepts of optics, as it is invisible to naked eyes (Mzoughi et al., 2007; Van Dam et al., 2000). Though the lab experiment could be helpful to improve students' understanding, it could be complex and costly for setting up a lab experiment as the class is getting larger. This situation makes computer visualization be an ideal tool to support the learning of the concepts of optics (Mzoughi et al., 2007; Van Dam et al., 2000). In response to this issue, Mzoughi et al. (2007) developed a 3D interactive computer graphics system – WebTOP for teaching and learning optics. WebTOP adopted the features of visual simulation to illustrate the optical concepts, such as reflection and refraction, lasers and scattering etc., allowing that learners better understand these phenomena. The system can support both classroom use and self-guided study and received generally positive affirmation from both educators and students. However, one general constraint for computer simulation is that users might be lack of motivation to engage in the repetitive process or they could easily get bored. This situation could diminish the educational benefits of using computer simulation for teaching and learning.

To improve users' motivation and level of engagement in learning, employing an educational game to support teaching and learning has been a trend in educational practices nowadays (Anderson & Barnett, 2013). An educational game adopts game mechanism in addition to computer simulation and visualization. Like computer simulation, computer games allow players to interact with the content and receive immediate feedback, which can be used to adjust their action or conception accordingly. Well-designed computer games provide clear goals with different levels of challenge that require players to explore the means to achieve. Rewards are provided as players reached the goals afterward. This process provokes intriguing experience that promotes players' motivation to and engagement in gaming (Mayo, 2009; Prensky, 2001; Young et al., 2012). In the educational context, games can be adapted to individual learners' pace and scaffold them through the learning process. For example, complex learning tasks can be decomposed to smaller tasks at the beginning of a game. As the level of challenge gradually gets easier, players would have opportunity to practice simple tasks several times before they deal with more complex learning tasks (Mayo, 2009). Through the repetitive practice, learners are expected to attain the mastery of particular cognitive skills (Prensky, 2001).

Witnessing the surging popularity of games and advantages of adopting games for learning, many educational games have been developed in recent years. However, there is relatively limited research on game-based STEM (science, technology, engineering, and mathematics) in the past decade (Young et al., 2012). In order to fill the literature gap, this study developed an educational game – *Shimmer©* to support the learning of optics and conducted a preliminary study to improve our knowledge about adopting game to support learning. The description of *Shimmer©* is presented in Research Method section.

To explore why and how people play games, previous game-related studies have used the idea of flow from psychology literature to assess the optimal gaming experience (e.g. (Barzilai & Blau, 2014; Hsu & Lu, 2004; Kiili, 2006). Flow experience refers to the perception of being totally absorbed to the activity in which people engaged. People would feel a sense of enjoyment, distortion of time, loss of self-conscious, and intrinsically motivated when they were in flow state (Csikszentmihalyi, 1994). Flow is therefore adopted as an important indicator of intriguing experience when playing games. Nonetheless, there are prerequisite conditions for the occurrence of flow experience. One essential element of flow is that the level of challenge has to match the player's level of skills. Otherwise, the player could get anxious when the level of challenge is significantly higher; on the contrary, the player would get bored when the level of skills is significantly higher. Moreover, people need to feel in control of what they are doing and feel that they are pursuing a clear goal (Kiili, 2006). This study adopted flow experience to assess the participants' evaluation of *Shimmer©* and their gaming experience. Besides, the effects of individual differences have been regarded as influential factors in educational research (Li, Cheng, & Liu, 2013; Yukselturk & Bulut, 2009). This study further looked into the plausible effects of individual differences on participants' learning outcomes, evaluation of game, and gaming experiences. Concluding from the above, the purposes of this preliminary study are summarized as follows:

1. To examine the difference of learning outcomes before and after playing the game - *Shimmer©* .
2. To explore the plausible differences of game evaluation and gaming experience of *Shimmer©* for students with different gender.

3. To explore the plausible differences of game evaluation and gaming experience of *Shimmer©* for students with different level of prior knowledge.

The initial findings of this preliminary study could help us better understand the effectiveness of *Shimmer©* , which visualizes the abstract and dynamic concepts of optics. These findings would also serve as a guideline for improving the future version of *Shimmer©* . Moreover, based upon the findings, suggestions for incorporating scientific visualization with game-based learning were purposed. The following sections are to delineate the introduction of game and the experimental design.

2. Research method

2.1 Procedure and Participants

This educational game used in this study was a 3D simulation game - *Shimmer©* , which was developed via Unity3D game engine. In the game, the player needs to protect civilians in a small village from the invaders. In each stage, there were a laser cannon, several plane mirrors and lenses, and one invader sat on a floating island as shown in figure 1. By clicking on the mirrors and lenses, players could calibrate their angle and location. Once the players finished the calibration, they could click on the laser cannon to launch a laser and check how it traveled. The goal of the game was to correctly set up the angle and location of the lenses so that the laser could shoot the invader. At current version of the game, players were allowed for unlimited trials. The concepts of optics introduced in the current version of *Shimmer©* are about reflection and refraction. At the beginning stages, there were plane mirrors only. Players could simply calibrate the angle of plane mirrors to achieve the goal. These stages were designed to help players to be acquainted with the game and learn the basic ideas of light reflection. At the further stages, convex and concave lenses as well as obstacles sited on the path of laser transmission were added, which would gradually increase the difficulty level of the game (as shown in figure 2). There were 15 levels of game challenges in current version of *Shimmer©* . A non-player character (NPC) - the Sage was created to provide relevant knowledge of light refraction. Players could click on the Sage to read the needed information at any time during the game. Screenshot of the provided information is shown in figure 3.

Figure 1: Game screenshot of the beginning stage.

Figure 2: Game screenshot of the advanced stage.

Figure 3: In-game optics knowledge guidance

The participants of this study were 50 seventh-grade students in a junior high school in northern Taiwan, including 28 males and 22 females. The experiment was conducted in a PC classroom. Before the experiment, students were asked to fill a pretest to assess their prior knowledge of light fraction. Students have to complete the pretest in seven minutes. The pretest was followed by a five-minute introduction of the game. Afterward, students were to freely play the game for 20 minutes. Finally, students were asked to fill a posttest as well as a flow instrument. The total time of the experiment was around 45 minutes.

2.2 Instruments

This study developed a test, which was used in the pretest and posttest session. The test included 8 multiple-choice questions, which were primarily adapted from the textbook, and 4 questions that asked students to draw the direction of light transmitted through the lenses. The example of drawing question is as shown as figure 4. In this example, the arrow showed the direction of the light. Participants were asked to draw the travel path of the light. Students were given one point for each correct answer.

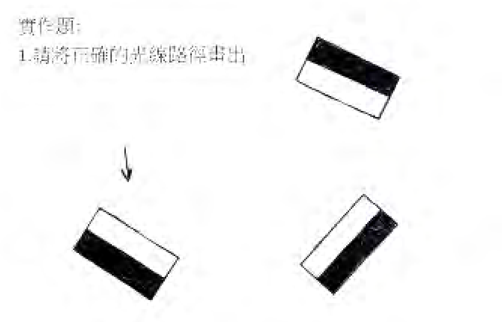


Figure 4: Example of a drawing question

The flow instrument was adapted from Kiili (2006). 22 items were used to assess two major components of the flow construct. The first component is flow antecedents, which represents players' evaluation of the game in terms of the level of challenge, goal, feedback, control, and playability. The second component is flow experience, which reflects players' perception of concentration, time distortion, autotelic experience, and loss of self-consciousness when playing game. The flow instrument was measured using five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree).

3. Data Analysis and Results

3.1 Learning outcomes

SPSS 20.0 was used to analyze the collected data. Table 1 summarized the participants' scores in pretest and posttest as well as the results of mean difference test. As shown in Table 1, overall, there was a significant difference between the pretest and posttest. However, further analysis revealed that there was no difference in the scores of multiple-choice. Nonetheless, after playing the game, students' performance of drawing question was significantly higher than before ($t = -5.14$, $p < 0.001$).

Table 1: Summary of students' performance in pretest and posttest.

Type of questions	Pretest (n=50)		Posttest (n=50)		t-statistics	p-value
	Mean	S.D.	Mean	S.D.		
Multiple-choice	4.38	1.92	4.46	2.11	-0.29	0.773
Drawing	0.46	0.68	1.18	1.02	-5.14	.000
Overall	4.84	2.24	5.64	2.58	-2.53	0.015

For the flow construct, the Cronbach's α of flow instrument was 0.971, suggesting high reliability (Nunnally, 1978). The participants generally possess positive evaluation of the game. Specifically speaking, the participants evaluated *Shimmer©* as challenging (mean = 3.64), with clear goals (mean = 4.10), providing prompt feedback (mean = 3.71), controllable (mean = 3.64), and playable (mean = 3.52). Meanwhile, all the means of sub-constructs of flow experience exceeded 3.5, suggesting that the participants generally have positive flow experience while playing game.

3.2 Individual differences

Individual difference was considered as an important factor in educational research (Price, 2006; Yukselturk & Bulut, 2009). In the context of gaming, the stereotypical view of males is that they might be more interested and engaged in playing video game than females (Chumbley & Griffiths, 2006; Yee, 2006). In addition, when playing game, males and females might behave differently. For example, girls would be more focus on causal interaction with other players, whereas boys would tend to focus on game-related conversation (Young et al., 2012). These differences would affect their evaluation of an educational game. In this regard, this study conducted a further analysis to explore the potential differences between boys and girls in evaluation of *Shimmer©* in terms of flow antecedents. The results showed that boys, in contrast to girls, showed higher evaluation of game challenge ($t = 2.96$, $p < 0.05$), feedback ($t = 3.62$, $p < 0.01$), and playability ($t = 2.72$, $p < 0.05$). However, there were no differences observed in their evaluation of game goal ($t = 1.69$, $p > 0.05$) and control ($t = 1.855$, $p > 0.05$). Regarding the flow experiences, boys reported higher degree of flow experience in concentration ($t = 2.22$, $p < 0.05$), autotelic experience ($t = 2.78$, $p < 0.05$), and loss of self-consciousness ($t = 2.51$, $p < 0.05$). Nonetheless, the difference in time distortion was not observed ($t = 1.71$, $p > 0.05$).

To further explore the plausible differences in students' flow experience, this study conduct a test for difference between the means of students with high and low prior knowledge. Based on their scores in the pretest of this study, students whose scores were above the top 27% were categorized into high prior knowledge group (N=14), whereas those who below the bottom 27% were categorized into low prior knowledge group (N=14). The results indicted students with high prior knowledge reported higher degree of experiencing time distortion than those with low prior knowledge ($t = 2.28$, $p < 0.05$). However, there were no differences observed in other flow dimensions. The interpretation and discussion of these results are to be presented in the following section.

4. Conclusion and subsequent research

With the technology advancement, modern computer games are able to simulate the scientific phenomena and create an immersing environment that follows the natural laws of physics for educational purposes (Anderson & Barnett, 2013). With this affordance, computer games have been used to support learning of various subjects. In particular, computer games are commonly used to support the learning of abstract concepts or complex procedures (Kebritchi & Hirumi, 2008). This study developed an educational game – *Shimmer©* to support the learning of optical concepts. The initial findings suggested that *Shimmer©* was helpful to improve students' conceptions of reflection and refraction, particularly in visual depiction. Participants generally reported positive flow experience, which was considered as an essential component of intriguing gaming experience, when playing *Shimmer©* . The discussion and implications of findings are delineated as below.

First of all, the game – *Shimmer©* helped the participants better capture the ideas of reflection and refraction. In particular, after playing the game, the participants were more able to correctly draw the path that light travels among mirrors, convex and concave lenses than before. *Shimmer©* allowed players to form their hypotheses of how light travels and then set up the angle and location of lenses to freely test and alter their hypotheses for unlimited times. This process could enhance students' autonomy in learning, which could promote their learning motivation and thus achieve better learning outcomes (Ryan, Rigby, & Przybylski, 2006; Vogel et al., 2006). Nonetheless, there was no significant difference in multiple-choice questions between the pretest and posttest. One plausible explanation may result from the form of problem representation in the game. In *Shimmer©* , students need to calibrate the lenses in order to make the light hit the target. In the process, they were able to observe the results of light transmitted through lenses of different angles and locations. Therefore, they would form the conceptions of reflection and refraction in visual way. *Shimmer©* didn't require players to memorize factual knowledge to achieve the game goal. In other words, students needn't to answer the traditional form of test when playing game. Thus, their performance on multiple-choice questions could be limited. Similar notions have been proposed by previous research. Masson et al. (2011) suggested that the engagement benefits of games are suited for training particular cognitive skills. Nonetheless, this immersing environment and repetitive process might make learners be less accomplished with regard to traditional test performance (Young et al., 2012).

Secondly, analysis of gender difference showed that boys tend to have higher evaluation in most dimensions of flow antecedents. As boys were regarded as typical gamers, they might be more able to adapt to game control than girls are (Chumbley & Griffiths, 2006; Yee, 2006). However, lacking of control could impede players from immersing in the game play, which in turn diminishes the level of engagement and the effects of game-based learning (Hou, Wang, & Tsai, 2013; Scoresby & Shelton, 2011). Nonetheless, in this study, there were no differences in boys and girls' perceptions of game goal and controllability. This may result from the game goal of *Shimmer©* , whose goal was to shoot the invader with a laser cannon, which was simple and clear enough for both boy and girl participants. Moreover, the participants only need to click on the lenses to calibrate the angle and location of them. No complicated operations were required. Regarding the flow experiences, boys seemed being more absorbed into the game as they generally reported higher degree of flow experience in concentration, autotelic experience, and loss of self-consciousness. However, the difference in sense of time distortion between boys and girls was not observed. This non-significant result could be attributed to the limited time of the game session (20 minutes) and a small sample size. Subsequent research is suggested to extend the game session and conduct a larger scale of experimentation to further investigate the gender differences in flow experience.

Lastly, regarding to individual differences, though the participants with high and low prior knowledge generally reported positive flow experience (i. e. mean of all sub-dimensions exceeded 3.5), results indicated that participants with higher prior knowledge tend to have significantly higher sense of time distortion. It seemed that they were more engaged in playing game. This finding was probably because students with higher prior knowledge were able to capture the ideas in game. Thus, they would be more focused on solving the problem than those with lower prior knowledge were. This finding also emphasized the importance of the match between the level of challenge in game and the skills of players, which is a prerequisite condition for the occurrence of flow experience (Csikszentmihalyi, 1994). Thus, it is suggested that the design of an educational game should consider the prior knowledge of learners in order to induce the optimal gaming experience while playing (Barzilai & Blau, 2014; Kiili, 2006).

5. Research Limitations and subsequent study

As a preliminary study, this study was limited in the time of game session and sample size; thus, these initial findings should be interpreted with cautions. First, the short game session could be a limitation for the students to properly evaluate the game and to have flow experience. In this manner, the longer game sessions are needed to further investigate the effects when using the game – *Shimmer©* to support the learning optical concepts. Secondly, an elaborated experimental design with more participants, which involves the control group provided for other instructional practice such as simulation, would help us better compare the effectiveness of game-based learning with that of traditional practice. Lastly, the advantages of game-based learning are not merely to promote learners' engagement. Mayo (2009) suggested using in-game assessment to track the sequences of players' behaviors, such as the number of attempts to solve a problem, the timing of seeking online help, or interaction among players. This approach can help instructors to further look into the learning process and provide feedback for game developers to improve the game design. In this manner, the subsequent research is suggested to employ sequential analysis to explore players' gaming patterns, which would help us better understand how players interact with the game and improve the future version of *Shimmer©* .

Acknowledgements

This research was supported by the projects from the National Science Council, Republic of China, under contract number MOST-102-2511-S-011-001-MY3, MOST-100-2628-S-011-001-MY4, MOST-103-2511-S-034-001 and NSC-102-2511-S-034-001.

References

Anderson, J., & Barnett, M. (2013). Learning physics with digital game simulations in middle school science. *Journal of Science Education and Technology, 22*(6), 914-926

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education, 70*(0), 65-79

Chen, Y.-L., Pan, P.-R., Sung, Y.-T., & Chang, K.-E. (2013). Correcting Misconceptions on Electronics: Effects of a simulation-based learning environment backed by a conceptual change model. *Educational Technology & Society, 16*(2), 212-227.

Chumbley, J., & Griffiths, M. (2006). Affect and the computer game player: The effect of gender, personality, and game reinforcement structure on affective responses to computer game-play. *CyberPsychology & Behavior, 9*(3), 308-316.

Csikszentmihalyi, M. (1994). *Flow: The Psychology of Optimal Experience.* New York, NY: Simon & Schuster Audio.

Hamam, H. (2003). An interactive web-based tool for teaching optics and vision. *Optics and Photonics News, 14*(8), 30-34.

Hou, H.-T., Wang, S.-M., & Tsai, D.-S. (2013). *The development and evaluation of a 3D simulation game for chemistry learning: exploration of learners' flow, acceptance, and sense of directions*. Paper presented at the The 21st International Conference on Computers in Education (ICCE 2013), Bali, Indonesia.

Hsu, C.-L., & Lu, H.-P. (2004). Why do people play on-line games? An extended TAM with social influences and flow experience. *Information & Management, 41*(7), 853-868.

Kebritchi, M., & Hirumi, A. C. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers & Education, 51*(4), 1729-1743

Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environment, 2*(2), 187-201.

Li, Z.-Z., Cheng, Y.-B., & Liu, C.-C. (2013). A constructionism framework for designing game-like learning systems: Its effect on different learners. *British Journal of Educational Technology, 44*(2), 208-224.

Masson, M. E. J., Bub, D. N., & Lalonde, C. E. (2011). Video-game training and naïve reasoning about object motion. *Applied Cognitive Psychology, 25*(1), 166-173.

Mayo, M. J. (2009). Video Games: A route to large-scale STEM education? *Science, 323*(5910), 79-82.

Mzoughi, T., Herring, S. D., Foley, J. T., Morris, M. J., & Gilbert, P. J. (2007). WebTOP: A 3D interactive system for teaching and learning optics. *Computers & Education, 49*(1), 110-129.

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.

Prensky, M. (2001). *Digital Game-based Learning*. New York: McGraw-Hill.

Price, L. (2006). Gender differences and similarities in online courses: Challenging stereotypical views of women. *Journal of Computer Assisted Learning, 22*(5), 349-359.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion, 30*(4), 344-360.

Scoresby, J., & Shelton, B. (2011). Visual perspectives within educational computer games: Effects on presence and flow within virtual immersive learning environments. *Instructional Science, 39*(3), 227-254.

Van Dam, A., Forsberg, A., Laidlaw, D. H., LaViola, J. J., & Simpson, R. M. (2000). Immersive VR for scientific visualization: a progress report. *Computer Graphics and Applications, IEEE, 20*(6), 26-52.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of educational computing research, 34*(3), 229-243.

Yee, N. (2006). Motivations for play in online games. *CyberPsychology & Behavior, 9*(6), 772-775.

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., . . . Yukhymenko, M. (2012). Our Princess Is in Another Castle: A review of trends in serious gaming for education. *Review of Educational Research, 82*(1), 61-89.

Yukselturk, E., & Bulut, S. (2009). Gender differences in self-regulated online learning environment. *Journal of Educational Technology & Society, 12*(3), 12-22.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

A Three-Stage Augmented-Reality-Facilitated Earth Science Instructional Process for Dispersing Learning Style Differences

Chang-Hwa WANG[a], Cheng-Ping CHEN[b]*
[a]*Department of Graphic Arts and Communications, National Taiwan Normal University, Taiwan*
[b]*Department of Visual Communication Design, Taipei College of Maritime Technology, Taiwan*
*chenjp0820@yahoo.com.tw

Abstract: Studies have proven that merging hands-on and online learning can result in an enhanced learning experience. However, the effects of learning styles have substantially affected online learning performance. As ICT continues to develop, an augmented reality (AR)-embedded instructional orientation could provide additional hands-on experiences to the classroom. In contrast to traditional online learning, multiple in-classroom activities may be involved in an AR-embedded e-learning process, thus could reduce the effects of individual differences. Using a three-stage AR-embedded comprehensive instructional process, an experiment was conducted to investigate the influences of student's learning styles. The results of the study showed that overall learning achievement was significant for the AR-embedded instruction. Nevertheless, as no significant difference found among different learning styles, indicating that our multiple activities oriented AR learning process may have helped disperse the effect of different learning styles.

Keywords: augmented-reality-facilitated learning, learning style, science learning, interactive learning environment

1. Introduction

Some researchers asserted that students learn more effectively with e-learning environments because students like interactive learning that provided by recent interactive technologies (Lee, Choi, & Park, 2009; Hatziapostolou, & I Paraskakis, 2010; Ali Karime, Hossain, A. S. M. M. Rahman, Gueaieb, Alja'am, & El Saddik, 2012). Hrastinski (2009) indicated if learner has an opportunity to control their learning environment, they would have more interest and willing to learn in classes. In an interactive e-learning environment, students would become more positive and active.

Augmented reality (AR) is one of such interactive technologies. It mixes virtual and real world by means of displaying virtual objects onto real images in accordance with target triggers (markers) that manipulating by users. In addition to visualization, users can interact with virtual objects (Chehimi, Coulton, & Edwards, 2007). Many studies revealed that AR systems have educational values because students enjoyed the interaction with virtual objects which is also effective to improve students' learning performance. Among various interactive technologies, interactions with AR were found to be particular helpful for learning spatial concepts (Kaufmann & Schmalstieg, 2003; Kirner, Zorzal & Kirner, 2006; Juan, Beatrice, & Cano, 2008). Researches have also indicated that learning style is important in laying the groundwork for understanding students' learning performance, especially for e-learning, in which learner characteristics is necessarily adapted to the interactive instruction (Huang, Lin, & Huang, 2012). Related works have been done with learning styles in relation to learner's participation, learning quality, and performance of e-learning (Shaw, 2012; Marković & Jovanović, 2011). The results of these researches exhibited there are significant relationship between learning styles and e-learning outcomes in general. Nevertheless, none of these discussions went beyond the realm of traditional online e-learning.

The adaptation of individual differences to e-learning has been discussed for several decades. Johansen and Tennyson asserted that adaptive advisement could help students in perceive knowledge in learner-controlled, computer-based instruction (Johansen & Tennyson, 1993). Magnisalis, Demetriadis, and Karakostas (2011) Claimed that Artificial Intelligence and Web 2.0 techniques could support collaborative learning in an online learning environment. However, all these techniques involve complicated computer algorithm and likely need to comply learning activities within a computer monitor, therefore lack of direct human interactions. Unlike traditional online learning, multiple activities may be involved in an AR-facilitate e-learning process. An AR-facilitate e-learning is able to take place in regular classroom settings, instructional activities may include life lectures, student manipulations of virtual objects, peer discussions, and even written exercises after experiencing the AR. These multiple activities could be more adaptive for learners with different learning styles. However, little research has been done on this issue.

The present study assumed that a comprehensive AR-facilitate learning process consists of various types of learning activities, therefore there will be less effects of learning styles on learners achievements. A quasi-experiment was performed to examine the effects of learning styles on learning achievement while the comprehensive AR- facilitate learning process was given. In the present study, a comprehensive AR-facilitate learning process includes lecture, hands-on AR experience, peer discussions, and written exercises.

2. Related Work

2.1 AR-facilitate Instruction

Azuma (2009) first recognized the AR as a technique that link between the real and virtual world. Yuen, Yaoyuneyong, and Johnson (2011) gave an up-to-date definition to AR. In their definition, AR has three distinctive characteristics: (a) it is the combination of real world and virtual elements, (b) it is interactive in real-time, and (c) it is registered in three dimensions. Thus, AR has some potential to influence instruction and learn knowledge from different fields.

Several researches have used AR systems in education, including mathematics, science, language, and medicine. For example, in their experiment, Kirner, Zorzal and Kirner (2006), a “Game of Word” used plates containing symbols of English alphabets, when setting up a word completed by the plates in front of the webcam, the related virtual object appears over it. They believed that this game was able to motivate the users to interact and create solutions in an attractive AR environment. In addition, Juan, Beatrice and Cano (2008) presented an AR system for learning the interior of the human body. Learners were able to “open” the abdomen of a virtual human body using their own hands. Learners also saw inside the human body virtually, and observed the areas where the stomach and the intestine are located. More recently, Matsutomo, Miyauchi, Noguchi, and Yamashita (2012) created a real-time visualization system, which can visualize a composite image of source materials and their generated magnetic field utilizing the AR technique. They claimed that with such a system, electromagnetics learners can observe the magnetic distribution in a virtual real-time manner.

More AR-facilitate science learning researches have been done in this decade. Recent discussions of instructional applications of AR have gone beyond the effectiveness of the AR per se. For example, in order to better understand the effective strategies that are appropriate for AR-facilitate learning Yoon, Elinich, Wang, Steinmeier, and Tucker (2012) compared four conditions for learning science in a science museum using AR and knowledge-building scaffolds. Results indicated that students demonstrated greater cognitive gains when scaffolds were used. The limitation of above research findings is that they only viewed AR learning as a standalone activity and fall short to vision the entire AR-facilitate learning process, including lecture, peer discussions, and other classroom activities, as a whole. Wang and Chi (2012) demonstrated a comprehensive AR-facilitate learning process, emphasizing in-classroom interactions, to teach fundamental earth science for junior highs. The entire learning process included teacher’s lecture, hands-on AR experiences, peer discussions, and written exercises. They thought that AR-facilitated instruction could improve the understanding of spatial concepts and be easier to acquire the course contents. Nevertheless, the differences between individual students were not discussed in this study. They suggested that further research on individual differences, for example, the learning styles is necessary.

A Synthetic work of AR research was done by Bujak, Radu, Catrambone, MacIntyre, Zheng, and Golubski (2013). They reviewed recent research on AR learning. They highlighted the potential benefits and limitations of using AR to deliver learning experiences, by presenting an analysis based on psychological constructs, and by comparing AR applications to physical and virtual manipulatives. They concluded that although AR shows great promise for extending the resources used for educating our students, there is much research to be done. Finally they suggested that researchers must more specifically address the usefulness of AR from a psychological perspective.

2.2 Learning Styles

Viewing from the psychological perspective, a line of research has found that learner characteristics had great effects on learning performance. Lamia and Mohamed Tayeb (2013) recognized that learning styles, thinking styles, and levels of knowledge and abilities are key learner characteristic that affects the successfulness of an e-learning. Among these learner characteristics, learning style is an key indicator of how a student learns and likes to learn, and how an instructor teaches to successfully address the needs of the individual students (Chang, Kao, Chu, & Chiu, 2009; Tseng, Chu, Hwang, & Tsai, 2008).

Learning style is a distinctive and habitual manner of acquiring knowledge, skills or attitudes through study or experience while learning preference is favoring of one particular mode of teaching over another (Marković, & Jovanović, 2011). There are a lot of learning style models developed in past fifty years. Witkin, Oltman, Raskin & Karp (1971) first systematically used a Group Embedded Figures Test to identify field independence of the learner. Kolb (1984) employed The Learning Style Inventory as the instrument to classify learner into four categories as convergent learners, divergent learners, assimilators, and accommodators. Keefe (1987) developed a learning style test. It can identify learners into four skill categories: Sequential Processing Skill, Discrimination Skill, Analytic Skill and Spatial Skill. Felder & Silverman’s (1988) model, however, comprises the category of intuitive/sensitive, global/sequential, visual/verbal, inductive/deductive and active/reflective, which can be used to discriminate 32 learning styles. Finally, Fleming defines learning style as “an individual’s characteristics and preferred ways of gathering, organizing, and thinking about information. VARK is in the category of instructional preference because it deals with perceptual modes (Marković, & Jovanović, 2011). The acronym VARK stands for Visual (V), Aural (A), Read/Write (R), and Kinesthetic (K).

Fleming (2012) further explained that life is multimodal. There are seldom instances where one mode is sufficient to describe complicated learner characteristics. For those who do not have a standout mode with one preference score well above other scores resulting from the VARK questionnaire, are defined as multimodal. Therefore he categorized VARK into fifteen learning style within three modes. The three modes are: single mode, dual-mode, and multimode. Fifteen learner styles are then categorized into three modes, they are: V, A, R, and K, for single mode, VA, VR, VK, AR, AK, and RK for dual-mode, finally, VAR, VAK, VRK, ARK, and VARK for multimode.

Huang, Lin, and Huang (2012) criticized that several studies investigated the relationship between learning style and performance most have adopted a dichotomous definition of learning style that does not offer sufficient information for an in-depth investigation of the relationship. The VARK learning style model, however, is among the few that allow categorizes learner into bi/multi-learning style modes. This unique characteristics of VARK classification scheme is particular suitable for the present study.

3. Method

3.1 Research Goal and Questions

The goal of the study was to examine whether a comprehensive AR-facilitate learning process, including lecture, hands-on AR experience, peer discussions, and written exercises, would have extensive adaptions of different learner styles. Based on the assumption that there will be less effects of learning styles on learners achievements while AR-facilitate learning process applied, under the VARK learning style classification scheme, the following research questions were issued:

1. Is there a significant effect of learning style, in terms of single, dual, or multi- mode, on learning achievement while AR-facilitate learning process applied?

2. Is there significant differences among the four single-mode learning styles (V, A, R, and K) on learning achievement while AR-facilitate learning process applied?
3. Is there significant differences among the six dual-mode learning styles (VA, VR, VK, AR, AK, and RK) on learning achievement while AR-facilitate learning process applied?
4. Is there significant differences among the four multi-mode learning styles (VAR, VAK, ARK, and VARK) on learning achievement while AR-facilitate learning process applied?

3.2 The AR Learning Kit

The AR learning kit used for the experiment consists of three components: a sun/earth relation turntable, a computer with screen, and a webcam that captures the birds-eye-view of the turntable. This AR learning kit is able to display a day/night sensitive map and a schematic diagram simulating the changes of pole shadows. The trigger image is the earth on a turntable that simulates the revolution of earth around the sun. There are three images synchronously display on a computer screen. Image on top of the screen displays the overlay images of the earth and the sun. Image on the bottom left side is the day/night sensitive map and image on the bottom right side is the pole shadow schematic diagram. Students are able to turn the earth on the turntable manually and three images will simultaneously simulate the situations in accordance with the date and time displayed on the most top of the screen. Please refer to Figure 1 for the actual orientation of the AR learning kit.

Figure 1. The Orientation of AR Learning Kit.

3.3 Lesson Plan

As AR learning were found to be particular helpful for learning spatial concepts, earth science phenomena: “Day, night, and seasons” that involved spatial orientations of was selected as the learning content. Specific learning objectives were designed as: 1) to understand the interchange of day and night is affected by the earth rotation; 2) to understand the alternation of four seasons is affected by the revolution of the earth; 3) to understand the position and the length of pole shadows are affected by the rotation and revolution of the earth, respectively. Anderson and Krathwohl’s (2001) revision of Bloom's taxonomy of educational objectives was employed as the guideline for constructing the instructional content. Knowledge, understanding, application, and analysis were taken as the four dimensions for instructional content. In knowledge dimension, the phenomena of rotation and revolution of the earth, the interchange of day and night, as well as the move of pole shadow are described and demonstrated. In understanding dimension, the reasons why the phenomena happen are described. In application dimension, students are required to operate the AR learning kit, observe the relationship between the move on the turntable and the coordinate change on AR displays on screen. Finally, in analysis dimension, exercises are provided to allow students use acquired knowledge to analyze given situations and resolve problems.

A three-stage AR instructional process was designed to bring about the learning objectives. The three stages are: gaining attention stage, learning with AR stage, and summarization stage. In the gaining attention stage, the instructor gave lecture on these natural phenomena and showed examples of the relationships between seasons and length of a day as well as the changes of the pole shadows. In the learning with AR stage, students were divided into groups by five, each group operated the AR-learning kit (as shown on graph 1 below). An assistant was assigned to each group to help the operation. These assistants also raised questions after hands-on AR experience for initializing within group peer discussions. Finally, in the summarization stage, students were obtained reinforcements of learned concept. A reflective sheet that required students to answer fill-in-blank questions by operating the AR learning kit was given. A summary of lesson plan for the experiment, including learning activities, amount of time spent, and the learning materials for each stage are shown on table 1.

Table 1: Summary of the Lesson Plan.

Stage	Learning activities	Time	Material
Gaining attention	Instructor gives lecture and shows examples,	25 min.	PowerPoint briefings
Learning with AR	Students operate the AR-learning kit in groups. Peer discussions	50 min.	AR learning kit
Summarization	Students obtain reinforcement of learned concepts	15 min.	Reflective sheet (written form)

3.4 The Instrument

Fleming's VARK learning style questionnaire employed to classify the multiple-tendency of learning styles [www.business.vark-learn.com]. The Younger Version revised on September, 2007 was used to fit the age range (13-15 years old) of our subjects. They are totally 16 questions in the questionnaire. Four selections of possible answers are available for each question. Each answer refers to one of the VARK category. Multiple selections of these answers are allowed. Totally 15 categories can be classified by this questionnaire. The reliability of the Chinese version of the questionnaire is Cronbach's α= .83.

A pretest and a posttest for evaluating the learning achievement were constructed. The revised taxonomy aforementioned was taken as the guideline for constructing test items. Both pretest and posttest have 18 single-answer multiple choice questions. The knowledge and understanding dimensions have four questions, and application and analysis dimensions have 5 questions. The pretest and posttest are designed as parallel forms. The reliability of the pretest and posttest are α= .74 and .84, respectively.

3.5 The Experiment

The experiment was done in two separate junior high schools in New Taipei City of Taiwan in a period of three months. Totally 144 students in five seven-grade classes were selected as subjects. One instructor and seven teaching assistances were involved in the instructional process. An independent three-stage AR instructional process was performed for each of the five classes. A pretest was given before, and a posttest was given after each instructional process was performed. Several students were randomly asked for a brief interview to understand student's interests and motivations on using the AR learning kit.

4. Results and Discussions

4.1 The Distribution of Learning Styles

According to Fleming's VARK classification scheme, students were categorized into 15 learning styles within single, dual, and multiple learning modes. The most prevalent learning style type was "VARK". 35 out of 144 students fell into this category. "V" and "VK" were the two categories that had least students fit in. In terms of learning style mode, most students had multi-modal learning style (80), and least students were single-modal (26). Detailed information please refer to Table 2.

Table 2: Distribution of Learning Mode and Learning Style.

Learning Mode	Learning Style	Number of Samples	
Single-modal	V	2	26
	A	11	
	R	4	

	K	9	
Dual-modal	VA	10	38
	VR	3	
	VK	2	
	AR	9	
	AK	11	
	RK	3	
Multi-modal	VAR	19	80
	VAK	15	
	VRK	4	
	ARK	7	
	VARK	35	
		144	144

4.2 The Effects of Learning Style Mode

The effects of the three modes of learning style: single, dual, and multiple modes on learning achievement were statistically analyzed. An ANCOVA was performed to examine the significance of the mean differences among the three modes. Pretest score was used as the covariance. The result indicated that there was no significant difference on learning achievement found among these three modes ($F_{2,140}$= .017, p=.983). As we also found that there is a significant pretest-posttest gains (t_{286}=10.346, p<.001), it was evident that the three-stage AR-facilitate learning process was adaptive for learners with any learning mode. Table 2 summarizes the ANCOVA.

Table 3: Summary of ANCOVA for Learning Style Mode.

Source	SS	df	MS	*F*	η^2	*p*
Mode	.327	2	.163	.017	.000	.983
Error	1336.534	140	9.547			

4.3 The Effects of Learning Styles On Achievement

We further examined the effects of learning style types on student's learning achievement. Three separate ANCOVA was performed for the three learning style modes.

There are four types of single-modal learning styles: "V", "A", "R", and "K". The result of ANCONA indicated a near significant result ($F_{3,21}$ =2.581, p=.081). In order to avoid possible type II error, LSD post hoc comparisons was done. The result of the post hoc analyses indicated that the achievement for "R" type learners was significantly better than "A" type (p=.030) and "K" type (p=.012). This result seems to be incoherent with Fleming's account that "R" type of learner prefers to use text-based materials. The possible reason is that the summarization stage helped more for "R" type of students in reviewing the concepts learned during the AR operation. Table 3 summarizes the ANCOVA and the LSD comparisons for the single-modal learning styles.

Table 4: Summary of ANCOVA and Post Hoc For Single-Modal Styles.

Source	df	*F*	η^2	*p*	Post hoc
Learning style	3	2.581	.269	.081	R > A, R > K
Error	21				

The differences among six types of dual-modal learning styles: "VA", "VR", "VK", "AR", "AK", and "RK". The result of ANCOVA showed a non-significant result ($F_{5,31}$=1.000, p=.434), there was no differences on learning achievement regarding dual-modal learning styles. Table 4 summarizes the ANCOVA for the dual-modal learning styles.

Table 5: Summary of ANCOVA for Dual-Modal Styles.

Source	SS	df	MS	*F*	η^2	*p*
Style	44.383	5	8.877	1.000	.139	.434
Error	275.140	31	8.875			

Finally, we examined the effects of multi-modal learning styles. Again, the result of ANCOVA exhibited that was no significant difference among "VAR", "VAK", "VRK" "ARK", and "VARK" types of learners. Table 5 summarizes the ANCOVA for the multi-modal learning styles.

Table 6: Summary of ANCOVA for Dual-Modal Styles.

Source	SS	df	MS	*F*	η^2	*p*
Style	85.252	4	21.313	1.858	.074	.127
Error	860.298	75	11.471			

5. Conclusion

An adaptive education combines the development of an individual's initial competence with alternative environments matched to different styles of learning. The Adaptation of individual differences to e-learning has been discussed for decades. However, most efforts have been made to develop intelligent programs to select appropriate instructional paths and/or to determine the amount of instruction to be given based on individual learner's on-task performance. To avoid complicated adaptive computer algorithms, we developed a comprehensive AR-facilitate e-learning process that is able to take place in regular classroom settings. Instructional activities included life lectures, student manipulations of virtual objects, peer discussions, and written exercises. An experiment employing the AR process was done with an earth science learning unit. The VARK learning style classification scheme was used. The results showed there was no significant difference in learning achievement of students with different mode of learning styles.

This result is promising. The present study provides an alternative rationale for developing adaptive e-learning without involving complicated adaptive algorithm. A comprehensive AR-facilitate e-learning process could make the regular classroom to be more adaptive to students with different learning styles.

Although findings of this study are potentially supporting the development of an alternative ICT-based e-learning strategy, some inherent limitations must be addressed. As we admitted the non-significant results related to the differences among learning styles, there were some risks of gaining type II errors. Although most of the non-significant decisions were made on a reasonably reliable basis ($p > .3$), sparse significant results were found in post hoc analyses of single-modal learning styles (R > A, R > K). Another limitation of the study is that the generalizability of the research findings is restrained because only a single learning unit within a single subject matter was implemented for the experiment. For future studies, we suggested that larger sample sizes and extensive subject matters need to be concerned.

Acknowledgements

Funding of this research work is supported in part by the Ministry of Science and Technology of Taiwan (under research numbers NSC 102-2511-S-003-023- and MOST 103-2420-H-003-031-).

References

Anderson, L. & Krathwohl, D. A. (2001). A *taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives.* New York, NY: Longman.

Azuma, R. T. (1997). A survey of augmented reality. *Presence, 6*(4), 335-385.

Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., & Zheng, R. (2013). A psychological perspective on augmented reality in the mathematics classroom, *Computers & Education, 68,* 536-544.

Chang,Y. -C., Kao, W. -Y., Chu, C. -P., & Chiu, C. -H.(2009). A learning style classification mechanism for e-learning. *Computers & Education*, *53*(2), 273-285.

Chehimi, F., Coulton, P., & Edwards, R. (2007). Augmented reality 3d interactive advertisements on smartphones.

In *Proceedings of the 2007 IEEE International Conference on the Management of Mobile Business* (pp.21-28).

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering education*, *78*(7), 674-681.

Fleming, N. D.(2012). *VARK: A guide to learning styles.* Retrieved from: http://www.vark-learn.com/english/index.asp.

Hatziapostolou, T. & Paraskakis, I. (2010). Enhancing the impact of formative feedback on student learning through an online feedback system. *Electronic Journal of E-Learning*, *8*(2), 111–122.

Hrastinski, S. (2009). A theory of online learning as online participation. *Computers & Education*, *52*(1), 78-82.

Huang, E. Y., Lin, S. W., & Huang, T. K. (2012). What type of learning style leads to online participation in the mixed-mode e-learning environment? A study of software usage instruction. *Computers & Education*, *58*(1), 338-349.

Juan, C., Beatrice, F., & Cano, J. (2008). An augmented reality system for learning the interior of the human body. In P. Diaz, Kinshuk, I. Aedo, & E. Mora (Eds.), *Proceedings* of *the 8th IEEE International Conference on Advanced Learning Technologies* (pp. 186-188).

Karime, A., Hossain, M. A., Rahman, A. S. M. M., Gueaieb, W., Alja'am, J. M., & El Saddik, A. (2012). RFID-based interactive multimedia system for the children. *Multimedia Tools and Applications*, *59*(3), 749–774.

Kaufmann, H. & Schmalstieg, D. (2003). Mathematics and geometry education with collaborative augmented reality. *Computers & graphics, 27*(3), 339-345.

Keefe, J. W. (1987). *Learning style: Theory and Practice.* Reston, VA: National Association of Secondary School Principals.

Kirner,C., Zorzal, E. R. & Kirner, T. G. (2006). Case studies on the development of games using augmented reality. In *Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics*, *2* (pp. 1636-1641).

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Lamia, M., & Tayeb, L. M. (2013). Discovering learner styles in adaptive e-learning hypermedia systems. *Journal of Universal Computer Science*, *19*(11), 1522-1542.

Lee, S. -H., Choi, J., & Park, J. -I. (2009). Interactive e-learning system using pattern recognition and augmented reality. *IEEE Transactions on Consumer Electronics*, *55*(2), 883–890.

Magnisalis, I. Demetriadis, S. & Karakostas, A. (2011). Adaptive and intelligent systems for collaborative learning support: A review of the field. *IEEE Transactions on Learning Technologies, 4*(1), 5-20.

Marković, S. & Jovanović,N. (2011). Learning style as a factor which affects the quality of e-learning. *Artificial Intelligent Review*, *38*(4), 303-312.

Matsutomo, S., Miyauchi, T., Noguchi, S. & Yamashita, H. (2012). Real-time visualization system of magnetic field utilizing augmented reality technology for education. *IEEE Transactions on Magnetics*, *48*(2), 531-534.

Shaw, R. -S. (2012). A study of the relationships among learning styles, participation types, and performance in programming language learning supported by online forums. *Computers & Education*, *58*(1), 111-120.

Tseng, C. -R., Chu, H. -C., Hwang, G. -J. & Tsai, C. -C. (2008), Development of an adaptive learning system with two sources of personalization information. *Computers and Education*, 51(2), 776-786.

Wang, C. -H., & Chi, P. -H. (2012). Applying augmented reality in teaching fundamental earth science in junior high schools. *Communications in Computer and Information Science, 352*, 23-30.

Witkin, H., Oltman,P., Raskin, E. & Karp, S. (1971). *A manual for the group embedded figures test*. Palo Alto, CA: Psychologists Press.

Yoon, S. A., Elinich, K., Wang, J., Steinmeier, C. & Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning*, *7*(4), 519-541.

Yuen, S., Yaoyuneyong, G., & Johnson, E. (2011). Augmented reality: An overview and five directions for AR in education. *Journal of Educational Technology Development and Exchange*, *4*(1), 119–140.

The Effects of AR-based Instruction on Students' Learning Performance, Motivation and Self-efficacy in Programming Learning

Gloria Yi-Ming KAO* & Cheng-An RUAN
Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan
*gloriakao@gmail.com

Abstract: Recent studies have reported benefits of Augmented Reality (AR)-based instruction in various learning domain. However, few studies were done to explore its effects in programming learning. In this study, we devolved an AR-based instruction with high (puzzle cards) or low (fixed card) interaction levels to assist programming learning. The results showed that students' learning performance and self-efficacy were improved after the experiment, which indicated AR-based learning did have positive effects though no difference between high and low interaction could be determined at this stage.

Keywords: augmented reality, programming learning, learning performance, motivation, self-efficacy

1. Introduction

1.1 Augmented Reality (AR)

Augmented Reality (AR) is a technique that adds virtual objects in a real environment, which can redeem the lack of information in a real environment (El Sayed, Zayed, & Sharawy, 2010). Physical and virtual tools are commonly utilized as assistant teaching tools in classrooms; however, it is hard to combine the strengths of the two tools. After AR was invented, which integrated physical and virtual functions, learners were supplied with better learning experiences (Bujak et al., 2013). AR enables learners get close to real environments from learning environments and supplies richer sensory experiences. Also, it makes learners have the opportunities to operate physical objects and then to interact with virtual ones (Wojciechowski & Cellary, 2013).

1.2 Programming Learning

Programming is a big role in science, technology, engineering and mathematics (STEM) fields. Especially, to those major in computer science, programming course plays an important role. Nevertheless, learning programming is difficult to the beginners who are not in the related fields (McCracken et al., 2001). Despite of learners' ages, programming is always difficult for beginners (Kelleher & Pausch, 2005). In Taiwan, traditional way to learn programming usually use textbooks or run sample program codes with computers, and then observe the execution results. However, beginners are not able to realize the programming process or the results. This can lead to poor learning motivation or performance.

1.3 AR in Education

Recently, many studies have shown that AR-based learning has positive influence on students' learning, such as learning performance (Lin, Duh, Li, Wang, & Tsai, 2013), learning motivation (Di Serio, Ibáñez, & Kloos, 2013) and self-efficacy (Kamarainen et al., 2013).

Therefore, AR is commonly used as an educational tool. Liu and Tsai (2013) employed teaching materials combined with AR in an English writing course, trying to reduce the difficulties of the students while learning second language. Results indicated that learners under this situation could construct contents and knowledge much easier; moreover, more meaningful articles were produced, and language learning performance was then increased. Chang, Chang, Sung, Chao and Lee (2014) developed an AR guide system and then used it in an art appreciation course. Compared to general audio guide and non-guide environment, learners in AR guide system group had more fluent experience and better learning performance. Chang, Wu and Hsu (2013) stimulated Fukushima nuclear disaster and explored the situation of nuclear pollution by using AR technique. The result show that AR based environment could improve students' comprehension and increase their sensorial immersion. Ibáñez, Di Serio, Villarán and Delgado Kloos (2014) implemented AR in a basic course of electromagnetism. The results showed that, in this environment, students could not only understand the phenomena and concepts of electromagnetism more efficiently but reach higher flow experience levels, compared to web-based learning environment. AR is used in many different subjects, yet cases of programming course are seldom found. Consequently, this study explores how AR-based learning influence students' learning performance and motivation in a programming course.

2. System design

2.1 Software Development

Aurasma, a cross-platform AR development system developed by Hewlett-Packard Development Company in 2011, supplies simple operating interface to make it easier for developers to produce AR contents. Aurasma enables users to connect to the database to get the latest AR contents by using the mobile device.

Based on the Aurasma, we devolved an AR-based application with high (puzzle cards) or low (fixed card) interaction levels to assist programming learning. Students could freely assemble puzzle cards to observe the corresponding results. With high interaction level, students are encouraged to try and figure out what combinations could work or not. With low interaction level, students could only watch default animations with fixed cards. We used pictures from the Scratch program, developed by Massachusetts Institute of Technology in 2003, in our design.

2.2 Operating process

Learners followed the teacher instruction, assembled different puzzle cards (as Figure 1 shows), and used the Aurasma in mobile device to observe and compare the differences among various combinations of puzzle cards. While the puzzle cards are assembled correctly, the programming operating animation will be displayed (as Figure 2 shows). However, the animation will not be displayed when the assembling is not correct.

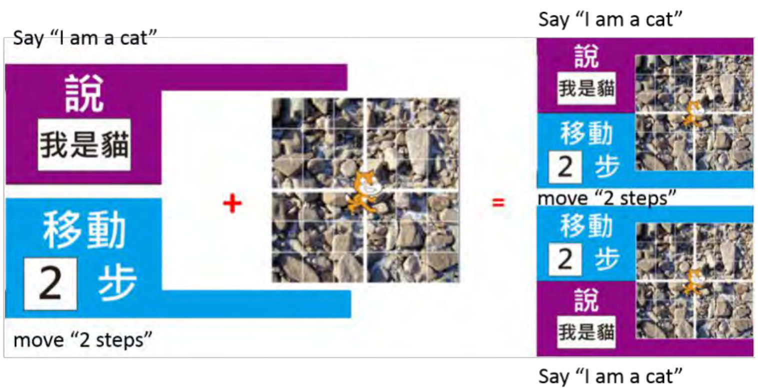


Figure 1. Assemble different puzzles.

Figure 2. The animation is displayed when the assembling is correct.

2.3 Learning Content

The learning content of the study was three main structures of flowchart (as Figure 3 shows) in computer programming. They were 1) Sequence structure, which followed certain order, operating the description separately; 2) Selection structure, which operates the program according to the conditional determination; 3) Iteration structure, which operates the descriptions repeatedly, until the descriptions match the breaking condition, the operation ends. Most Algorithms can be consisted of these three structures. This course is not only an important role but an essential part in programming courses. Several combinations of puzzles were supplied. Puzzles can be assembled with two combinations in Sequence and selection structures, but they can be assembled with three combinations in Iteration structure.

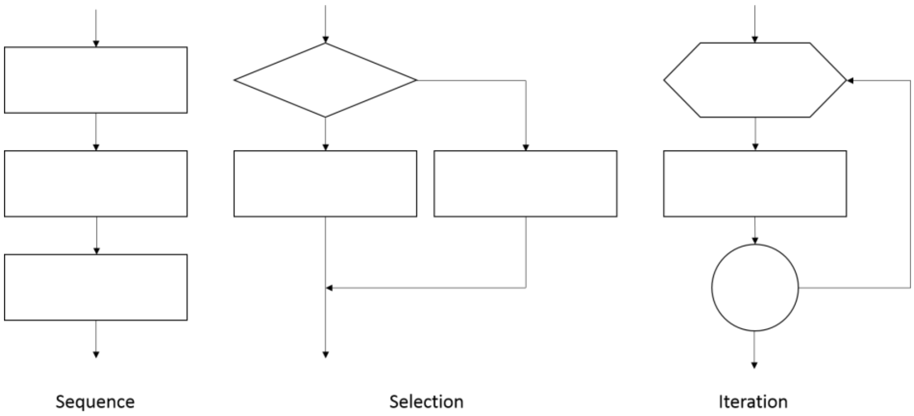


Figure 3. Three main structures of flowchart.

3. Experiment design

3.1 Participants

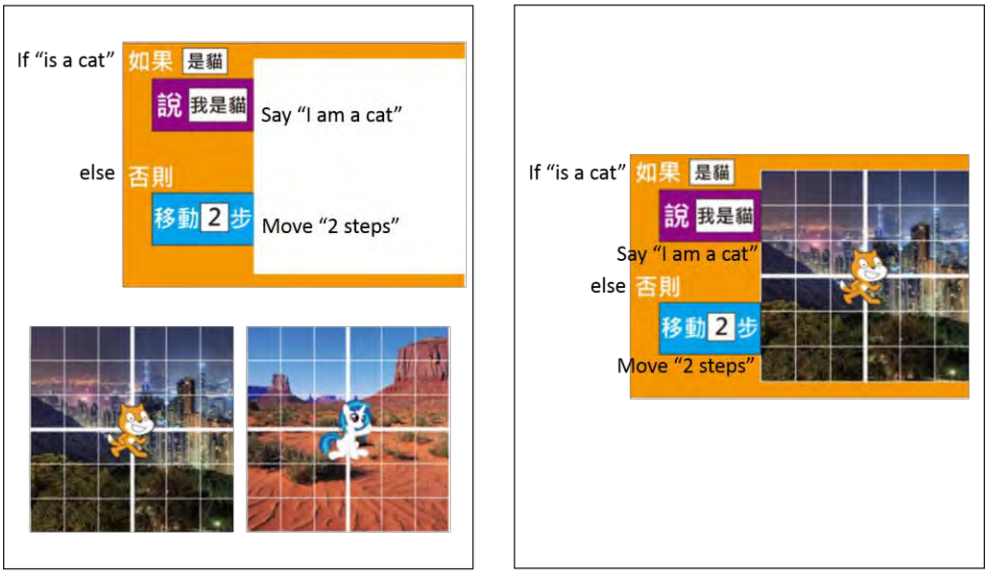
Sixty-five seventh graders from two mixed-ability classes in a junior high school in North Taiwan participated in the experiment, the mean age was 13. Thirty-one (15 females and 16 males) were in the experimental group, while thirty-four students (16 females and 18 males) were in the control group. High interactive AR-based learning systems (puzzle cards) were used in the experimental group, while low interactive AR-based learning system (fixed cards) was in the control group.

Mobile devices were not used in previous teaching process, so the students were instructed to operate the mobile devices and the AR system first before the experiment was carried out. In both of the two groups, there were 7 students had learned Scratch, a software which can program interactive animations. Their average learning time was one year. Therefore, a pretest was adopted to exclude the differences among the students caused by prior knowledge.

3.2 Learning activities and environment

In order to fit in the teacher’s original teaching style, the experiments were taken place in a computer classroom. During the experiment, computers were only used for broadcasting PowerPoint slides for instructional purpose. For both the experimental and control groups, each two students were equipped with one mobile device (iPad) and were allowed to discuss with each other.

As Figure 4 shows, high interactive AR-based learning systems were used in the experimental group. This group was given puzzle cards and was asked to freely assemble those puzzles to yield and observe different program execution outcomes. While students in the control group with low interactive AR-based learning system could only get fixed cards and observe the default program execution results.



Experiment group　　Control group

Figure 4. Experiment group with puzzle cards and control group with fixed cards.

3.3 Experiment procedure

Figure 5 shows the four stages of this study and describe as below:

- Stage 1: A pretest was adopted to examine the students’ prior knowledge about programming; pre-questionnaires were employed to probe students’ learning motivation and self-efficacy. Students were firstly explained the ways to fill out the sheets and were not allowed to discuss during the test. This stage took ten minutes.
- Stage 2: With one mobile device equipped, two students grouped and were taught how to operate the mobile devices and the learning system. This stage took ten minutes. The researcher checked the students’ status to make sure they were able to operate the tools.
- Stage 3: Programming course was started. There were four main units in this course. Students observed and learned with mobile devices after the teacher’s instruction. This stage took sixty minutes.
- Stage 4: A learning performance posttest, a learning motivation post-questionnaire and self-efficacy post-questionnaire were adopted. Students were not allowed to discuss during the test. This stage took twenty minutes.

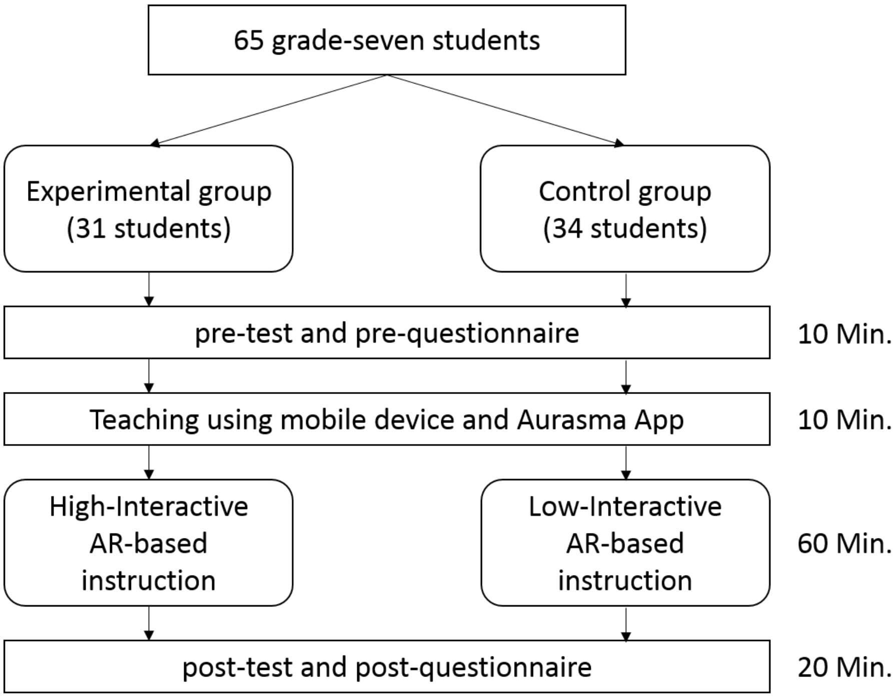


Figure 5. Experiment procedure.

4. Analysis methods and results

4.1 Analysis tools

Learning performance test: A self-edited performance test sheet was used, and the expert validity was constructed after the scale was revised by three computer and information science teachers. A pretest was employed to examine the students' prior knowledge about programming. Each of the ten matching questions in the sheet was ten points. A post-test was adopted to explore students' knowledge and applied abilities about programming. There were five matching questions and five applicant questions, and each question was ten points. From this test, we wanted to understand if the learning performance was improved after the experiment.

Learning motivation and self-efficacy: The learning motivation questionnaire adopted in the study was a revision from Hwang, Yang and Wang (2013) and included seven questions. The self-efficacy questionnaire was a revision from Wang and Hwang (2012) which included eight questions. Both the questionnaire adopts five-point Liker rating scheme, "1" means very disagree while "5" means very agree. All the descriptions in the questionnaire were positive (e.g. I think this course is meaningful and worthy learning, I believe that I can get great score in the assignments). For the purpose to explore the difference before and after the intervention, the two questionnaires were employed both before and after the experiment to investigate students' perceptions toward learning motivation and self-efficacy. The original Cronbach's alpha of the motivation questionnaire was 0.79, while the revised one was 0.883. The original Cronbach's alpha of the self-efficacy questionnaire was 0.916, while the revised one was 0.919.

4.2 Method

The present study analyzed and processed the data by using statistical software. Firstly, the descriptive statistic, mean and standard deviation, would show the differences among learning performance, learning motivation and self-efficacy. Secondly, after pre-test and pre-questionnaire, independent t-Test will be employed to examine if there are any differences between the two groups. Thirdly, ANCOVA will be used to explore if there are any differences in post-test and post-questionnaire between the two groups. Above three steps will be adopted to observe how high and low interactions influence AR-based instruction. Lastly, how AR-based instruction influence programming learning will be discussed. Also, dependent t-Test will be used to examine whether there are any significant differences between pre-and-post test, and pre-and-post questionnaire.

4.3 Results

Several independent t-Tests were conducted to examine the difference between experimental and control group before the experiment. As shown in Table 1, there is no significant difference between experimental and control group (pre-test: t=-0.403, $p>0.05$, pre-motivation: t=1.103, $p>0.05$, pre-self-efficacy: t=0.136, $p>0.05$) which means before the learning activity, students in the two groups have equivalent prior knowledge, learning motivation and self-efficacy.

Table 1: Independent t-Test result of Pre-test, Pre-motivation and Pre-self-efficacy of Experimental group and Control group.

		N	Mean	SD	t	Sig.
Pre-test	Experimental group	31	47.097	19.008	-0.403	0.689
	Control group	34	48.824	15.524		
Pre-motivation	Experimental group	31	4.143	0.655	1.103	0.274
	Control group	34	3.975	0.574		
Pre-self-efficacy	Experimental group	31	3.347	0.948	0.136	0.892
	Control group	34	3.320	0.582		

After the learning activity, several analysis of covariance (ANCOVA) were used to evaluate the difference between experimental and control group in terms of learning performance, motivation or self-efficacy. The pre-test score of each evaluation was used as covariate while the post-test score of each evaluation was used as dependent variable. Table 2 shows that the results are not significant (post-test: F=0.082, $p>0.05$, post-motivation: F=1.216, $p>0.05$, post-self-efficacy: F=0.142, $p>0.05$) which means high or low interactive AR-based leaning would not impact on students' learning performance, motivation or self-efficacy.

Table 2: ANCOVA result of Post-test, Post-motivation and Post-self-efficacy of Experimental group and Control group.

		N	Mean	SD	F	Sig.
Post-test	Experimental group	31	69.839	19.701	0.082	0.775
	Control group	34	68.618	17.132		
Post-motivation	Experimental group	31	4.014	0.654	1.216	0.274
	Control group	34	3.870	0.554		
Post-self-efficacy	Experimental group	31	3.581	0.935	0.142	0.707
	Control group	34	3.522	0.607		

Several dependent t-Tests were conducted on exploring how AR-based learning influences students' learning performance, motivation and self-efficacy, as shown in Table 3. There is a significant difference between learning performance (t=-8.634, $p<0.05$) and self-efficacy (t=-2.557, $p<0.05$), yet learning motivation is not significantly different (t=1.559, $p>0.05$). The results showed overall after the learning activity, students' learning performance and self-efficacy were increased significantly, although no difference between the puzzle cards and fixed cards could be determined.

Table 3: Dependent t-Test result of Pre-test, Post-test, Pre-motivation, Post-motivation, Pre-self-efficacy and Post-self-efficacy.

	N	Mean	SD	t
Pre-test	65	48.000	17.157	-8.634***
Post-test	65	69.200	18.266	
Pre-motivation	65	4.055	0.615	1.559
Post-motivation	65	3.938	0.603	

Pre-self-efficacy	65	3.332	0.772	-2.557*
Post-self-efficacy	65	3.550	0.775	

***p<.001, *p<.05

5. Conclusions and discussion

The present study explores how high and low interactive AR-based learning influence students' learning. As the results showed, different interactive levels with puzzle or fixed cards did not impact on students' learning performance, motivation and self-efficacy. Instead, overall students' learning performance and self-efficacy were improved after the experiment, which indicated the innovative use of AR into programming learning did have positive effects although no difference between high and low interaction could be determined at this stage.

Convenience sampling was conducted in this study, which was a limitation of the study. The results can only represent the learning performance of the students in the experiment, but cannot infer the overall students' in other areas. Further, time duration was another limitation of the study. Owing to the time, the states of students' learning and how much did they exactly learned were not thoroughly considered.

Future studies are suggested dividing the participant groups into high-interactive AR, low-interactive AR, and traditional learning to further investigate how AR-based learning influences students' learning.

Acknowledgements

This research was funded by the National Science Council of the Republic of China, Taiwan, under contract number NSC 102-2511-S-011 -005 -MY3.

References

Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., Zheng, R., & Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. Computers & Education, 68, 536-544.

Chang, K. E., Chang, C. T., Hou, H. T., Sung, Y. T., Chao, H. L., & Lee, C. M. (2014). Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum.Computers & Education, 71, 185-197.

Chang, H. Y., Wu, H. K., & Hsu, Y. S. (2013). Integrating a mobile augmented reality activity to contextualize student learning of a socioscientific issue.British Journal of Educational Technology, 44(3), E95-E99.

Di Serio, Á., Ibáñez, M. B., & Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. Computers & Education, 68, 586-596.

El Sayed, N. A., Zayed, H. H., & Sharawy, M. I. (2011). ARSC: Augmented reality student card. Computers & Education, 56(4), 1045-1061.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. Computers & Education, 69, 121-130.

Ibáñez, M. B., Di Serio, Á., Villarán, D., & Delgado Kloos, C. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. Computers & Education, 71, 1-13.

Kamarainen, A. M., Metcalf, S., Grotzer, T., Browne, A., Mazzuca, D., Tutwiler, M. S., & Dede, C. (2013). EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. Computers & Education, 68, 545-556.

Kelleher, C., & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers.ACM Computing Surveys (CSUR), 37(2), 83-137.

Lin, T. J., Duh, H. B. L., Li, N., Wang, H. Y., & Tsai, C. C. (2013). An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. Computers & Education, 68, 314-321.

Liu, P. H. E., & Tsai, M. K. (2013). Using augmented- reality- based mobile learning material in EFL English composition: An exploratory case study.British Journal of Educational Technology, 44(1), E1-E4.

McCracken, M., Almstrum, V., Diaz, D., Guzdial, M., Hagan, D., Kolikant, Y. B. D., ... & Wilusz, T. (2001). A multi-national, multi-institutional study of assessment of programming skills of first-year CS students. ACM SIGCSE Bulletin, 33(4), 125-180.
Wang, S. L., & Hwang, G. J. (2012). The role of collective efficacy, cognitive quality, and task cohesion in computer-supported collaborative learning (CSCL). Computers & Education, 58(2), 679-687.
Wojciechowski, R., & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. Computers & Education, 68, 570-585.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Implementation of Student-associated Game-based Open Inquiry in Chemistry Education: Results on Students' Perception and Motivation

Keeratika MEESUK[a], **Niwat SRISAWASDI**[b*]

[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*

[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*

**niwsri@kku.ac.th*

Abstract: Educational computer game refers to the use of digital technology to promote learning performance and experience through game-based activity. Currently, researchers mentioned that digital game-based learning can promote students' interest and enhance learning outcomes. As such, this study aims to develop inquiry-based learning process with the support of digital game for chemistry learning about ionization energy. This paper reports research findings on two studies. First, a digital game regarding chemistry concept of ionization energy has been created and 29 twelfth-grade students participated to play the game. The students' perceptions towards the game were measured and results showed that the game have significant effect in fostering their perceptions. Another, 87 tenth-grade students were recruited in this study, and they were divided into an experimental group (N=43) and a control group (N=44). The experimental group participated with student-associated game-based open inquiry, called SAGOI, class and another were assigned to conventional class. The results indicated that students in SAGOI class have changed science motivation over the SAGOI learning experience. This could be implied that the student-associated game-based open inquiry could be an alternative way for promoting chemistry learning in school science.

Keywords: Digital game, inquiry-based learning, ionization energy, science motivation

1. Introduction

Edutainment is an educational concept which aims to make a combination between education and entertainment. As such concept, educational computer game could be considered as an alternative form of edutainment that learners can learn lessons of subject matter by playing the game. By the way, researchers mentioned that educational game is different from other edutainment by its nature in requiring strategies, proving hypothesis, solving problems, and it usually use higher-order thinking rather than memorization. The main characteristics of educational game are the challenging to achieve the objectives, and providing a specific situation and reward for engaging and motivating learners which act as the players (Prensky, 2001; Papastergiou, 2009). A recent study reported about how digital game support learners' motivational and cognitive processing. Huang (2011) indicated that learners have more confident in learning after playing with educational game. In an addition, using game in education increased students' perceived learning, enjoyment and flow of learning experience (Barzilai and Blau, 2014).

For chemistry class, not many study investigated influence of educational game on students' learning outcomes. In fact, chemistry knowledge is often abstract, complex, and complicate in representations of the chemistry knowledge. The use of digital game may be help student increasing learning interest in chemistry, motivation, and attitude towards chemistry learning. Therefore, this study aims to create educational computer game as an inquiry tool to learn chemistry in concepts of ionization energy.

2. Research Questions

Based on the above mentioned rationale, the goals of this study was a couple: (i) to investigate students' perceptions towards an educational computer game, called "The IE War"; and (ii) to investigate students' science motivation delivered in a proposed open-inquiry learning process with support by an educational computer game, called "Student-associated Game-based Open Inquiry (SAGOI)". Specifically, the following questions were answered:

- Do the students interacted with "The IE War" perform significantly better by perceptual constructs i.e. perceived learning, flow, enjoyment, and satisfaction?
- Do the students engaged in "SAGOI" perform significantly better by motivational constructs towards science learning i.e. intrinsic motivation, self-determination, self-efficacy, career motivation, and grade motivation?

3. Backgrounds and Research Hypotheses

2.1 Educational Computer Game and Students' Perceptions

2.2.1 Perceived Learning

Perceived learning relates to a retrospective evaluation of the learning experience and can be defined as a set of beliefs and feelings one has regarding the learning that has occurred (Caspi and Blau, 2011). The perceived learning is about the new information was obtained and person can get the new understanding, subjective evaluation of learning by learners themselves. Researchers mentioned that perceived learning is connected to emotion as flow, enjoyable, and satisfaction (Chu and Hwang, 2010). Regarding in context of educational computer game, when learners are immersed in game-based learning environment, they can judge themselves in the learning process and quality of how to get the knowledge from game, so game can help learned and practiced (Giannakos, 2013). According to the abovementioned, a research hypothesis was set to examine in this study as following.

H1. Students' perceived learning between pre- and post-interacting with educational computer game has development.

2.1.2 Flow

Flow is a state of deep concentration in which thoughts, intentions, feelings, and all of the senses are focused on the same goal (Csikszentmihalyi, 1990; Barzilai and Blau, 2014). The experience of flow would happen when person who take part in challenge situations or activities that need skills. Flow depends on a chance to concentrate, an immediate feedback, a sense of control, and a clarify goal (Barzilai and Blau, 2014). As such, if learners concentrate with the learning experience of educational computer game, the flow of learning would occur during playing the game. According to the abovementioned, a research hypothesis was set to examine presented as following.

H2. Students flow between pre- and post-interacting with educational computer game has development.

2.1.3 Enjoyment

Enjoyment is the condition of having and using technology, e.g. educational computer game that is good or pleasant. The enjoyment of player is a key goal, related with an easy to use of game and enjoyment was found to have valuable in explaining objective to use applications (Giannakos, 2013).

When learners which act as players of game fail to pass the game task, they would get disappointment and attempt to replay again. As such, the enjoyment can help reduce worry to learn and feel more confident when learners success. Accordingly, if the educational computer game can enjoy learners, then they would like to learn more and think positive to the subject. Based on this aspect, below is a research hypothesis for examining the impact of educational computer game on enjoyment.

H3. Students' enjoyment between pre- and post-interacting with the educational computer game has development.

2.1.4 Satisfaction

Satisfaction is the individual awareness of how well a learning environment supports academic success (Lo, 2010). It is relevant to instructional method that learners can think and learn, so their satisfaction can help to get how academic success. At the moment to learn with educational computer game, if it gets positive response from learners that means they reach to positive learning experience with also. In an addition, satisfaction can yield positive of learning performance and can improve learning outcome (Giannakos, 2013). To investigate an impact of educational computer game, a hypothesis was set as the following.

H4. Students' satisfaction between pre- and post-interacting with the educational computer game has development.

2.2 Student-associated Game-based Open Inquiry and Science Motivation

Science motivation refers to the motivation of students to learn science within their emotional which stimulate, control and support in science learning behavior. Therefore, science motivation could be achieved to learners when activate their behaviors by asking the questions, doing experiments, and collaborative learning (Schunk, Pintrich and Meece, 2008; Glynn et al., 2011). Researchers stated that science motivation consists of five motivational constructs: *intrinsic motivation*, an internal state of satisfaction to learn science because it will be good for itself; *self-determination*, the controlling of students' belief that they have when learning science; *self-efficacy*, students can bring their belief connect and manage to achieve learning science; *career motivation*, students learn science to get a good work in the future; and *grade motivation*, learning science to have a good score (Glynn et al., 2011). The following research hypothesis was another one which the researchers expected to examine in this study.

H5. Students' science motivation between pre- and post-participating in the student-associated game-based open inquiry has development.

4. Research Design

This study used two different research designs: one-group pretest-posttest design to examine impacts of the proposed digital game, The IE War, on students' perception and a simple two-group comparison to examine the effects of an instructional intervention, The SAGOI, on students' science motivation, as study 1 and 2 respectively.

5. Study 1

5.1 Participants

The subjects of this study were 29 twelfth-grade students in a public school at the northeastern region of Thailand. The age range of the students was 17-18 years, and all of them were females. They were attending a chemistry course for basic education level and all of them have satisfactory skills on basic computer and information and communication technology, but they have no experience yet using educational computer game in chemistry learning.

5.2 Learning Materials

The digital game on chemistry of ionization energy was designed to comprise three stages and ten levels of playing. The IE War is the first stage and it was created in style of shooting game. The goal of this educational game is to facilitate student getting the definition of ionization energy. To complete The IE War, there were five playing levels as followings. Figure 1 illustrates a flow chart of The IE War game.

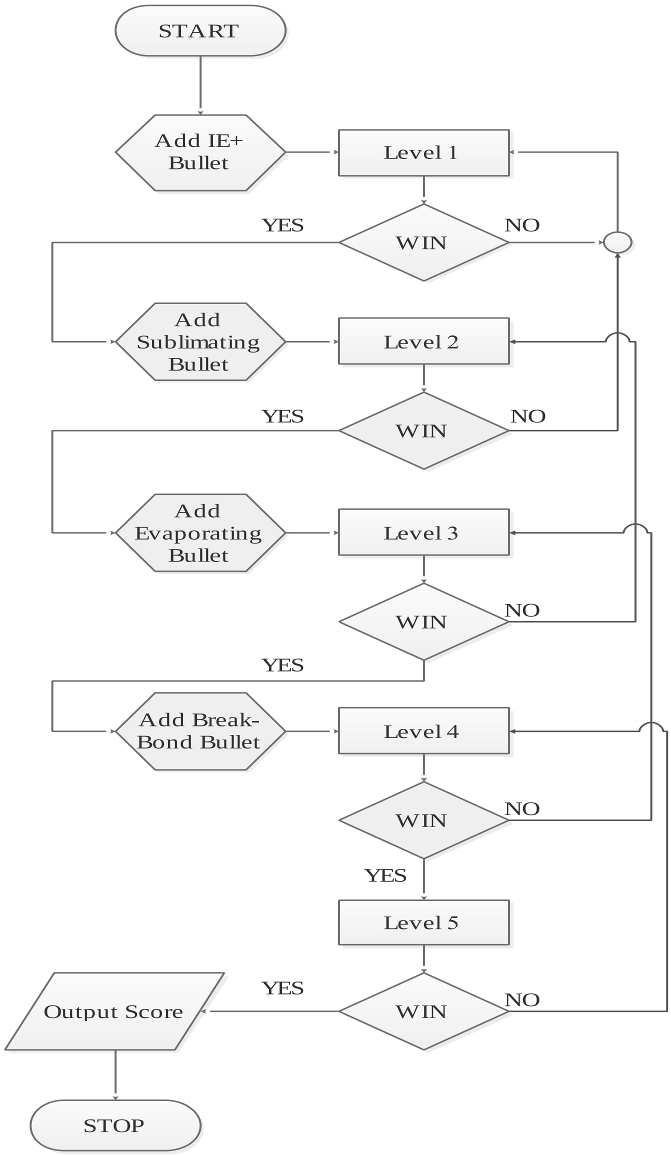


Figure 1. Flow chart of The IE War game.

Level 1: The players have only two bullets, which are IE and IE++ bullet and various gaseous atoms fall down to the ground. The players are able to shoot the atoms by the IE bullet, and then the atom changes its state to be gaseous ion if its energy is corrected. For one round, the players have only three hearts power and they must play to get 10 points for passing any level. Finally, the players who can pass this level would get another bullet, a sublimating bullet.

Level 2: The players have three bullets, which are IE, IE++, and sublimating bullet. In this level, various solid atoms fall down to the ground. The players must choose the sublimating bullet to sublimate solid into a gas, before using IE and IE++ respectively. In the end of this level, the players who complete this level would get the evaporating bullet for the next level.

Level 3: The players have four bullets, which are IE, IE++, sublimating, and evaporating bullet. Likewise, various liquid atoms fall down to the ground and the players must use the evaporating bullet to shot the liquid into a gas, before using IE and IE++. Finally, the players who complete this level would get another bullet called break bond.

Level 4: The players have five bullets, which are all previous four and the break bond bullet. Gaseous molecules fall down and the player must shot them by the break bond bullet to break the gaseous molecules into gas, before using IE and IE++ in next.

Level 5: In this level, all state of matter including gas, solid, liquid, and gaseous molecules fall down to the ground. The players have all five bullets, and must think how to use the bullet correctly and then if the points are high enough, the players would finish this game.

5.3 Instrument

An 18-item perception questionnaire was used to measure students' perception on four subscales: perceived learning, flow, enjoyment, and satisfaction. All of these 5-point Likert scale items obtained from Chu and Hwang (2010) and Barzilai and Blau (2014). From the English version, an identical version in Thai was constructed and two experts were recruited to identify communication validity. For the Thai version, reliability for the overall items was 0.902. There were four items on perceived learning subscale and its reliability was 0.754. The five items on flow subscale indicated that its reliability was 0.661. For enjoyment and satisfaction subscales, the reliability for them was 0.751 and 0.857 respectively.

5.4 Data Collection and Analysis

The participants were asked to complete the perception questionnaire to measure their pre-perceptions on perceived learning, flow, enjoyment, and satisfaction for 10 minutes. After completing the instrument, they were exposed to play The IE War game for 20 minutes. After finishing the game, the students' post-perceptions were examined by the same questionnaire for 10 minutes. To compare the pre-post perceptions on each subscale, paired t-test in SPSS was used to indicate the difference.

5.5 Results and Discussion

Compared to pre-perception scores, the results of post-perception indicated significant higher of students' perception scores on perceived learning ($t = 3.324$, $p < .01$), flow ($t = 5.158$, $p < .01$), enjoyment ($t = 2.480$, $p < .01$), and satisfaction ($t = 4.297$, $p < .01$). This indicated that students have increased their positive perception towards playing the game. Figure 2 illustrates a graphical representation on students' pre- and post-perception scores. The figure showed that the IE War game can affect students' perceived learning, flow, enjoyment, and satisfaction to learn science effectively.

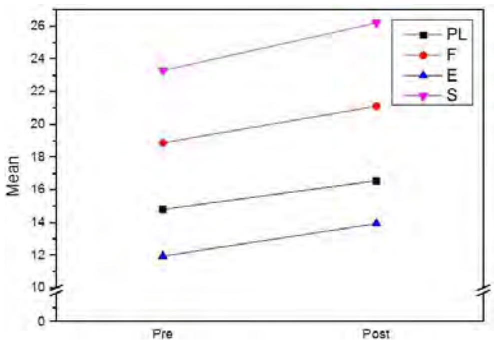


Figure 2. Compare mean scores between pre- and post-questionnaire of four scales.

6. Study 2

6.1 Participants

A total of 81 tenth-grade students were recruited into this study. They were divided into an experimental group (N=43) and a control group (N=44). The experimental group participated with SAGOI class and another was assigned to conventional class as a control group. They were attending a chemistry course for basic education level and all of them have satisfactory skills on basic computer and information and communication technology, but they have no experience yet using educational computer game in chemistry learning before.

6.2 Learning Materials and Activity

The IE War game and another two game related factors that affect to ionization energy were incorporated as instructional tool into open-inquiry learning process, called SAGOI. In this learning process, students will collaborative work together in small groups of three to five members. This pedagogy begins with an open-ended driving question targeted to alternative conceptions about ionization energy commonly found in students. To assist the process of hypothesis generation addressed the question, essential scientific backgrounds are provided to students. Then, students are required to perform generating testable hypotheses, designing an investigation with the games. During playing the game, each group was assigned to access Google Drive spreadsheet, preparing by instructor, for recording scores and what they found into a predetermined table. In an addition, each group was assigned to analyze the recorded data by comparing individual score and also use Google Chat for discussing in the group. When they finished the game, all groups have to communicate findings among groups by creating a PowerPoint presentation via Google Drive presentation. Finally, instructor induces students into a forum for drawing a conclusion based on evidence and collaborative explaining the result of hypotheses testing.

6.3 Instrument

A 25-item science motivation questionnaire was used to measure students' motivation to learn science on five subscales: intrinsic motivation (IM), self-determination (SDT), self-efficacy (SEC), career motivation (CM), and grade motivation (GM) (Glynn et al., 2011). From 25 items English version, the translation an identical version in Thai was constructed and Cronbach's alpha of Thai version were 0.79, 0.81, 0.89, 0.81 and 0.85 for IM, SDT, SEC, CM and GM respectively (Srisawasdi, submitted).

6.4 Data Collection and Analysis

Students were investigated science motivation by the 5-point Likert-scale questionnaire before giving SAGOI intervention for 10 minutes. In the SAGOI class, students participated biology learning of ionization energy using the developed digital game for 230 minutes. After the instruction, students were administered by the same questionnaire again as post-test. The statistical data techniques selected for analyzing students' science motivation was repeated-measures MANOVA in SPSS.

6.5 Results and Discussion

A repeated-measures MANOVA was conducted to determine students' science motivation scores on the five subscales. The assumption of homogeneity of variance-covariance was tested with Box's M Test which was not significant and indicated that homogeneity of variance-covariance was fulfilled (p = .313). The results for the repeated-measures MANOVA indicated significant main effect for group (Wilks' lambda =0.713, $F(5, 75) = 6.029$, $p<0.001$, $\eta^2=0.287$), and time [Wilks' lambda=0.663, $F(5, 75) = 7.634$, $p<0.001$, $\eta^2=0.337$], but interaction effect of time and group were not significant. Thus, these results indicated that students in SAGOI class have changed science motivation over the SAGOI learning experience. Univariate MANOVA on each subscale was conducted as follow-up tests to the one-way MANOVA. The results of the univariate test for control- and experimental group students are summarized in Table 1.

Table 1: The students' subscale means of science motivation by time and univariate MANOVA.

Subscale	Time				F	Sig.	η^2
	CG Pre-test: Mean (SD)	CG Post-test: Mean (SD)	EG Pre-test: Mean (SD)	EG Post-test: Mean (SD)			
Intrinsic motivation (IM)	17.10 (2.73)	17.61 (2.95)	17.95 (2.64)	19.23 (2.51)	8.96	0.004	0.10
Self-determination (SDT)	15.39 (3.71)	16.54 (3.20)	15.65 (2.91)	18.00 (2.92)	26.27	0.000	0.25
Self-efficacy (SEC)	13.17 (3.29)	14.56 (3.65)	15.83 (3.99)	16.98 (3.36)	21.11	0.000	0.21
Career motivation (CM)	17.39 (3.30)	18.07 (3.25)	18.12 (2.86)	19.73 (3.17)	14.87	0.000	0.15
Grade motivation (GM)	17.44 (3.38)	17.44 (3.67)	20.00 (2.70)	20.72 (2.26)	1.15	0.285	0.01

In Table 1, the univariate MANOVA from pre- to post-questionnaire of four subscale scores consists of IM, SDT, SEC and CM were significant differences across time. The univariate results pointed out a significant effect on IM ($F_{1,79} = 8.960$, $p < 0.01$, partial $\eta^2 = 0.102$), SDT ($F_{1,79} = 26.273$, $p < 0.001$, partial $\eta^2 = 0.250$), SEC ($F_{1,79} = 21.113$, $p < 0.001$, partial $\eta^2 = 0.211$) and CM ($F_{1,79} = 14.873$, $p < 0.001$, partial $\eta^2 = 0.158$), but another one GM, the univariate result displayed insignificant ($F_{1,79} = 0.285$, $p < 0.01$, partial $\eta^2 = 0.014$). The result suggested that the increase of science motivation regarding intrinsic motivation, self-determination, self-efficacy, and career motivation from the pre- to post-questionnaire was different between control group and experimental group after participating with the learning instruction. Grade motivation was no effect of time difference on science motivation in learning.

These findings confirm with previous studies (Tuan et al., 2005) that inquiry lessons can increase students' learning motivation. In addition (Erhel and Jamet, 2013) found that game can support motivation and learning which offer it with features that immediate learners to actively procedure in the content and (Ebner and Holzinger, 2007) also found that learning will not be successful if there is a lack of motivation. Therefore we needed some tactics to motivate the students to play the game repeatedly.

5. Conclusion

This study reported impacts of educational computer game on students' perceptions and effects of student-associated game-based open inquiry on students' science motivation. The findings revealed successful of improving students' perceptions and science motivation in context of digital game-based learning experience. This implies that the student-associated game-based open inquiry can be effective in fostering chemistry learning of ionization energy. The results from this study could lead us to conclude that the student-associated game-based open inquiry could be an alternative way for promoting chemistry learning in school science.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, *70*, 65-79.

Caspi, A., & Blau, I. (2011). Collaboration and psychological ownership: how does the tension between the two influence perceived learning?. *Social Psychology of Education*, *14*(2), 283-298.

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, Judy C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education*, *55*(4), 1618-1627.

Ebner, M. and Holzinger, A. (2007). Successful implementation of user-centered game based learning in higher education: An example from civil engineering. *Computers & education*, *49*(3), 873-890.

Erhel, S. and Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, *67*, 156-167.

Giannakos, M. (2013). Enjoy and learn with educational games: Examining factors affecting learning performance. *Computers & Education*, *68*, 429-439.

Glynn, S., Brickman, P., Armstrong, N. and Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159-1176.

Chittleborough, G., & Treagust, D. (2007). The modelling ability of non-major chemistry students and their understanding of the sub-microscopic level. *Chem. Educ. Res. Pract*., *8*(3), 274-292.

Huang, W. (2011). Evaluating learners' motivational and cognitive processing in an online game-based learning environment. *Computers in Human Behavior*, *27*(2), 694-704.

Lo, C. (2010). How student satisfaction factors affect perceived learning. *Journal of the Scholarship of Teaching & Learning*, *10*(1), 47-54.

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education*, *52*(1), 1-12.

Srisawasdi,N. (submitted). Development and Evaluation of Computer-assisted Inquiry Laboratory Learning Module on Self-cleaning Nanotechnology. *Journal of Nano Education*.

Tuan, H., Chin, C., Tsai, C. and Cheng, S. (2005). Investigating the effectiveness of inquiry instruction on the motivation of different learning styles students. *International Journal of Science and Mathematics Education*, *3*(4), 541-566.

The Development and Evaluation of the Online Science Fair Inquiry System based on Scaffolding Design

Li-jen Wang, Chien-yu, Cheng, Chiu-ming, Hu & Ying-Tien Wu*
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*ytwu@cl.ncu.edu.tw

Abstract: Science fair is one of the most common open inquiry activities which can facilitate learners to construct their science knowledge and develop science literacy in school. However, there are a great deal of difficulties and challenges in Taiwan's science fair. For example, novice teachers may neither effectively guide learners to conduct science fair inquiry activities nor facilitate learners to construct related knowledge. To scaffold teachers' instruction and students' learning in science fair inquiry effectively, the "Online Science Fair Inquiry System based on scaffolding design" (OSFIS v 2.0) adapted from (OSFIS v 1.0) was developed in this study. After the development of the system, this study also conducted a series of system evaluations on it. To this end, questionnaire survey was conducted. The participants of the system evaluation in this study were 61 elementary school teachers. The participants expressed highly satisfactory perceived usefulness on teachers' and students' scaffolding design. Also, they had high intention to use the OSFIS v 2.0. Moreover, this study reveals that the system may facilitate both teachers and students to understand the process of science fair inquiry, solve the limitation of activities times, record the portfolio during inquiry activities, and complement teachers' professional knowledge. Finally, some suggestions and implications for teachers to conduct open inquiry activities, system design, and future work are also proposed.

Keywords: Science fair; Inquiry; Online Science Fair Inquiry System, Technology-enhanced inquiry tool; Scaffolding

1. Introduction

There is no doubt that inquiry is the core of modern science education. The major educational goal of inquiry-based teaching or inquiry-based instruction is to help learners study science inquiry skills and enhance the understanding of science inquiry (NRC, 2000). In general, there are five stages of an inquiry activity in science classroom; namely questioning, planning, implementing, concluding, and reporting (Lee et al., 2006). According to the openness of inquiry activities, Bell et al. (2005) categorized four different levels of inquiry activities: confirmation, structured inquiry, guided inquiry, and open inquiry. For science educators, K-12 students are expected to be able to conduct open inquiry. In practice, science fair is the most common open inquiry activity in science classrooms. In many countries, science fair is adopted to help student learn science (Bencze & Bowen, 2009). However, the literature revealed that many teachers may lack of professional knowledge, time, recourses, and assistance when conducting science fair instruction (Anderson, 2002). Only few science teachers know how to guide students to conduct science fair projects or inquiry activities effectively (Justi & Gilbert, 2002). In particular, in recent years, lower quality of the science fair projects conducted by elementary school students have been found in Taiwan. Therefore, how to scaffold elementary school teachers' instruction and students' learning in conducting science fair projects is crucial for educators.

Recently, various technology-enhanced inquiry tools have been developed to scaffold inquiry activities for different science users, activities, and openness (Chung, 2012). For user type, IQWST (Investigating and Questioning our World through Science and Technology) and WISE (Web-based Inquiry Science Environment; Linn et al., 2003) were developed for scaffolding inquiry activities in junior and senior high school science curriculum while OSFIS v.1.0 (Online

Science Fair Inquiry System v.1.0; Chung; 2012) was specifically developed for elementary science fair. For the inquiry activity type, IQWST was mainly for in-class inquiry, WISE was not only for in-class inquiry but online inquiry and experimental inquiry. And the OSFIS v.1.0 was for science inquiry which was different from IQWST and WISE. And for the degree of the openness, IQWST was structured inquiry, WISE was guided inquiry, and the OSFIS v.1.0 was learner-centered open inquiry. However, technology-enhanced inquiry tool for scaffolding elementary science fair inquiry especially based on scaffolding design is still not yet available. In order to scaffold elementary school science teachers' science fair instruction and students' learning in science fair inquiry, this study aimed to develop a new "Online Science Fair Inquiry System" (OSFIS v 2.0) based on scaffolding design. After the development of the OSFIS v 2.0, this study also conducted a series of system evaluations on it.

2. System development

2.1 Participants of system development

The OSFIS v 2.0 developed in this study aimed to provide a platform for elementary teachers who are interested in personal professional development regarding science fair instruction. They can enhance their professional knowledge by using this platform to guide learners to conduct science fair inquiry activities or facilitate learners to construct related knowledge. The development of this system is coordinated by science education and e-learning researcher, in-service science teacher, and system designer. By combining researcher's professional knowledge and in-service teacher's practical experience with system designer's software skills the system design therefore can be more practical for teachers to use.

2.2 System flow chart

In order to help teachers conduct science fair effectively, five inquiry stages suggested by Lee et al. (2006) are included in this system. As shown in Fig. 1, the student flow chart is composed of a series instruction module. Once the students finish each stage of science fair inquiry, they can submit their work to the work reviewing module. If teacher approves the stage, then students can move on to the next one. If not, they need to revise their work according to their teacher's comments and resubmit their revised work. After students finish the five stages, they have completed a science fair project.

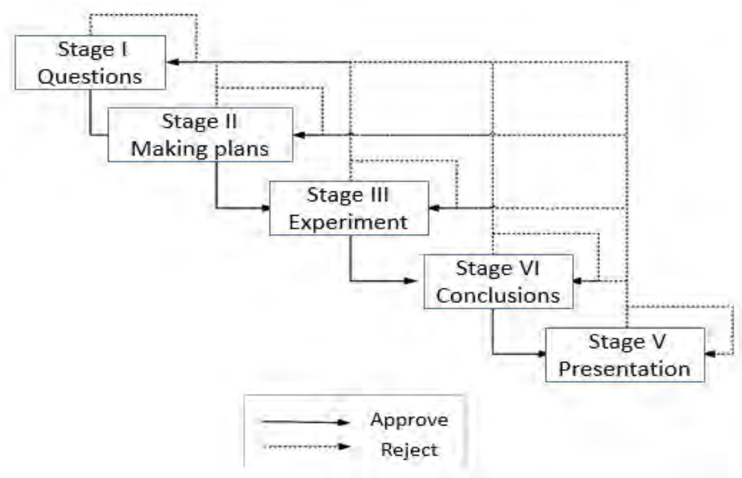


Figure 1. System flow chart

2.3 System framework

The system framework of the OSFIS v 2.0 consists of five main modules and six databases. The six databases store science expert knowledge, student science fair mission, group discussion, interaction, reflection, and user information database. The five modules include user information, interaction and reflection, collaboration, science fair project, and teacher supervision module.

2.4 Modification of Student's and teacher's modules

As shown in Fig. 2, the functions derived from OSFIS v 1.0 are presented with a solid line while new functions of OSFIS v 2.0 are presented with a dotted line. In OSFIS v 2.0, "SF Knowledge" (Science Fair Knowledge) is departed from "SF Management" (Science Fair Management) which may assist students acquire more science fair knowledge. "Learning Process" is also added to help students monitor all learning activities that may help students view their own learning process, including "Previous awarded science fair projects", "Result History", "Grading History", "Log book History", "Discussion History", and "Reflection History". More functions which may help students collaborate and interact with others are added, such as "Sharing Data", "Acquiring Comments", "Acquiring Replied Hints", "Teachers' Hints", "Account Registration", and "Add groups".

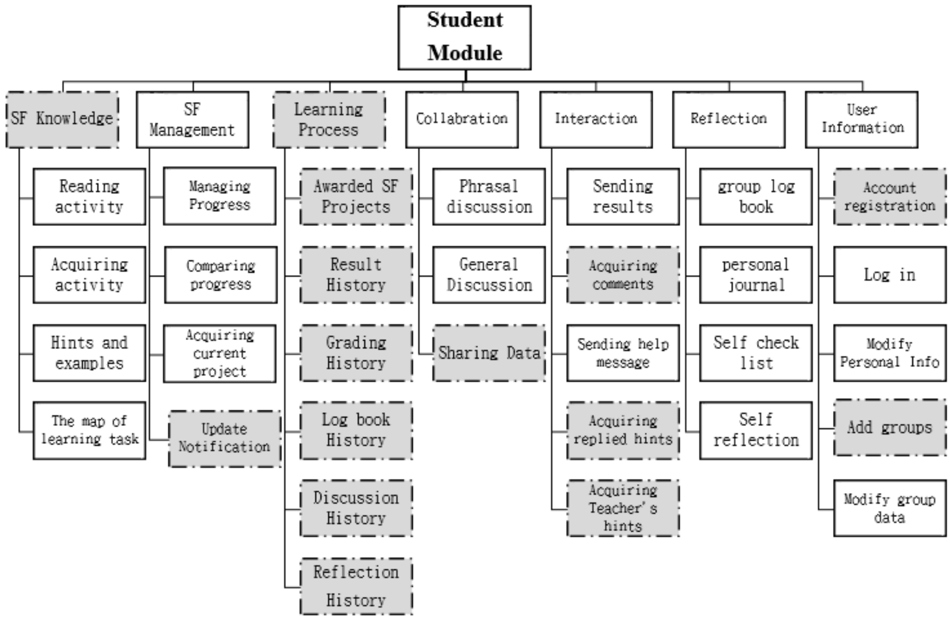


Figure 2. The comparison of student module between OSFIS v 1.0 and v 2.0

For teacher module, as shown in Fig. 3, the functions derived from OSFIS v 1.0 are presented with a solid line while new functions of OSFIS v 2.0 are presented with a dotted line. In OSFIS v 2.0, "SF Knowledge" (Science Fair Knowledge) is departed from "SF Management" (Science Fair Management). Within this module, two more functions, "Read awarded SF projects" and "Add related links" and are added in order to help teachers acquire more science fair knowledge. In "Monitoring & Management module", group modification is added. Teachers can modify groups' learning status according to their progress. "Teaching module" is a brand new module which may help teachers review instructional records. In addition, three more functions, "Remind learning progress", "Project Evaluation", and "Add friends" are added in OSFIS v 2.0. Teachers have the privilege to remind students with their learning progress and they can also evaluate each project's phrasal outcome anytime. Teachers can view other teachers' science fair projects by adding them as good friends.

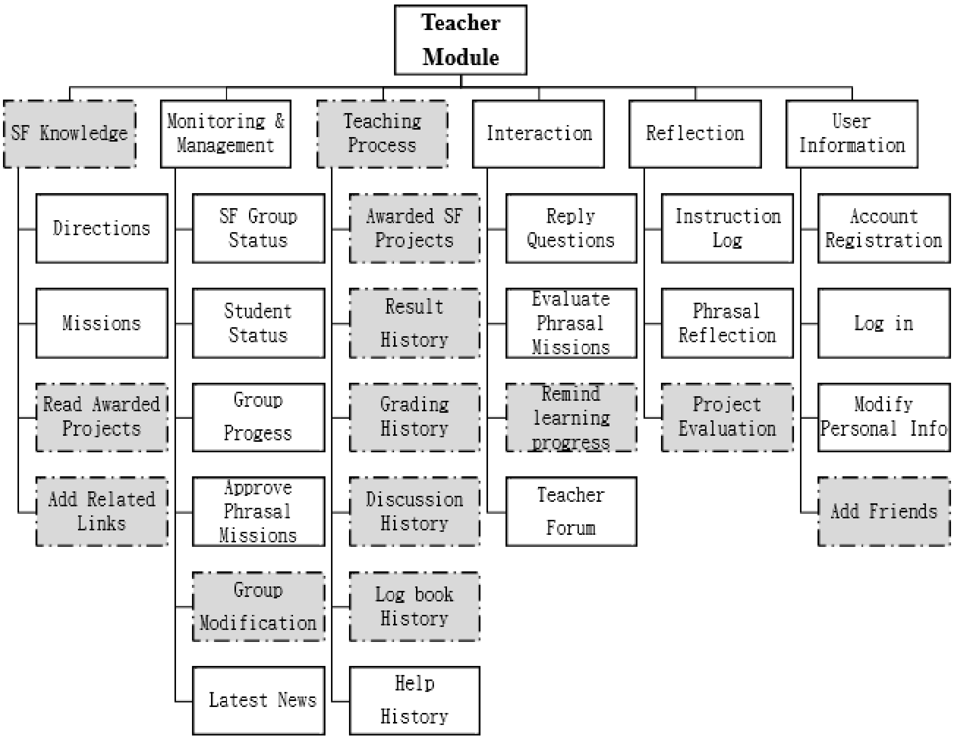


Figure 3. The comparison of teacher module between OSFIS v 1.0 and v 2.0

2.5 Student/ Teacher module and scaffolds

The system is designed to help teachers guide students to conduct science fair projects or inquiry activities by adding scaffolds. As a result, it is necessary to elaborate the connection between the system modules and the scaffolds. The relations between student /teacher module and scaffolds are showed as follows: Teacher's module and scaffolds (Table 1), Student's module and scaffolds (Table 2):

Table 1: Teacher module and scaffolds

Scaffolds	Content	Teaching Challenges	System Modules
Teacher Conceptual scaffolds (TCS)	Offering related Knowledge of Science Fair instruction	Teachers do not understand science fair's relevant knowledge and structure.	Teaching knowledge module
Teacher Procedural Scaffolds (TPS)	The procedure of inquiry instruction	Teachers do not know how to Conduct science fairs.	Inquiry structure
Teacher Metacognitive Scaffolds (TMS)	Making plans	Teachers do not know how to make Instruction plans	Monitoring & Management module & Teaching module
	Monitoring and adjustment	Teachers do not know how to monitor their instructional process	
		Teachers do not know how to make sure students can finish phrasal learning missions.	
		Teachers do not know how to help students modify phrasal missions in order to achieve learning goals	
		Teachers cannot monitor students'	

		learning process	
	Reflection	Teachers do not know how to Reflect their teaching abilities	Reflection module
Teacher Strategic Scaffolds (TSS)	Teacher-student interaction	Teachers increase the chances to interact with students asynchronously	Interaction module
	Peer interaction	Teachers share their progress with other teachers	

Table 2: Student module and scaffolds

Scaffolds	Content	Teaching Challenges	System Modules
Student Conceptual scaffolds (SCS)	Offering related Knowledge of Science Fair instruction	Students do not understand science fair's relevant knowledge and structure.	Teaching knowledge module
Student Procedural Scaffolds (SPS)	The procedure of inquiry instruction	Students do not know how to Conduct science fairs.	Inquiry structure
Student Metacognitive Scaffolds (SMS)	Making plans	Students do not know how to make Instruction plans	Monitoring & Management module & Teaching module
	Monitoring and adjustment	Students do not know how to monitor their instructional process	
		Students do not know how to finish phrasal learning missions effectively.	
		Students do not know how to adjust phrasal missions in order to achieve learning goals	
		Students cannot focus on the current learning tasks	
	Reflection	Students do not know how to reflect their learning tasks	Reflection module
Student Strategic Scaffolds (SSS)	Peer interaction	Group leaders need to arrange tasks Group members need positive interaction	Group collaboration module
	Teacher-student interaction	Students have difficulties learning tasks	Teacher-student interaction module

3. Methodology (System evaluation)

3.1 Participants

The participants of this study consisted of 61 elementary teachers (20 males and 41 females) whose teaching experience ranged from 1 to 32 years, with an average of 15.81 years. 38 (62.3%) teachers had the experience of instructing science fairs. 21 (34.4 %) teachers had no experience of instruction. In general, most participant teacher in this study did not have enough science fair instructional experience which was perfect for this system evaluation.

3.2 Instruments

In this study, the participant teachers' perceived usefulness and usability of the OSFIS v 2.0 as well as their willingness of using the OSFIS v 2.0 were evaluated. To this end, the 6 Likert-scale questionnaire developed in Yuen & Ma (2008) was adapted and used in this study. The modified instrument consists of three scales: usefulness (5 items), usability (5 items), and willing of use (4 items). All the alpha reliability values of the three scales are greater than 0.8, and the overall alpha reliability value of the instrument is 0.96 (Table 3).

Table 3: Item numbers, reliability, and sample items of the instrument scales

Scale	Item	α	example item
ITU	5	0.93	I would like to use OSFIS to conduct science fair.
PU	5	0.96	I find using OSFIS can enhance my teaching.
PEOU	4	0.94	It is easy for me to master the operation of OSFIS.

ITU, Intention to Use; PU, Perceived Usefulness; PEOU, Perceived Ease of Use)

Over all α = 0.96

3.3 Data collection

There were three phases of data collection. In the first phase (30 minutes), the authors gave directions of five stages of learning tasks, system operations, and student interface, and teacher interface of the system. In the second phase (30 minutes), teachers freely explored the scaffolding design, student interface, and teacher interface. In the last phase (10 minutes), overall system evaluation, usability of learning tasks and scaffolding design, and teachers' feedback were evaluated by using a questionnaire developed in this study (Fig 4).

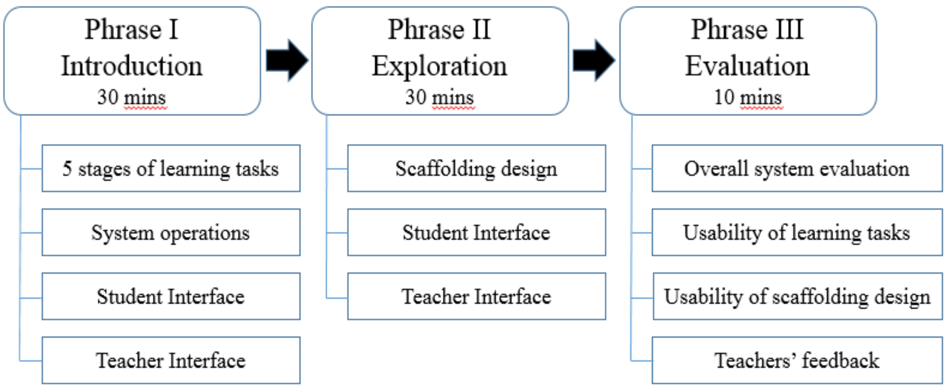


Figure 4. Data collection procedure

4. Major findings and Conclusions

4.1 Major findings

Table 4 shows that the teachers' average scores on perceived usability of teachers' scaffolds are between 4.80 to 4.92. The result was higher than the 6 Likert scale average score (i.e., 3.5).
It indicates that participants in this study perceived high usability on teachers scaffolds.

Table 4: Teachers' average scores on perceived usability of teachers' scaffolds (n=61)

Scaffolds	Mean	S.D.
Teacher conceptual scaffolds（TCS）	4.98	0.66
Teacher procedural scaffolds（TPS）	4.90	0.78
Teacher metacognitive scaffolds（TMS）	4.91	0.76
Teacher strategic scaffolds（TSS）	4.90	0.77
Overall	4.92	0.69

Table 5 shows that the teachers' average scores on perceived usability of students' scaffolds are between 4.90 to 4.98. The result was also higher than the 6 Likert scale average score (i.e., 3.5). It indicates that the participants in this study generally held positive attitude toward the system and were willing to use it.

Table 5: Teachers' average scores on perceived usability of students' scaffolds (n=61)

Scaffolds	Item	Mean	S.D.
Student conceptual scaffolds（SCS）	4	4.83	0.77
Student procedural scaffolds（SPS）	3	4.92	0.81
Student metacognitive scaffolds（SMS）	6	4.83	0.76
Student strategic scaffolds（SSS）	5	4.80	0.76
Overall	18	4.83	0.71

4.2 Conclusions

To scaffold teachers' instruction and students' learning in science fair inquiry, the "Online Science Fair Inquiry System" (OSFIS v 2.0) based on scaffolding design was developed in this study. After the development of the system, this study also conducted a series of system evaluations on it. The participants expressed satisfactory perceived usefulness and ease of use of the system. Also, they had high intention to use the OSFIS v 2.0. Moreover, this study reveals that the OSFIS based on scaffolding design may facilitate both teachers and students to understand the process of science fair inquiry, solve the limitation of activities times, record the portfolio during inquiry activities, and complement teachers' professional knowledge. Some teachers also gave concrete suggestions of the system design which may be provided as the future revision of the system. Based on the findings in this study, the OSFIS v 2.0 may be improved and applied to different levels of school in the future work.

Acknowledgment
This study was funded by the National Science Council, Taiwan, ROC, under grant contract number NSC 101-2628-S-008-001-MY3, but the opinions expressed in this article do not reflect the position of the National Science Council.

References

Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman,R., Hofstein, A., Niaz, M. Treagust, D., Tuan, H.-L.. (2004). Inquiry in science education: International perspectives. *Science Education, 88*(3), 397 – 419.

Abernethy, K., Gabbert, P. & Treu, K. (1998). Inquiry-Based Computer Science Instruction : Some Initial Experiences. *In Proceedings of the 3rd Conference on Integrating Technology into Computer Science Education and on 6th Annual Conference on the Teaching of Computing*, pp.14-17. Ireland.

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education, 13*(1), 1-12.

Banchi, H. & Bell, R. (2008). The Many Levels of Inquiry. *Science and Children, 46*(2), 26-29.

Bencze, J.L., & Bowen, G. M. (2009). A national science fair: Exhibiting support for the knowledge economy. *International Journal of Science Education, 31*(18), 2459-2483.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. Science Education, *86*(2), 175 – 218.

Flick, L., & Bell, R. (2000). Preparing tomorrow's science teachers to use technology: Guidelines for science educators. *Contemporary Issues in Technology and Teacher Education, 1*(1), 39 – 60.

Justi, R. & Gilbert, J.K. (2002). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education, 24*(4), 369–387.

Hamelin, D. (2004). *Searching the Web to Develop Inquiry and Collaborative Skills*. Annual Joint

Conference Integrating Technology into Computer Science Education. Working group reports from ITiCSE on Innovation and technology in computer science education, United Kingdom.
Hill, J. R. & Hannafin, M. J. (2001). Teaching and learning in digital environments: The resurgence of resource-based learning. *Educational Technology, Research and Development*, *49*(3), 37-52.
Kim, M. C., Hannafin, M. J., & Bryan, L. A. (2007). Technology-enhanced inquiry tools in science education: An emerging pedagogical framework for classroom practice. *Science education, 91*(6), 1010-1030.
Kyza, E. A., Constantinou, C. P., & Spanoudis, G. (2011). Sixth graders' co-construction of explanations of a disturbance in an ecosystem: Exploring relationships between grouping, reflective scaffolding, and evidence based explanations. *International Journal of Science Education*, 33, 2489–2525. doi: 10(1080/09500693),2010,550951.
Land, S. M., & Greene, B. A. (2000). Project-based learning with the World Wide Web: a qualitative study of resource integration. Educational Technology *Research and Development, 48*(1), 45–67.
Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional science, 26*(1), 49-63.
Mettas, A. C., & Constantinou, C. C. (2008). The Technology Fair: a project-based learning approach for enhancing problem solving skills and interest in design and technology education. *International Journal of Technology and Design Education, 18*(1), 79-100.
National Research Council. (2011). The national science education standards. Washington, DC: National Academy Press.
Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences, 13*(3), 273 – 304.
Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Slotta, J.D. (2004). The Web-based Inquiry Science Environment (WISE): Scaffolding Knowledge Integration in the Science Classroom. In M.C. Linn, P. Bell and E. Davis (Eds). *Internet Environments for Science Education*. 203-232. Lawrence Erlbaum & Associates.
Windschit, M. (2004). Folk theories of inquiry: How preservice teachers reproduce the discourse and practices of an atheoretical scientific method. *Journal of Research in Science Teaching, 41*(5), 481-512.
Yuen, AHK & Ma, WWK (2008). Exploring teacher acceptance of e-learning technology. *Asia-Pacific Journal of Teacher Education, 36*(3): 229–243.

The difference in Sudoku puzzle-solving ability between undergraduates and postgraduates

Hong-Mei HU[a*], Ling-Jin LI[b], Li-Sha WANG[c] & Feng-Kuang CHIANG[d]
[a]*the faculty of education, Beijing Normal University, China*
[b]*the faculty of education, Beijing Normal University, China*
[c]*the faculty of education, Beijing Normal University, China*
[d]*the faculty of education, Beijing Normal University, China*
*hongmeihu1990@gmail.com

Abstract: As Sudoku is sweeping around the world, it seems to be increasingly prevalent that Sudoku can contribute to the cultivation of logical thinking ability. In the study, empirical approach was adopted to investigate differences in Sudoku puzzle-solving ability and metacognitive ability of mathematical problem-solving between undergraduates and postgraduates, and to examine the relationship between Sudoku puzzle-solving ability and metacognitive ability of mathematical problem-solving. All participants in the study were students studying in Beijing Normal University. The results indicate: (a) no difference between undergraduates and postgraduates in solving the same level Sudoku puzzles and metacognitive ability of mathematical problem-solving, and (b) metacognitive ability of mathematical problem-solving having no significant effect on Sudoku puzzle-solving ability. However, the number of Sudoku puzzles participants had ever finished had appreciable impact on Sudoku puzzle-solving.

Keywords: Sudoku, undergraduates, postgraduates, metacognitive ability

1. Introduction

Sudoku is a logic-based combinatorial (Lenstra, Kan, & Shmoys, 1985) number-placement puzzle. The objective is to fill a 9×9 grid with digits so that each column, each row, and each of the nine 3×3 sub-grids that compose the grid contain all of the digits from 1 to 9. The puzzle setter provides a partially completed grid, and every puzzle has a unique solution (Wayne, 2006).

Problem-solving is a kind of mental activity of seeking methods to deal with problems when facing new situations and new tasks or realizing a lack of ready-made countermeasure to deal with the contradiction between subjective and objective needs in daily life and social practice. Mathematical problem-solving is one of the most important basic processes in mathematic thinking. Researchers defined mathematical problem-solving from the perspective of cognitive theories that it's a process of a series of perceiving (information processing) which depend on cognitive operator with the guidance of target problem and the information provided by the problem (Jinqiu, 1995).

As to metacognition, different scholars define it differently. Metacognition theory considers human learning a procedure of not only conceiving, memorizing, understanding and processing memory materials but self-perceiving, controlling and modulating cognition process. All in all, metacognition is participants perceiving their cognitive activities (Xueying, 2008). Berardi-coletta et al. (1995) tested college students by a traditional question named tower of Hanoi and got a conclusion that metacognition training is good for solving problems. Ping Yu's (2002) research indicated that metacognition level shows a close correlation with mathematic problem-solving ability for senior high school students. Besides, there are researches which indicate that we can improve students' mathematical problem-solving ability by developing their metacognitive ability (Yan, 2005).

Sudoku puzzles are popular all around the world. According to the media that children can improve their logical thinking ability and develop intelligence by solving Sudoku puzzles, while middle-aged people can bring back their energetic mind and throw away life pressure (Song, 2005). Currently some schools of China organize students to solve Sudoku puzzle at math class regularly, expecting to improve students' math scores. But can Sudoku puzzles really improve students' math scores? In the existing literature there are a lot about strategies and arithmetic in solving Sudoku puzzles such as a 9*9 solving strategy introduced by Davis (2006). Some studied how Sudoku puzzles influence humans. For example, Baek, et al. (2008) indicated that Sudoku puzzles can help people develop their logical reasoning ability. There are other researchers focusing on other topics with Sudoku puzzles. Such as Chuen-Tsai Sun et al. (2011), they used Sudoku puzzles as the digital scaffolding to study problem-solving behavior with the premise that if a player can solve the Sudoku puzzles from one level to a higher one, his ability is improved in solving Sudoku puzzles.

However, there is more experience than empirical studies about weather Sudoku puzzles can lead to those changes. No information illustrates a specific function of Sudoku puzzles neither facilitation in students' math scores and mathematical problem-solving ability.

This study aims at probing the relationship between metacognitive ability of mathematical problem-solving and Sudoku puzzle-solving ability. By comparing science undergraduates and postgraduates' metacognitive ability of mathematical problem-solving and Sudoku puzzle-solving ability, we expect to reveal the relationship between the two kinds of ability and explore whether Sudoku puzzles has a close correlation with math scores and mathematical problem-solving ability. The follows are hypotheses of the study.

(a) There is significant difference in metacognitive ability of mathematical problem-solving between undergraduates and postgraduates.

(b) There is significant difference in Sudoku puzzle-solving ability between undergraduates and postgraduates.

(c) Metacognitive ability of mathematical problem-solving has appreciable effect on Sudoku puzzle-solving.

(d) Average math scores in senior high school have positive correlation with Sudoku puzzle-solving ability.

2. Methods

2.1 Participants

30 students of Beijing Normal University were selected by convenient sampling, 15 science undergraduates and 15 science postgraduates. 29 participants came from the school of educational technology, 1 from mathematics. As to gender, postgraduate group consists of 5 males and 10 females, and undergraduate group consists of 4 males and 11 females.

2.2 Materials

2.2.1 Metacognitive ability questionnaire

Two parts form the questionnaire, one part for fundamental information including participants' gender, major, the number of done Sudoku puzzles before the study and average score of math in high school, and the other part for testing the participants' metacognitive ability of mathematical problem-solving which is a scale including 37 items.

The 37 items come from a metacognition scale in mathematical problem-solving compiled by Jianlan Tang et al. of Guangxi Normal University in 2005, which is a comprehensive analysis of Panaeula's questionnaire, Skalower and Sbingling's questionnaire, Jianyue Zhang's self-monitoring ability questionnaire in math among middle school students, Ping Yu's self-monitoring ability questionnaire in mathematical problem-solving, and Haiyan Guo's dynamic and static metacognitive questionnaire with a series of unstructured questionnaire verbal reports in problem solving process. This scale is a structured original questionnaire adopting a five Likert scale. In the scale, Metacognitive ability is divided into 3 main factors and 37 items. The main factors are metacognitive knowledge, metacognitive experience, and metacognitive strategy. The sum score of these 37 items a participant get is seemed as metacognitive ability score. The scale's α coefficient of the total internal consistency is 0.901, indicating a good reliability and validity (Jianlan, Ying & Fucheng, 2005).

As mentioned earlier, metacognitive ability has a close correlation with mathematic problem-solving ability (Ping, 2002. Yan, 2005). In the study, we regarded the metacognitive ability score as participants' mathematic problem-solving ability.

2.2.2 Sudoku puzzles

Sudoku is a number-placement puzzle with the rule that each column, each row, and each of the nine 3×3 sub-grids that compose the grid contain all of the digits from 1 to 9 but no repetition is allowed.

We selected 9 basic Sudoku puzzles from Sudoku puzzles (2006) compiled by Wayne Gould in New Zealand. Every puzzle is presented on a half of A4 paper.

The maximum score of each puzzle is 10. The score of each puzzle a participant got is according to the participant finished percentage in total blanks. A participant's Sudoku score is the sum of 9 Sudoku puzzles' scores he got in 1 hour. In the study, we regarded the Sudoku score a participants got as his/her Sudoku puzzle-solving ability.

2.2.3 Interview question

A brief interview question was designed to investigate participants' attitude towards Sudoku puzzles. The question is "Do you want to do Sudoku puzzles in the future?"

2.3 Research process

In order to make the study more reasonable and achievable, a pilot study was conducted in which 2 undergraduates participated, and then we revised some details of the study with their advice. The 15 undergraduates didn't include the 2 undergraduates participated in pilot study. Every participant decided specific time respectively from Dec. 20th, 2013 to Jan. 6th, 2014. It took each participant about 70 minutes. Below was the detailed process.

(a) A researcher explained to the participant the study content, purpose and what they would do.

(b) The participant filled in the metacognitive ability questionnaire. It took about 3 minutes.

(c) A researcher gave the 9 Sudoku puzzles to the participant, explained what was Sudoku and the rule, and remind the participant that he/she could choose a few to do, not all.

(d) The participant got down to Sudoku puzzle-solving for 1 hour with pencils and erasers.

(e) A researcher asked the participant the interview question and recorded the answer.

3. Results

SPSS was used to analyze the data of the study. Significance level was 0.05.

3.1 Differences in metacognitive performance and Sudoku puzzle-solving performance

Independent-Samples T test was used to test the differences in metacognitive performance and Sudoku puzzle-solving performance between undergraduates and postgraduates (Table 1, Table2).

Table 1: Two-tailed t-test results for undergraduates and postgraduates' metacognitive scores.

Dimension	Group	N	Mean	SD	Sig.(two-tailed)
Metacognitive knowledge	Undergraduates	15	27.00	6.36	0.665
	Postgraduates	15	26.13	4.31	
Metacognitive experience	Undergraduates	15	19.40	5.53	0.944
	Postgraduates	15	19.53	4.75	
Metacognitive strategy	Undergraduates	15	60.27	12.31	0.447
	Postgraduates	15	63.40	9.78	
Total score	Undergraduates	15	106.67	21.52	0.729
	Postgraduates	15	109.07	15.64	

As shown in Table 1, the difference of undergraduate and postgraduates' total mean scores is 2.40. The mean scores of undergraduate and postgraduates are very close in metacognitive knowledge and metacognitive experience dimensions, and score differences in metacognitive strategy dimension is 3.13. The standard deviations of postgraduates' score in metacognitive knowledge, metacognitive experience and metacognitive strategy are less than those of undergraduates. Overall, the results indicate that no statistical significance ($p>0.05$) in metacognitive ability between undergraduates and postgraduates, and invalidate hypothesis a.

Table 2: Two-tailed t-test results for undergraduates and postgraduates' Sudoku scores.

Group	N	Mean	SD	Sig.(two-tailed)
Undergraduates	15	24.86	11.91	0.656
Postgraduates	15	27.14	15.57	

As shown in Table 2, the difference of undergraduates and postgraduates' Sudoku mean scores is 2.28. The results indicate no statistical significance ($p>0.05$) in Sudoku puzzle-solving ability between undergraduates and postgraduates, so hypothesis b is invalid.

3.2 Correlations of Sudoku score with other variables

The linear regression analysis was used to analyze the correlations of Sudoku score with other variables (Table 3).

Table 3: Coefficient[a] of linear regression results for correlations of Sudoku scores with other variables.

Model	Unstandardized Coefficients		t	Sig.
	B	Standard Error		
（Constant）	-8.721	25.537	-0.342	0.736
Sex	3.448	5.322	0.648	0.524
Grade	2.482	5.012	0.495	0.625
The number of done Sudoku puzzles	8.689	3.108	2.796	0.011*
Average math scores in senior high school	1.377	4.865	0.283	0.780
Metacognitive knowledge	1.701	1.093	1.556	0.134
Metacognitive experience	-1.710	1.141	-1.499	0.148
Metacognitive strategy	-0.041	0.286	-0.143	0.888

a. Dependent variable: Sudoku score
*p<0.05

Data in Table 3 indicates that only the number of done Sudoku puzzles before the study has a significant effect on Sudoku score ($p<0.05$), and metacognitive ability has no significant effect on Sudoku puzzle-solving ability. Hypothesis c is invalid.

The linear regression analysis was used to analyze the correlations of Sudoku score with the number of done Sudoku puzzles before the study, as ANOVA results shown in Table 4.

Table 4: ANOVA[a] results for correlations of Sudoku scores with the number of done Sudoku puzzles.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1379.710	1	1379.710	9.567	0.004[b]
Residual	4038.049	28	144.216		
Total	5417.759	29			

a. Dependent Variable: Sudoku score
b. Predictors: (Constant), the number of done Sudoku puzzles before the study

Data in Table 4 shows that the number of done Sudoku puzzles before the study can be used to predict Sudoku score. The linear regression equation is $Y = 14.246 + 7.668 X$.

Wenling Li et al. (2008) divided correlativity into four levels according to Pearson correlation coefficient: (a) No correlation or weak correlation, (b) normal correlation, (c) high correlation, and (d) strong correlation. Pearson correlation coefficient for all variables shows that Pearson correlation coefficient of Sudoku score and the number of done Sudoku puzzles before the study is 0.505, which means the two variables have high correlativity. Pearson correlation coefficient of average math scores in senior high school and Sudoku score is 0.092, which means that these two variables have no correlativity. So hypothesis d is invalid.

3.3 Interview results

According to the results of metacognitive ability questionnaire, there were 20 participants hadn't done any Sudoku puzzles, 11 undergraduates, 9 postgraduates. 17 participants who hadn't done any Sudoku puzzles before answered this question. 12 of them said they wanted to do, 3 said they didn't want to do any more, 2 said they might want to do. 9 participants who had done some Sudoku puzzles before answered this question. 5 of them said they wanted to do.

4. Discussion

4.1 Average math scores in senior high school and Sudoku score

On the basis of this study results, there is no correlation between average math scores in senior high school and Sudoku score. So that we doubt whether it is worth students' while solving Sudoku puzzle at math class regularly. Class time is very important for students. If solving Sudoku puzzle cannot improve students' ability or scores, maybe school should reconsider how to help students make better use of class time.

However, we cannot draw the conclusion that solving Sudoku puzzles can't improve math abilities as this study has several limits. (a) This study only include average math scores in senior high school, not including math scores in junior high school and primary school. So it is illogical to say that solving Sudoku puzzles cannot improve math scores. (b) We divided average math scores in senior high school into five score section which is not a subtle rational division. (c) We gave only 9 puzzles to participants and they did Sudoku puzzles just for 1 hour. They cannot form a steady Sudoku puzzle-solving ability in such a short time, so the data maybe represents participants' ability accurately.

In the future, we can conduct a long-term tracking study in primary school and junior high school to explore whether solving Sudoku puzzles has impact on math scores.

4.2 The number of done Sudoku puzzles before the study and Sudoku score

According to the result, the number of done Sudoku puzzles before the study has a significant impact on Sudoku score. We can infer that to some extent, the larger number of Sudoku puzzles a participant had done the more experience and strategies he might get, and he would get the higher Sudoku score in the study. However, we should consider whether a certain number may exist. Participants may get equivalent scores when they had done more than the certain number of Sudoku puzzles. This requires a deeper study.

4.3 Metacognitive ability and Sudoku score

The results show that there is no significant difference in metacognitive ability between undergraduates and postgraduates. A possible reason is that two groups of participants are adults and in the stage of formal operational stage defined by Piaget. They have a steady cognitive level, and there is no significant difference in metacognitive ability among them. Another possible reason is that the metacognitive ability questionnaire doesn't work well to detect the metacognitive level.

Data show that metacognitive ability has no statistical significant effect on Sudoku score. However, we can't draw the conclusion that metacognitive ability has no significant impact on Sudoku puzzle-solving ability, because we don't have clear evidence that metacognitive ability of mathematical problem-solving is the same with metacognitive ability in Sudoku puzzle-solving. The metacognitive ability questionnaire detects the former not the latter. It is necessary to design a special scale to test metacognitive ability in Sudoku puzzle-solving in further study.

4.4 Attitude towards Sudoku puzzles

According to results of interview question, although 71.4 % of participants who had never done Sudoku puzzles before the study, 70.6% of them wanted to try it later. So we can speculate that Sudoku puzzle is an attractive game. Why is it so attractive? And why do some people think solving Sudoku puzzles is good for them improving ability subjectively? We can do another study to investigate the reason.

5. Conclusion

According to the study, there is no significant difference in metacognitive ability of mathematical problem-solving between undergraduates and postgraduates, and no significant difference in Sudoku puzzle-solving ability between undergraduates and postgraduates. Besides, metacognitive ability of mathematical problem-solving has no appreciable effect on Sudoku puzzle-solving. In addition, average math scores in senior high school have no positive correlation with Sudoku puzzle-solving ability. However, the number of done Sudoku puzzles before the study has a significant effect on Sudoku puzzle-solving ability.

There were limitations in the study. We had a small number of participants and more proper materials should be used.

Acknowledgements

This study thanks sincerely to professor Keding Zhong for his guidance and suggestion to study goals and data analysis method, Yehui Wang for detailed advice on data analysis and discussion, and vice professor Yin Wang for her kindness to provide research facility. Besides, we thank to the 32 participants.

References

Baek, Y., Kim, B., Yun, S., & Cheong, D. (2008, October). Effects of two types of sudoku puzzles on students' logical thinking. In *Proceedings of the 2nd European Conference on Games Based Learning* (pp. 19-24). Academic Conferences Limited.

Berardi-Coletta, B., Buyer, L. S., Dominowski, R. L., & Rellinger, E. R. (1995). Metacognition and problem solving: A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 205.

Davis, T. (2006). The mathematics of Sudoku.

Fang Jinqiu. (1995). Cognitive definition of mathematics problem-solving. *Journal of Beijing Institute of Education* (pp. 22-24).(in Chinese)

Han Yan. (2005). Develop metacognition, and bring up junior high school students' ability of mathematics problem-solving (Master's thesis, Capital Normal University).(in Chinese)

Lenstra, J. K., Kan, A. R., & Shmoys, D. B. (1985). *The traveling salesman problem: a guided tour of combinatorial optimization* (Vol. 3). New York: Wiley.

Li Wenling, Zhang Houcan, Shu Hua. (2008,12). *Education and psychology's quantitative research method and statistical analysis.* Beijing: Beijing Normal University Publishing Group.(in Chinese)

Ping Yu. (2002). Mathematical problem solving cognitive mode and the teaching theory research. Nanjing *Normal University doctoral dissertation*. (in Chinese)

Sun, C. T., Wang, D. Y., & Chan, H. L. (2011). How digital scaffolds in games direct problem-solving behaviors. *Computers & Education, 57*(3), 2118-2125.

Tang Jianlan, Zhou Ying, & Tang Fucheng. (2006). The design of metacognition questionnaire in mathematics problem-solving. *Journal of Mathematics Education, 14*(4), 44-48.(in Chinese)

Wayne Gould. (2006). Sudoku.

Wo Song. (2005, October 19). The global digital puzzle lights up students' interest in mathematics: Sudoku is coming. *China Teacher*, 10.(in Chinese)

Yang Xueying. (2008). The differences between the arts and science students in solving the fundamental problems of inequality(Master's thesis, Soochow University).(in Chinese)

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Investigating the role of self-explanation and co-explanation in 4th graders' game-based science learning

Chung-Yuan HSU[a*], Hung-Yuan WANG[b], & Shih-Hsuan WEI[c]
[a]*Department of Child Care, National Pingtung University of Science and Technology, Taiwan*
[b] *Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan*
Department of Mathematics Education, National Taichung University of Education, Taiwan
*jackohio@gmail.com

Abstract: This study extends our previous studies on investigating the effects of embedding self-explanation principle into game-based science learning. In order to enhance the students' generating their own explanations during the game, we replaced the multiple choice questions with the design of allowing dyads to co-explain their causes of failure in the game via utilizing the technique of online chat. The participants were 60 4th graders recruited from an elementary school in southern Taiwan. They were randomly assigned to dyads of either an experimental group (conducting co-explanation via online chat) or a control group (conducting self-explanation via multiple choice questions). The measurements included the pretest, posttest, and a three-week retention test. The results show that both games had a positive impact on facilitating the students' acquisition of scientific concepts. But, the players who performed co-explanation via online chat did not outperform those who used multiple choice questions as self-explanation prompts. Through analysis of dialogue of the players in the experiment group, we found that the quality of the dyads' dialogue was poor; they rarely discussed the causes of failure when the prompts appeared in the game.

Keywords: Game-based learning, self-explanation, science learning, multiplayer game

1. Introduction

Researchers (Chi, Bassok, Lewis, Reimann, & Glaser, 1989) found that students learned well when they were asked to generate explanation to themselves. This constructive learning process enables learners to generate inferences to fill in information gap, integrating information, and monitoring and repairing faulty knowledge (Roy & Chi, 2005). In the recent years, a growing number of researchers attempt to integrate self-explanation principle into educational games and investigate its impacts on players' learning outcomes (Adams & Clark, 2014; Hsu, Tsai, & Wang, 2012 & Hsu, Tsai, & Wang, in press; Johnson & Mayer, 2010). This study extends our previous studies (Hsu et al., 2012, in press) on investigating the effects of embedding self-explanation principle into game-based science learning. Although both studies as well as the previous research (Adams & Clark, 2014; Johnson & Mayer, 2010; O'Neil et al.,2014) have identified the positive impacts of using multiple choice questions as self-explanation prompts, it might still risk limiting learners' generating inferences and hinder robust learning outcomes. Thus, to enhance the students' generating their own explanations, Hsu et al. (in press) suggested replacing the multiple choice questions with the design of allowing dyads to co-explain their causes of failure in the game via utilizing the technique of online chat. Through interaction with peers in the game, we hypothesize that the experimental condition would outperform the control condition since the dyads in the experimental group could share diverse perspectives, co-construct knowledge, and benefit from explaining another person's reasoning. In sum, this study attempted to examine how different forms of self-explanation influence students' game-based science learning. The guiding questions are:

1. What are the effects of self-explanation and co-explanation in game-based science learning?
2. How is the quality of co-explanation during game playing?

2. Methodology

2.1 Participants

The participants were 60 4th graders recruited from an elementary school in southern Taiwan. Without receiving formal instruction regarding light and shadow concepts, they were randomly assigned to dyads of either an experimental group (conducting co-explanation via online chat) or a control group (conducting self-explanation via multiple choice questions). There were 13 females and 17 males in the experimental condition and 16 females and 14 males in the control condition. Both groups played a multiplayer game with self-explanation embedded.

2.2 The game

The game of this study was developed by the researchers to support forth graders' learning of shadow and light concepts. The game consisted of three stages and each one was designed to instruct a core concept, such as the relationship between the height of a light source and the length of the shadow produced, shadow change throughout the day, and shadow intensity, respectively for Stage 1 to 3. The participants were required to play the game with a peer randomly assigned by the researchers (see Figure 1). That is, neither of them knew who their partner was or where she or he was situated. During game playing, a self-explanation prompt appears whenever a mistake is made. Both players had to stop playing and respond to the prompt. The participants in the experimental group were encouraged to discuss the causes of failure via online chat. When the discussion was completed, they could click a button and continue the game. However, the students in the control group used multiple-choice questions as self-explanation prompts in the game context. The time limitation for all the three stages was 35 minutes. The players would be directed to the posttest when failing to meet the limitation.

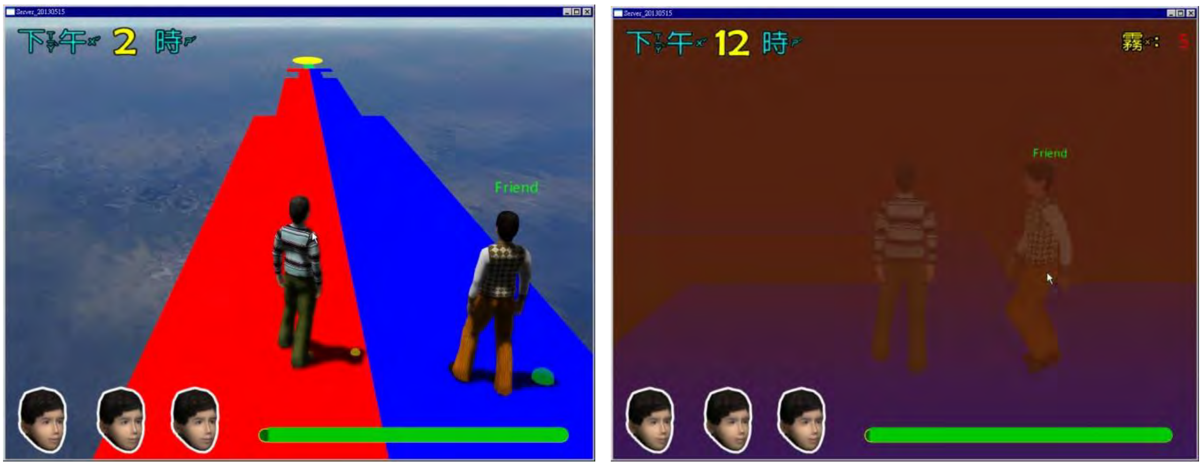


Figure 1. Screenshot of the game.

2.3 Measurement

A 10-item test was used to measure the participants' understanding regarding light and shadow covered around the main concepts in the game. Each student took the test before the game, right after the treatment, and three weeks after the treatment. Sharing the same questions, the test only varied in the order of displaying the questions and options. These items were also used in Hsu et al.'s (2011, in press) studies. The reliability coefficient was 0.60 in Hsu et al. (in press), suggesting acceptable reliability.

2.4 Procedure

Participants were randomly assigned to either the experimental or control group and individually seated at a computer when entering the computer classroom. A researcher introduced the study and the tasks to the class. Following the introduction, the students took a pretest without a time limit (averagely less than six minutes). Later, the researchers helped the students build up an online connection of the game with their partners, and log in Skype (a technology allows users to communicate with peers by using a microphone over the Internet) for those in the experimental condition. Their narration during the game playing would be recorded for further analysis. The students then played the game for 30 minutes. They received a posttest when passing the three stages or over the time limit. Each student also took a retention test after three weeks.

2.5 Data analysis

First of all, a series of paired t-test were conducted to compare students' improvement from the pretest. This study later investigated the score difference of both groups by using the pretest scores as a covariate. A content analysis was utilized to probe the dialogue of the players in the experimental group during game playing.

3. Results

Table 1 shows the results of paired t-tests. As indicated, the students' posttest and retention scores were significantly higher than their pretest scores in both experimental and control condition. This finding suggests that both games could positively facilitate the students' acquisition of scientific concepts.

Table 1: Paired t-tests for the scores of the control and experimental groups.

Group	Test Type	N	Mean	SD	*t*
Control	pretest	30	5.83	1.80	-4.87*
	posttest	30	7.63	1.50	
	pretest	30	5.83	1.80	-4.82*
	retention	30	7.53	1.85	
Experimental	pretest	30	5.37	2.16	-5.96*
	posttest	30	8.23	1.63	
	pretest	30	5.37	2.16	-5.46*
	retention	30	7.97	1.94	

*<.001

This study further examined the score difference between the two groups by using the pretest scores as a covariate and the posttest and retention score as dependent variables. The assumption of homogeneity of regression was tested and was not violated (F =1.16, p > .05; F = 1.44, p < .05). The ANCOVA results of the posttest and retention are shown in Table 2. As shown, no statistically significant difference was identified. That is, the players who performed co-explanation via online chat did not outperform those who used multiple choice questions as self-explanation prompts.

Table 2: Descriptive data and ANCOVA results for the posttest and retention scores.

Type	Group	N	Adjusted mean	Std. error	*F*
Posttest	Control	30	7.61	.29	2.57
	Experimental	30	8.26	.29	
Retention	Control	30	7.47	.33	1.44
	Experimental	30	8.03	.33	

As aforementioned, the players' communication during the game would be recorded and transcribed for further analysis. In this preliminary analysis, we focus on the players' narration right after the failure in the game. The results show that the dyads rarely discussed the causes of failure when the self-explanation prompt appeared, such as:

Participant 12: I am dead.
Participant 10: I am dead, too.
Participant 12: No problem, let's play again.
Participant 10: Well, this time we should walk slowly.

In addition, they tended to blame their partner for the cause of mistakes. Take Participant 13 for instance, "I hate you. I only make one mistake but you make two. It is annoying that we keep failing." Although some dyads might come up with the tricks to pass the game, these tricks were not absolutely correct. An example is:

Participant 1: Oops, I am completely dead
Participant 6: I told you not to move but you never listen. Maybe you should walk on the red lane. Be careful! Do not fall in the sea.

Regarding the above example, the players should pay attention to shadow change throughout the day, rather than the difference in the lanes they walk.

4. Discussion

Self-explanation effects become effectively when learners can generate inferences to fill in missing information, integrate information and repair faulty knowledge (Roy & Chi, 2005). However, the previous research pointed out that utilizing multiple choice questions as self-explanation prompts in the game context was likely to limit the players' generating inferences (Hsu et al., 2014). To solve this problem, this study implemented a design by having dyads co-explain their causes of failure during game playing and investigate its impact on the participants' learning outcomes. However, no statistically significant difference was identified. Players who co-explained via online chat did not perform better than those who used multiple choice questions as self-explanation prompts. In addition, through analyzing the dyads' dialogue, we found that the quality of the dyads' narration was not satisfied and they rarely discussed the possible causes of failure when the prompt appeared. They chatted all the time and blamed their partner for failure. Although some of them could identify some tricks to pass the game, they might not be accurately linked to the targeted concept.

According to Chi's (2009) framework of passive-active-constructive-interactive learning strategies, interacting with a peer in a computer-based environment can be classified as interactive learning activities only when the dialogue includes substantive contributions from both partners, as well as learners respond to scaffoldings and modify errors based on feedback. It seems like that the participants of the experimental condition simply taking turns speaking, which could not be categorized as an interactive learning event. To sum up, having dyads collaboratively construct knowledge in game-based science learning is a one of ultimate level of learning strategies. But, future studies still need to think about ways to promote quality of players' interaction, such as designing events to confront or challenge the partner's statements, or encourage involvement into deeper discussion.

Acknowledgements

This research was supported by the projects from the National Science Council, Republic of China, under contract number NSC 102-2511-S-020-001 and MOST 103-2511-S-020 -001 -MY3.

References

Adams, D. M., & Clark, D. B. (2014). Integrating self-explanation functionality into a complex game environment: Keeping gaming in motion. *Computers & Education*, *73*, 149-159. doi: 10.1016/j.compedu.2014.01.002

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science, 13*, 145-182.

Hsu, C.-Y., Tsai, C.-C., & Wang, H.-Y. (2012). Facilitating third graders' acquisition of scientific concepts through digital game-based learning: The effects of self-explanation principles. *The Asia-Pacific Education Researcher*, 21, 71-82.

Hsu, C.-Y., Tsai, C.-C., & Wang, H.-Y. (in press). Exploring the effects of integrating self-explanation into a multi-user game on acquisition of scientific concepts. *Interactive Learning Environments*. doi: 10.1080/10494820.2014.926276

Johnson, C. I., & Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior, 26*, 1246-1252. doi: 10.1016/j.chb.2010.03.025

O'Neil, H. F., Chung, G. K. W. K., Kerr, D., Vendlinski, T. P., Buschang, R. E., & Mayer, R. E. (2014). Adding self-explanation prompts to an educational computer game. *Computers in Human Behavior, 30*, 23-28. doi: 10.1016/j.chb.2013.07.025

Roy, M., & Chi, M.T.H. (2005). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 271-286). New York: Cambridge University Press.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Preliminary Requirements Analysis towards an Integrated Learning Analytics System

Byung-gi CHOI[a], Yong-sang CHO[b], Jaeho LEE[c*]
[a]*School of Electrical and Computer Engineering, University of Seoul, Republic of Korea*
[b]*Korea Education & Research Information Service*
[c]*School of Electrical and Computer Engineering, University of Seoul, Republic of Korea*
**jaeho@uos.ac.kr*

Abstract: An integration of various information and processes for learning analytics into a united framework is the key to the development of an open and extensible learning analytics system. Recently, we have taken a step towards developing such a framework by starting to build a reference software architecture which in turn will allow us to identify the structure and the workflow of learning analytics systems. Our final goal is to develop an explicit specification of the learning analytics architecture as the international standard so that open and extensible learning analytics systems can be built for worldwide interoperability. In this paper, we present the result of our preliminary requirement analysis towards such an open and interoperable learning analytics system. The analysis focuses mainly on the system aspects of existing well-known frameworks such as IMS Global learning analytics platform.

Keywords: Learning Analytics, Requirement Analysis, Reference Architecture, Standards

1. Overview

Learning analytics (LA) systems require integration of the processes of measurement, collection, analysis and reporting of data about learners and their contexts, and thus involve multi-disciplinary areas including artificial intelligence, information science, statistics, visualization, and so on.
An effective integration of various information and processes for learning analytics into a united framework is the key to the development of an open and extensible learning analytics system because the system should be *open* to several related areas of research such as academic analytics, action research, educational data mining, personalized adaptive learning, and more. The system also should be *extensible* as new methodologies and technologies emerge rapidly. Especially, the big data technology has been evolved markedly to make it possible to collect data massively, analysis instantly, and visualize appropriately for learning analytics field.

Recently, we have taken a step towards developing such a framework by starting to build a reference software architecture which in turn will allow us to identify the structure and the workflow of learning analytics systems.

Our final goal is to develop an explicit specification of the learning analytics architecture as the international standard so that open and extensible learning analytics systems can be built for worldwide interoperability. In this paper, we present the result of our preliminary requirements analysis towards such an open and interoperable learning analytics system. The analysis focuses mainly on the system aspects of existing well-known frameworks such as IMS Global learning analytics platform.

The remainder of this paper is organized as follows. Firstly, we briefly review the purpose of learning analytics and survey related works and standardization activities in learning analytics field. Based on this survey, we then discuss the basic requirements of the reference software architecture. Finally we gives a summary of the main results of this paper and highlight directions for future work.

2. Backgrounds

In this section, we introduce the basic backgrounds that are needed for requirement analysis for learning analytics including the analysis levels of learning analytics and standardization activities.

2.1 Learning Analytics and Big Data

Learning analytics, as stated in the “Call for Papers of the 1st International Conference on Learning Analytics & Knowledge (LAK 2011)", is “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs.” The tasks of measurement, collection, analysis, and reporting in this definition correspond closely to the major activities in big data, that is, collection, processing, analysis, and visualization of data, as shown in Figure 1.

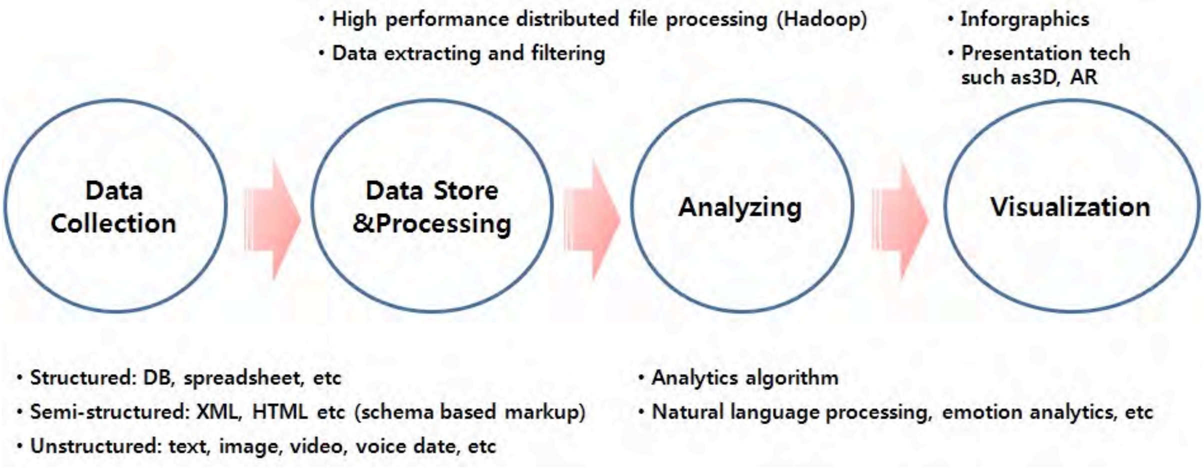


Figure 1. Big data workflow

Such correspondence is not coincidental but suggests that learning analytics can take advantage of the technological advancement of big data in building a learning analytics framework. Our requirement analysis for learning analytics is also largely borrowed from that of the big data framework.

2.2 Range of Learning Analytics

According to Buckingham Shum (2012), the range of learning analytics can be defined as macro-, meso- and micro-levels.

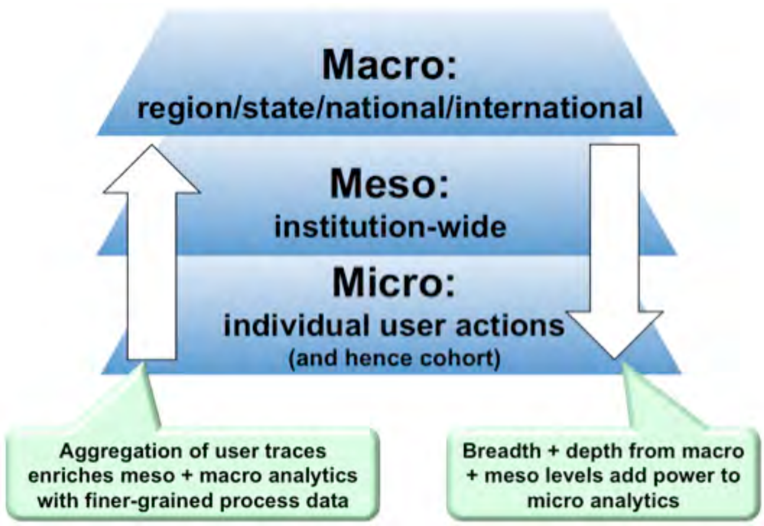


Figure 2. Levels of Learning Analytics (Buckingham Shum, 2012)

Buckingham Shum (2012) describes the levels of learning analytics as follows.

- Macro-level analytics is the cross-institutional analytics over region, state, national, or international institutions for students' lifetimes. Macro-analytics becomes increasingly real-time, incorporating more data from lower meso- or micro-levels, utilizing data integration methodologies that are developed in non-educational sections.
- Meso-level analytics operates at institution level, benefiting from the common business processes for business intelligence, utilizing the tools to integrate data silos in enterprise warehouse, optimize workflows, generate dashboards, mine unstructured data, predict future trends, and so forth.
- Micro-level analytics supports the tracking and interpretation of process-level data for individual learners or groups. This data is of primary interest to learners themselves and correspondingly the most personal, since it can disclose online activities as well as physical activities such as geolocation, library loans, purchases, and interpersonal data such as social networks.

As shown in Figure 2, while the aggregation of user date from the micro-levels enriches meso- or macro-level analytics, the breadth and depth at the macro- or meso-levels add power to micro-level analytics in building predictive models or providing feedback to learners. We thus believe that an effective integration of data and activities among these layers is essential requirement for mutual enrichment.

2.3 Reference Model for Learning Analytics

Chatti, Dyckhoff, Schroeder, and Thüs (2012) describe a reference model for learning analytics based on four dimensions: data and environments (what?), stakeholders (who?), objectives (why?), and methods (how?) as depicted in Figure 3.

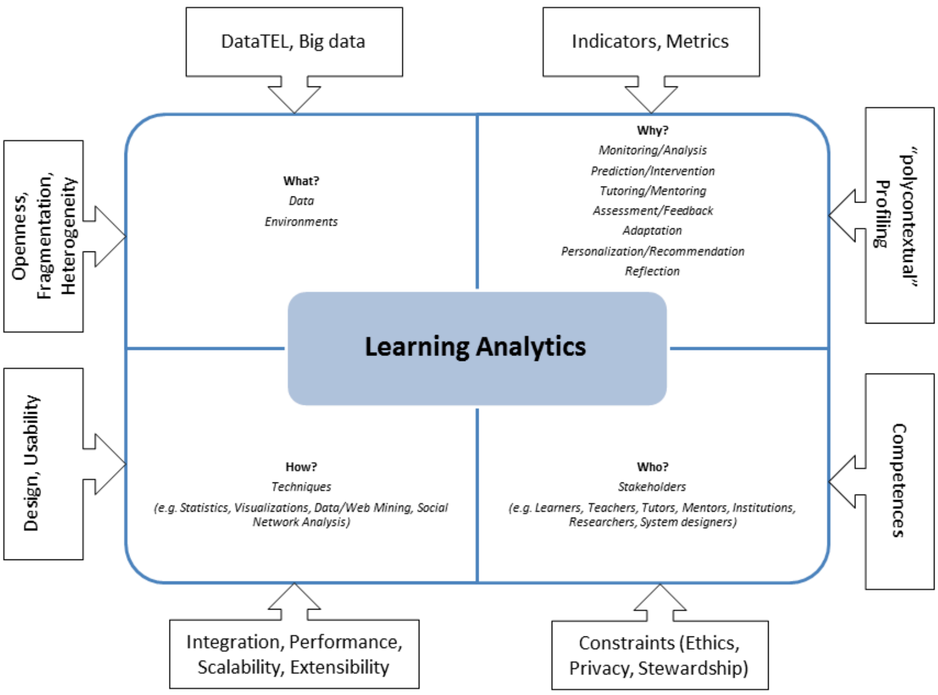


Figure 3. Learning Analytics Reference Model (Chatti et al., 2012)

This reference model is going to be utilized in our requirement analysis as it provides a classification schema on software components in our ongoing development of reference software architecture.

2.4 Standardization Activities

As the importance of the interoperability of learning analytics is recognized in Korea, a three-year-long research project funded by Telecommunications Technology Association (TTA) of Korea is recently launched to build *a reference model of learning analytics based on educational content and unstructured data*. In this project, Korea Education and Research Information Service (KERIS) runs the project in close coordination with ISO/IEC JTC 1/SC 36 and IMS Global.

ISO/IEC JTC 1/SC 36, Information Technology for Learning, Education and Training is a standardization subcommittee (SC), which is part of the Joint Technical Committee ISO/IEC JTC 1, that develops and facilitates standards within the field of information technology (IT) for learning, education and training (LET) to support individuals, groups, or organizations, and to enable interoperability and reusability of resources and tools. Recently, SC 36 established an *Ad Hoc Group on Learning Analytics Interoperability* to ascertain the necessity of standards to facilitate interoperability among diverse learning analytics components such as data collection, analysis, privacy protection, qualification of data and accessibility. We expect to submit the result of requirement analysis to this Ad Hoc Group as contributions for assessment, and anticipate to participate in the future Working Group, if established, with the reference software architecture.

3. Related Work

3.1 IMS Global

IMS Global Learning Consortium (usually referred to as IMS GLC, IMS Global or simply IMS) is a global, nonprofit, member organization that strives to enable the growth and impact of learning technology in the education and corporate learning sectors worldwide. IMS GLC members provide leadership in shaping and growing the learning industry through community development of interoperability and adoption practice standards and recognition of the return on investment from learning and educational technology. Their main activity is to develop interoperability standards and adoption practice standards for distributed learning, some of which like Learning Tools Interoperability (LTI), Question & Test Interoperability/Accessible Portable Item Protocol (QTI/APIP), Common Cartridge, Learning Information Services and Content Packaging are very widely used (Wikipedia, 2014a).

Especially, IMS Caliper (IMS Global Learning Consortium, 2013) is a work in progress to define a learning measurement framework using existing IMS specifications to provide a standardized representation, capture, and marshaling of learning activity generated metrics targeted for consumption by any conforming sensor API endpoint enabled analytics store/ services. IMS Caliper is built around the following concepts:

- **IMS Learning Metric Profiles** that provide a Learning Activity centric focus to standardize on metrics (actions and related context) captured across consumer and producer learning tool's delivery activities and delivery platforms that consume and orchestrate activity based curriculum, while providing for custom extensions and future additions to the profiles;
- **IMS Learning Sensor API and Learning Events** drive standardized instrumentation and metric capture and marshal between tools and their delivery platforms and/or associated analytics service solution aggregating metrics;
- **IMS LTI™/LIS/QTI™ leverage and extensions** enhance and integrate granular, standardized learning measurement with tools interoperability and the underlying learning information models, inclusive of course, learner, outcomes and other critical associated context.

In a recent IMS publication, "Learning Measurement for Analytics Whitepaper" (IMS Global Learning Consortium, 2013), they claim standards for learning analytics are required so they can be combined across all of the educational sources by asserting that *"equipped with a standards based common foundation for learning measurement, the quality, efficacy and performance derived analytics for the online curriculum across the ecosystem can be achieved more effectively."*

3.2 UNESCO IITE

UNESCO Institute for Information Technologies in Education (IITE) identifies three kinds of predictors and indicators, and two kinds of interventions as follows.

- **LMS/VLE Analytics Dashboards**: The first kinds of analytics that many institutions will encounter will be the analytics dashboards now appearing in most online learning platforms., Data logs are now rendered via a range of graphs, tables and other visualizations, and custom reports designed for consumption by learners, educators, administrators and data analysts.
- **Predictive Analytics**: From the pattern of learners' static data (e.g. demographics; past attainment) and dynamic data (e.g. pattern of online logins; quantity of discussion posts) one can classify the trajectory that they are on (e.g. "at risk"; "high achiever"; "social learner"), and hence make more timely interventions (e.g. offer extra social and academic support; present more challenging tasks).
- **Adaptive Learning Analytics**: Adaptive learning platforms build a model of a learner's understanding of a specific topic (e.g. algebra; photosynthesis; dental surgical procedures), sometimes in the context of standardized tests which dictate the curriculum and modes of testing. This enables fine-grained feedback (e.g. which concepts you have grasped and at what level), and adaptive presentation of content (e.g. not showing material that depends on having mastered concepts the learner has failed on).
- **Social Network Analytics**: Social network analysis (sometimes called Organizational Network Analysis in corporate settings) makes visible the structures and dynamics of interpersonal networks, to understand how people develop and maintain these relations. People may form 'ties' of different sorts, ranging from extended, direct interaction reflecting significant ties, to more indirect ties.
- **Discourse Analytics**: Analytics could go beyond simple quantitative logs, and provide feedback to educators and learners on the quality of the contributions. Researchers are beginning to draw on extensive prior work on how tutors mark essays and discussion posts, how spoken and written dialogue shape learning, and how computers can recognize good argumentation, in order to design analytics that can assess the quality of text, with the ultimate goal of scaffolding the higher order thinking and writing that we seek to instill in students.

4. Requirements of the Reference Software Architecture

In this section, we present preliminary results of requirement analysis of the reference software architecture based on the survey described in the previous sections, Section 2 and 3. Even though, the requirements are multifold over data, analysis, and application requirement, the results can be summarized as design requirements of reference software architecture as follows:

- **Open and extensible**: It should be open to incorporate new sensors or analytics functionality, desirably without interrupting the task being serviced. It also should ensure incorporation or modification of new workflows at the task level.
- **Distributed**: It should be able to handle multiple sources of data and functionalities distributed over multiple systems. It is also desired to be able to distribute data and to delegate functionality dynamically and transparently.
- **Interoperable:** It should provide compatibility for various learning platforms or VLE by providing interoperable interface to the data and operations.
- **Reusable and configurable**: The functional components and data interfaces should be modular and thus reused and configured for different tasks or more complex tasks as building blocks.
- **Real-time and predictable**: Learning analytics should be performed satisfying the real-time constraints and should be able to estimate the time to completion.
- **Usable:** It should acceptable user experience (UX) by providing appropriate data visualization and user interfaces for monitoring and tasking throughout the learning analytics process.
- **Secure and traceable:** It should protect personal user information to secure privacy and preserve confidential information. Some analytics functionality should be ensured not to be performed as

required. Furthermore, the history of execution of analytics functions and access to data should be recorded, if needed, to ensure traceability.

5. Reference Software Architecture

A reference software architecture is a software architecture where the structures and respective elements and relations provide templates for concrete architectures in a particular domain or in a family of software systems. A reference architecture often consists of a list of functions and some indication of their interfaces (or APIs) and interactions with each other and with functions located outside of the scope of the reference architecture (Wikipedia, 2014b).

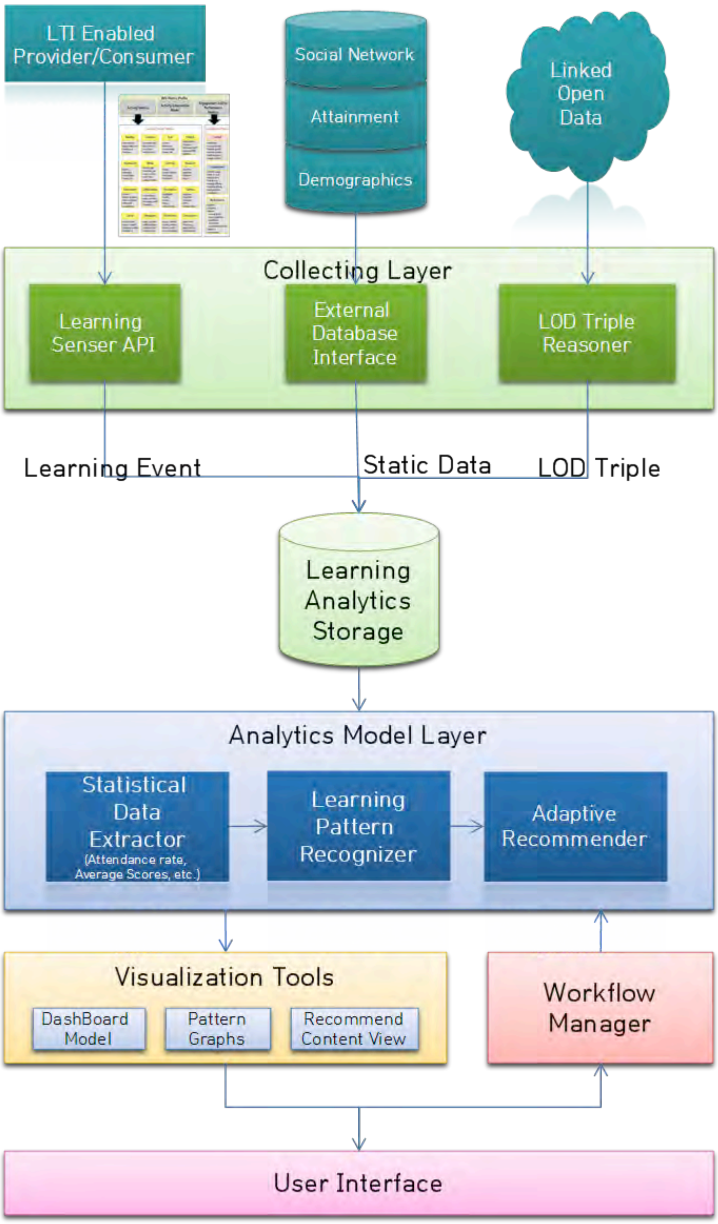


Figure 4. a preliminary reference software architecture

In this section, we present our preliminary reference software architecture as shown in Figure 4. The main purpose of this initial work is to identify the necessary components and data interactions of the components before the results of the requirement analysis are fully applied. The initial architecture is then instantiated with supporting software tools to validate the effectiveness. Table 1 summarizes the basic input and output interfaces for each component and applicable software to implement the functionality of the component.

Table 1: Specifications of the Reference Software Architecture

Component	Description	Input	Output	Applicable Software
Learning Sensor API	External API to Collecting Learning Activity Data	Metric Profile	Learning Event Data	Apache Storm
External Database Interface	Collect Structured Data from RDB or Web Services	Well-Formatted Data		Apache Sqoop
LOD Triple Reasoner	Collect LOD and extension Triple Data	LOD Triple	LOD Triple	Webpie
Statistical Data Extractor	Calculate Statistical Information	Literal data set	Analyzed Information	*R*
Learning Pattern Recognizer	Recognize Specific pattern by Data Mining	Literal Data Set, Analyzed Information		*R*
Adaptive Recommender	Recommend Associated Content	Analyzed Information	Literal Data	*R*
Visualization Tools	Visualize each Types of Analyzed Data	Analyzed Information	HTML	Google Charts

6. Discussion and Future Work

Realizing learning analytics systems indispensably requires an integrated and holistic approach. In this paper, we first identified the basic requirements of reference software architecture to capture the holistic and integrated view. The primary requirements includes the system to be open, distributed, interoperable, reusable, real-time, usable, and secure. So far these requirements were derived from the survey of the state of the art in learning analytics field, but they need to be refined and applied to the reference software architecture by the real use cases and application scenario. We are currently collecting such use cases and application scenarios of learning analytics. We are well aware that our goal to develop the explicit specifications of the learning analytics architecture is far away from this preliminary work, but we believe this is an essential step toward the goal.

Acknowledgements

This research was supported by the ICT Standardization program of MISP (The Ministry of Science, ICT & Future Planning)

References

Buckingham Shum, S. (2012). Learning analytics. *IITE Policy Brief*.
Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5), 318-331.
IMS Global Learning Consortium. (2013). Learning Measurement for Analytics Whitepaper.
Wikipedia. (2014a). IMS Global. 2014, from http://en.wikipedia.org/wiki/IMS_Global
Wikipedia. (2014b). Reference Architecture. 2014, from http://en.wikipedia.org/wiki/Reference_architecture

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Learning Analytics Interoperability – looking for Low-Hanging Fruits

Tore Hoel [a*] & Weiqin Chen [b]
[a]*Oslo and Akershus University College of Applied Sciences, Norway*
[b]*University of Bergen, Norway*
*tore.hoel@hioa.no

Abstract: When Learning Analytics is seeking a wide community, the challenge of efficiently and reliably moving data between systems becomes important. This paper gives a summary of the current status of Learning Analytics Interoperability and proposes a framework to help structuring the interoperability work. The model is based on a three dimensional Enterprise Interoperability Framework mapping concerns, interoperability barriers and potential solutions. The paper also introduces the concept of low-hanging fruits in prioritising among solutions. Data gathered from a small group of Norwegian stakeholders are analysed, and a list of potential interoperability issues is presented.

Keywords: Interoperability, Standards, Learning Analytics, Educational Data Mining, Sharing Data Sets

1. Introduction

Learning Analytics (LA) is an emerging research field where we are starting to see contributions from a diversity of research disciplines, and development of a range of tools, techniques and applications used by LA researchers and practitioners (Siemens, 2013). However, large-scale implementations of LA in an educational sector, a region, an industry, – or even in an institution, remain to be seen. Scaling up LA means to go beyond research prototypes or the innovative solution of a single vendor who keeps the data under tight control within a closed ecosystem, the analytics magic in a black box, and only exposes the results to the users in colourful dashboards. Unless scaling up means 'winner takes all', we need to address a range of new issues posed by the needs of actors, systems, organisations, and cultures to interoperate.

Data lies at the heart of learning analytics. This does not necessarily mean that data sharing and interoperability has been a main concern for LA research or development till now. Interoperability involves different aspects of how systems at large (both organisations and ICT systems) communicate on different levels (e.g., technical, semantic, organisational, political, and legal). New challenges are posed when scenarios foresee third party LA tools analysing data from diverse sources by national and international organisations sharing and comparing data. In addition, moving from prototypes to large-scale implementations opens up a raft of new issues, – organisational capacity and privacy being only two of them (Scanlon et al., 2013; Siemens, 2013).

We suggest using the concept of interoperability as an overarching term for this new level of discourse on scaling up applications of LA. By doing so, we bring a new set of actors to the table, underlining that user groups, implementers, standardisation experts and bodies, local authorities, and others have a role to play in order to reap the benefits of bringing analytics to education (MacNeill & Campbell, 2014). Interoperability as a term will be discussed below. However, we would also suggest to apply the concept of "low-hanging fruits" in framing the discourse on LA interoperability, as "the way ahead to get results sometime soon requires care (..) a middle way seems necessary, in which a little time is spent on discussing the most promising and the best-understood targets, i.e. to look for the low hanging fruit" (Cooper, 2013b).

This proposal to identify low-hanging fruits addresses the problem of how to conceptualise the solution space for learning analytics. Our background in the standards community has made us wary of big and "complete" designs never leaving the researchers' drawing pad. LA have the potential to associate numbers with any aspect of the learning process, and as a consequence requires immensely complex data models for exchange of information. With ambitions of large-scale implementations based on an ill-defined problem space, a piecemeal and lightweight approach might be more advisable (Hoel, 2014a; Sales et al., 2012). Consequently, there is a need to find a way to identify the low-hanging fruits.

In this paper we have carried out a pilot study exploring the Learning Analytics Interoperability (LAI) problem space by interviewing a small number of representatives of LA stakeholders. The rest of this paper is organised as follows: First, the concepts of interoperability and low-hanging fruits are reviewed. Then a small explorative study of stakeholder groups' views on interoperability in the context of learning analytics is presented. The results are discussed in relation to an Enterprise Interoperability Framework, searching for approaches to interoperability that could be characterised as low-hanging fruits.

1.1 Learning Analytics Interoperability

A search in Google Scholar on 'learning analytics' AND 'interoperability' gives in mid-2014 just above 400 hits; while searching for 'learning analytics' AND 'data sharing' gives less than 100 hits. In 2013, Cooper surveyed academic and formal publications as well as informal publications and noted that "only a small group of people, largely researchers, have drawn attention to LAI and a significant amount of the literature has been produced by a few people" (Cooper, 2013a). He also found no references to LAI from software suppliers.

> *"The way LAI is covered by these works will be identified as being of three kinds: assertion or argument in favour of interoperability in general; references to interoperability for a particular purpose or context; interoperability as a significant or key topic. Assertion and argument about interoperability are usually concerned with the lack of it" (Cooper, 2013a).*

There has been little work on interoperability specifications by the educational technology community; first in August 2014 the ISO committee working on learning technology standards established an ad hoc group to develop scope for new work items on LAI.

Interoperability is a multidimensional term with many interpretations and definitions. According to the Institute of Electrical and Electronics Engineers interoperability is "the ability of two or more systems or components to exchange information and to use the information that has been exchanged" (Geraci, 1991). Cooper (2014) states that "a broad interpretation of "systems" that includes people and the activities they undertake using these digital technologies captures the true essence of interoperability as a means to achieve human aims and objectives". However, without describing the different dimensions of interoperability the term tends to get a merely technical interpretation, leading to a focus on exchange of data when there is a need to zoom out and look at the social, political and organisational motivators and barriers to interoperability. We would suggest that a perspective inspired by Enterprise Interoperability (EI) should be applied at this early stage of exploring LA Interoperability (LAI) challenges. In the EI setting interoperability is defined as the "ability to (1) communicate and exchange information; (2) use the information exchanged; (3) access to functionality of a third system" (Chen & Daclin, 2006).

Applying an enterprise perspective to interoperability foregrounds the two dimensions that make up the problem space (barriers and concerns), and highlights the need to explore the solution space looking into alternative approaches, as illustrated in Figure 1 (Chen & Daclin, 2006). Analysing the barriers it will make sense to group them in the broad categories of conceptual, technological and organisational barriers. The concerns however, need to be derived by studying the domain characteristics of education as a particular instance of an enterprise. The dimensions identified in the ATHENA Interoperability Framework (http://athena.modelbased.net) with concerns related to data, services, processes and business, might help the analysis.

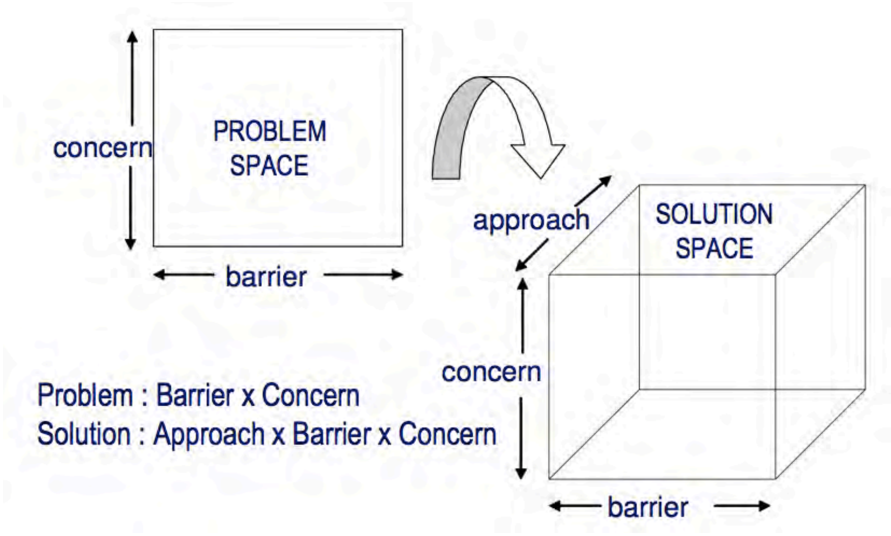


Figure 1. Problem space vs. Solution space (Chen & Daclin, 2006)

A key task in the analysis of the problem space and the solution space will be to identify LA solutions that address stakeholder concerns and overcome interoperability barriers. When the solution space is mapped, the next step is to choose strategy; and it is here we will introduce the concept of low-hanging fruits.

1.2 Low-Hanging Fruits

By using tools and data that are already in place the community could benefit right away from the development of new knowledge and new designs. This is the basic idea behind the approach of reaping the low-hanging fruits. To extend the fruits metaphor, one should refrain from extensive pruning (e.g., changing the context or the system) until the gardener knows more about the trees and the garden. Within Learning Analytics and Educational Data Mining this may make sense, since it is difficult getting data out of information systems. However, the hunger for tasting the benefits of LA is great; the potential data sources are diverse; and the range of methods and experience is growing (Cooper, 2013b). By going for the low-hanging fruits we allow stakeholders time to argue their case for specific LA solutions before deciding on approaches with far-reaching implications.

2. Related work

In 2011 SOLAR, the Society for Learning Analytics Research, issued a proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques under the name of Open LA (Siemens et al., 2011). The proposal was a high level argument in favour of openness of process, algorithms and technologies; and modularized integration, – asking for development of common language for data exchange and open repository of anonymised data. It is still early days to deliver on this proposal; in 2014 a follow up meeting was organised where SOLAR joined forces with the Apereo Foundation, an umbrella organisation for a number of open source projects. Their aim for a LA Initiative is now to “accelerate the operationalization of Learning Analytics software and frameworks, support the validation of analytics pilots across institutions, and to work together so as to avoid duplication” (Cooper, 2014a).

LA Interoperability initiatives are also launched by the standards community. Since 2010 the Advanced Distributed Learning (ADL) initiative have developed an eXperience API (xAPI), also called TinCan API, based on the idea of tracking activity streams (ADL, 2014). A similar approach is adopted by IMS Global in their Caliper project (IMS, 2013), initiated late 2013. When ISO/IEC JTC 1/SC36 late 2014 starts to work on LAI it is assumed that they will begin by defining an abstract framework in what eventually could become a multipart international standard (Hoel, 2014b).

The Open LA and the xAPI initiatives represent the opposite parts of the LAI continuum. While the former is more an interoperability dream, the latter represents a very concrete approach to exchange data on any activity that is related to learning, storing statements of the form “I did this”, linking an actor to an object via a verb. However, the main parts of the LAI continuum are still to be addressed. Cooper (2014) has reflected on the “big picture” of LAI exploring what should be the scope of work in this field. He identified three areas of discourse defined by these questions:

- **Models and Methods:** How can we transfer information about statistical and data mining methods, the parameters used, and the predictive models?
- **Analytical Results:** How can we transfer individual-level and grouped numerical results? How can we track data provenance, quality and processing methods?
- **Data for Analysis:** How can we get data out of the operational systems? (Cooper, 2014)

One takeaway from Cooper’s briefing (2014) is that LAI is very complex and involves interoperability specifications that are not generally known within the educational technology community. As an example, Cooper mentions PMML, the Predictive Model Markup Language (dealing with interoperability of models and methods); and SDMX, the Statistical Data and Metadata eXchange standard (dealing with interoperability of analytical results). To further LAI one needs to invest in knowledge building to deal with this complexity, which is not only of technical nature, but also have issues concerning consensus about the objectives of LA as a whole.

3. Soliciting stakeholders’ interoperability requirements

When Big Data is promoted by global consultancy firms as “the next frontier for innovation, competition and productivity” (Manyika et al., 2011), and research is pointing to large datasets as the key to improving learning and the environments in which it takes place (Ferguson, 2013), it is essential to solicit the views of stakeholders who play vital roles in delivering and analysing the data and using the results. Data sharing and interoperability between IT systems are about tearing down unwanted barriers. It is important to note that this is only one perspective, and that the flip side of barriers could be boundaries people have established to protect themselves. We only know what perspectives will influence design and implementation of LA applications when we have involved the stakeholders.

This paper presents a pilot study aiming to structure the discourse on interoperability and give input to scoping of the first work items for standardisation in this field. We have interviewed eight representatives of students, teachers, support staff, and policy makers in Norway asking them to elaborate on interoperability in the context of LA. The semi-structured interviews focussed on what are the things to agree upon (that need to interoperate) to realise the potential of LA, aims of LA applications, and what data sources could be used for LA. The interviews were supported by an online sticker board, resulting in graphical summaries of the interviews giving a rough sense of the priorities of the respondents.

This research is positioned in the first Relevance Cycle of the three research cycles of Design Science (Hevner, 2007; Hevner et al., 2004), addressing requirements and field testing. The purpose is to come up with candidate concepts that describe the problems and opportunities in the application domain from a people, organisational systems, and technical systems perspective. The study is offered as an approach that could be replicated within other regional or national communities or sectors in order to gather requirements for LAI.

4. Results

The first results are an unordered list of concerns of the respondents being asked to reflect upon LA. The interviews are by no means representative for Norwegian communities or for the different stakeholder groups. Nevertheless, this study indicates that it is still early days for the idea of using LA in schools and higher education. There is little experience with LA solutions, and there is a general need for overview and understanding of the basic ideas. Even if the concerns reflected the respondents’ role in education, the issues they brought forward reflected more or less the same views on the problem space. Support for the individual learner, and emphasis on privacy and control of the data generated through interaction with the systems were in the foreground of every interview. The concerns mentioned were:

Student interviews

- Personal development and support for learning and career planning as primary aim; helping the institutions, e.g., to improve retention, as a secondary aim
- Non-intrusive guidance (not being evaluated & tested all the time)
- Privacy & Control over personal information
- Trust - school or university as a trusted "partner" in LA
- Consent to allow data flow between systems - transparency to who sees what
- Access to data for LA: Students have a mixed use of tools that do not exchange data
- Educational tools policy: More emphasis on institutional tools, like LMS, will prevent students from using social networking tools for learning. Latter group of tools important for life-long learning.
- Control over LA results: Students want to be in control of interaction with their data - ownership to analysis and results, not only to data
- Data should be open, based on agreements between the partners involved
- Coordination and coherence among services: Students enrolled in different courses, learning on different platforms - in order to get a coherent picture teachers, data, etc. need to be coordinated
- LA solutions should allow for two-way interaction, user control, consent, time to think before giving data away, etc.
- Non-intrusive LA (no extra time on using LA tools)

Teacher interviews

- Understanding the affordances of LA: To get a conceptual understanding of the domain, from different perspectives, not only technological aspects
- Prioritising benefits in this order: Individual (adaptive), teaching, institutional / organisational
- Making sense of 'contexts' and 'activities': One cannot make sense of data, unless one knows their context. How to describe contexts and activities?
- LA implementation without losing control of «pedagogy» (technology or market driven vs. pedagogically motivated innovation)
- Learner Control and ownership to data; and control over how data are used, e.g., through anonymisation / pre-processing (removing personal information)
- User Control over tools & services
- Getting the statistics out of the tools that are currently used
- In listing potential sources of LA data, institutionally controlled data sources, e.g., LMS, are mentioned first

Support staff interviews

- Interpretations of LA results: What do we measure - and what does the results mean related to the different stakeholders' use of the LA results?
- Control over data: Do students have the right to reserve themselves against participating in LA, sharing data from LA, etc.?
- Agreeing about contexts for analytics. What to do with 'context free data' that make no sense for analysis?
- Primary LA beneficiaries should be the learner
- Need to improve the interoperability of legacy systems in order to get data for LA
- Need to agree upon realistic aims for the use of LA
- Promote institutional control of data generated in (cross-institutional or international) MOOC systems
- Develop and introduce systems that give enough data to allow LA, e.g., MOOCs
- Start using non-controversial data - e.g., data showing if watched videos are too long, before using data identifying the individual

Policy maker interviews

- Creating a culture for LA - mapping the incentives to make use of LA and develop a strategy
- Privacy and Ownership to data - Open Badges approach to data (owning your own data)
- Interpreting data: How to avoid measuring the wrong data and making invalid conjectures?
- The ultimate aim is to get empirical support for pedagogical choices, improve quality of learning resources, and further adaptive learning. However, also institutional aims are important, e.g., better retention and early warnings of drop outs
- We should discuss who should store the data; should it be nationally controlled or distributed?

Concerns can be grouped along a continuum starting with data in the bottom row, climbing up via Tools and Technology Support (Services), Learning Activities (Process), and ending with Aims (Business) at the top row. We have chosen these dimensions as a refinement of the categories used in Enterprise Interoperability Analysis (Chen & Daclin, 2006). This tradition looks at Conceptual, Technological and Organisational barriers as the categories for the barrier axis fitting well with the investigation of our study.

4.1 Problem space

Table 1 presents the results of populating the LA problem space with the data from the stakeholder interviews. The numbers in the table both refer to the concern/barrier nexus described in the text below, and give an indication of urgency or priority gleaned from repeated rounds of qualitative analysis of the interviews by two independent researchers (1 representing the highest urgency/priority).

Table 1. Learning Analytics Problem Space (concern and barrier matrix)

Concern/Barriers	Conceptual	Technological	Organisational
Aims			(1)
Learning activities	(3)		(2)
Tools and Technology support			(5)
Data		(4)	

(1) **Privacy, Trust & Control of Data:** This nexus between enterprise aims and organisational barriers relates to the complex issues of how interoperating LA systems get access to data without violating the privacy of users; and how to maintain legitimacy of these systems while giving the users control of their data.

This problem is situated in the most abstract corner of this two dimensional space. It relates barriers to interoperability on all levels. However, there is no conceptual or technological fix to this problem; it is clearly up to organisations and their members to agree upon questions like how much private information has to be exposed in order to reap the benefits of more adaptive systems, more support to learning, etc.; what institutions within education can be trusted to manage personal information for which groups of learners, with what kind of procedures; and what kind of control will the system give the originator of data throughout the LA cycle. And perhaps most importantly, what aims should have priority, e.g., if there are conflicts between the aim of an institution to reduce drop out and the privacy of the student?

(2) **LA affordances and application domains:** This interoperability problem arises when there is no consensus about the benefits of LA, and what domains LA should be applied to.

This problem is related to strategies for policy development and implementation for institutions, sectors and governments. Even if the barriers are organisational, at this early stage of LA conceptual barriers (e.g., lack of shared vocabularies) are part of this problem.

(3) **LA Context & Learning Activities:** This is the “blind data” problem that arises when there is no context information provided with data from a learning activity.

This is a conceptual barrier due to the lack of linkage between learning activity streams and their pedagogical contexts.

(4) **Legacy system interoperability - information model for LA data exchange:** This is the classic learning tools interoperability problem where systems have data in silos without any possibility of aggregating data to get a coherent view of the activities in a class or a school.

When LA is added as yet another system this data integration problem is brought to the attention again.

(5) **LA implementation best practice guide:** This interoperability problem relates to the market vs. policy driven implementation of LA systems and lack of open institutional or regional LA policies.

Technical interoperability is not always first priority for an enterprise developing an innovative LA solution. Organisations may find they need guidance to best practice for implementing LA solutions in order to support other educational policies.

4.2 Searching for solutions and low-hanging fruits

Having mapped the problem space, the next step is to develop the solution space, adding a new approach dimension to our model. As explained in Figure 1, a Solution Space is formed when Approach, Barrier, and Concern intersect. At this stage we are not searching for *any* solution but the one that can be picked as a low-hanging fruit.

Our interviews only indirectly pointed to solutions. However, the results reported in Table 1 give a prioritisation of problems, which is a first step towards designing an approach. Our data clearly shows that the non-technical issues are seen as the most imminent and important problems that could jeopardise uptake of learning analytics. The barriers are mainly organisational and conceptual, with the only identified technological problem being related to legacy systems.

The following approaches are based on a second analysis of our data to extract possible ideas for solutions.

Privacy, Trust & Control of Data

Strictly speaking, this complex set of issues is not related to LA in particular. Some experts may even say it is out of scope for LAI. Our respondents, however, see this issue as major stumbling block that needs to be dealt with, even before discussing the potential benefits of innovative LA solutions now being marketed. It is also an issue that is manageable, at least from a conceptual and technological point of view. The challenge is to engage in this work a new group of stakeholders (e.g. teachers, policy makers) whose primary interest is not with technology development.

Privacy laws are in place for schools and universities. Too strict interpretations may stifle new use of technologies and new learning practices as seen in reluctance to use non-institutional controlled services (e.g., cloud services). There is a need to clarify rules and practices (e.g., lay down principles for which services to trust). This may prove difficult, as the new learning practices that push the boundaries away from institutionally controlled, teacher-led education towards learner-centred, socially situated life-long learning may challenge existing privacy and data protection paradigms.

Our respondents pointed to learner control over data and teacher control over pedagogy as principles to pursue. We will come back to this point in the Discussion section of this paper.

LA affordances and application domains

Our respondents see the benefits of LA, but have a lot of questions mixed with fears about potential adverse effects. Are we measuring the right things? Do we discern between causation and correlation? Are we able to draw the right conclusions from the LA dashboards? To improve interoperability between LA systems one needs a more granular picture of LA activities. Academic analytics to prevent dropouts from school is a different activity to Learning analytics capabilities built into a digital textbook to support adaptive learning. As one respondent put it, “if we only have data for when a video loses the audience (not knowing it was Peter), we should focus on improving the video, not dreaming of supporting a particular person”.

The interviews requested more conceptual clarity to the field of LA. This is a feasible task that could be organised as a consensus process.

LA Context & Learning Activities

The approach to solve the problem of missing context descriptions in relation to learning activities is to launch a traditional standardisation project developing the necessary vocabularies and find ways to describe relations among the concepts. These are continuous activities within learning technology development. However, the new interest in LA may contribute to speed up this activity and bring new stakeholders on board.

Legacy system interoperability - information model for LA data exchange

LA could potentially be a driver for revisiting the interoperability problems due to siloed legacy systems. However, as one respondent said, nobody is going to rebuild the student information system that is working well for all higher education institutions in Norway. The solution may be in establishing some kind of aggregated system, raising the question of who should run such a system, and how should the data be stored. Again, in order to make progress on this problem, the first problem cluster discussed above needs to be sorted out.

LA implementation best practice guide

To lead by example is a good principle. A best practice guide may help institutions to implement LA, as such guides need to be clear about what the actors of the implementations should be. Developing a guide as a consensus document may help the different levels of the educational system to identify other interoperability problems and to prioritise among them. Starting in the nontechnical end of interoperability work may first, help address the concerns of the stakeholder groups represented in our survey; and second, take the focus away from technical LA challenges, making it easier to address the semantic, organisational, political and legal interoperability problems posed by putting numbers in the service of learning, education and training.

5. Discussion

Traditionally, ISO/IEC JTC 1/SC36 had published a number of multipart standards within the learning technology domain. When launching an LA project, this international standardisation group would be expected to fall into the old pattern starting with a first part being a LA Framework standard, and the other parts filling the different puzzles of an all encompassing LA jigzaw. This paper can be read as a warning against such an approach. The explorative interviews we have reported make it clear that detailed information models alone will not ease the uptake of learning analytics in schools and higher education. In order to foster interoperability among actors in this sector there is a need to find solutions on all levels where two systems interoperate. The big question, however, is to find which puzzle to start with that will make it easier to see the pattern and find the solution to the other pieces.

We have suggested designing a solution space by soliciting input from stakeholders that do not necessarily know much about LA, but who eventually will play a crucial role in its adoption. In this study, the respondents highlighted above all the softer issues related to privacy, trust and control. It is worth exploring if the solution space related to this issue could be established as a kind of baseline for further design of LA systems.

Often the Big Data hype is used to sell Learning Analytics (Ferguson, 2012), triggering stories of aggressive and subtle marketing and manipulation from a commercial context. LA needs to distance itself from the setting of marketing and sales, as learning and education have very little in common with the motivational arm-twisting of commerce. From a liberal, and some may say Western perspective, the ethos of learning and education is that the learner should be in control and the supporting institutions should only do what is in the interest of the learner. Therefore, it should be easy to argue for learner control over her own data; transparency controlled by the individual learner; trust built bottom-up; etc. The counterargument would be that such an approach will not give the amount of data needed for LA, as only centrally run systems where all are subscribed would give enough data. However, in a civic society the individual has some rights to opt out of education. And privacy protection and trust are not opposites. Trust does not need to be blind; it can be a dynamic property, ultimately controlled by the learner, but also maintained by institutions, e.g., the school, the university or the educational authorities.

What happens if the learners (and their parents) lose trust in LA systems is vividly illustrated by the US InBloom case:

> *Protests began in earnest when it was discovered that inBloom's software had more than 400 optional data fields that schools could fill out—asking for potentially sensitive information such as the nature of family relationships, learning disabilities, and even Social Security numbers. Although there were no reported leaks, parents were uncomfortable without an absolute guarantee of that data's safety or a clear indication of who could access it. (Slate future tense, 24 April 2014)*

InBloom "was meant to extract student data from disparate school grading and attendance databases, store it in the cloud and funnel it to dashboards where teachers might more effectively track the progress of individual students" (New York Times, 21 April 2014). In April 2014, after a period of heated public debate, the system had to close down after the New York state passed legislation prohibiting the state department of education from giving student information to data aggregators like inBloom.

Giving priority to the solution of privacy, trust and control issues could help identifying the LA systems with the best possibilities to succeed within the nearest timeframe. Trust is built in concentric movements starting with the learners, co-learners, school, other community of learners, etc. Local, distributed, transparent and adaptive systems supporting the learner seem to be easier to sustain than systems that are more distant and leave the user with more questions of who is in control. On the other hand, we know that learners use cloud services and social media systems where they have minimal control.

If the systems are found useful, they tend to be used and the users freely give access to their data. In the case of social media, however, the educational institutions are not acting as intermediators between the users and the systems. In formal education, institutions have to follow standards, and it is therefore problematic to mandate use of tools with poor or unknown data protection policies. It would therefore be helpful to have a consensus about how tools outside institutional control are used and what privacy, trust and control models education would promote (Slade & Pinsloo, 2013). This is an argument for engaging in a process on clarifying LA affordances and application domains, the second approach coming out of our stakeholder survey.

Establishing learner control as a design baseline would help identify which LA contexts that need specification. One might assume that smaller systems that are able to demonstrate benefits to the learner would be easier to introduce, and as such represent the low-hanging fruits of learning analytics. While ideas about more complex and institutionally motivated systems, e.g., with institutional, regional or even national learning record stores, should be left to ripen before brought to standardisation. It is also reasonable to think that once the idea of complex and integrated systems are put on the back-burner, new ideas could be foreseen how analytics can be carried out to improve education with existing data and systems. In an emergent field there is a need to showcase and demonstrate best practices that work before investing too heavily in wild dreams.

6. Reflections and Outlook

This research was conducted in accordance with design science guidelines (Hevner et al., 2004) to develop a support framework for structuring work on LAI. Stakeholders were interviewed about their concerns about LA in order to construct a problem space along the dimensions of concerns and interoperability barriers. The interview data was then analysed in order to identify approaches that could be dealt with within a reasonable timeframe, given the dynamic nature of current learning analytics development. The analysis gave five candidate issues that are potential new work items for LAI standardisation.

The results of this study need to be validated through further field-testing in order to see if the same issues are prioritised by more representative selections of stakeholders and respondents from other countries. Furthermore, the relevance of the barriers used in this study should be tested. It is possible that a more fine-grained categorisation may be needed, especially to understand the technological barriers to LAI.

The concept of low-hanging fruits should also be further developed. In this paper we have used the concept to add a strategic dimension to the approach axis of the solution space. We have also indicated that selecting the low-hanging fruits may alter the space itself, giving priority to a certain group of applications, repurposing others. In developing this framework related existing work on quality models for standards looking into aspects of product quality, process quality, and quality in practice (Folmer, 2011) should be explored.

References

ADL (2014). Training & Learning Architecture (TLA): Experience API (xAPI). Retrieved September 1st, 2014, from www.adlnet.gov/tla/experience-api

Chen, D., & Daclin, N. (2006). Framework for enterprise interoperability. Proc of IFAC Workshop EI2N.

Cooper, A. R. (2013a). Learning Analytics Interoperability - a survey of current literature and candidate standards. Online at http://blogs.cetis.ac.uk/adam/2013/05/03/learning-analytics-interoperability, accessed 2014-08-01

Cooper, A. R. (2013b). Learning Analytics Interoperability – some thoughts on a "way ahead" to get results sometime soon. Online at http://blogs.cetis.ac.uk/adam/2013/10/17/learning-analytics-interoperability-some-thoughts-on-a-way-ahead-to-get-results-sometime-soon, accessed 2014-08-01

Cooper, A. R. (2014). Learning Analytics Interoperability – The Big Picture in Brief. Online at http://laceproject.eu/publications/briefing-01.pdf, accessed 2014-08-01

Cooper, A. (2014a). Open Learning Analytics – progress towards the dream, blog post of 2014-04-14, online at http://www.laceproject.eu/blog/open-learning-analytics-progress-towards-dream/, accessed 2014-08-02

Geraci, A. (1991). IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries. IEEE Press, Piscataway, NJ, USA.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4(5/6), 304. doi:10.1504/IJTEL.2012.051816

Ferguson, R. (2013). Learning Analytics for Open and Distance Education, 1–8.

Folmer, E. (2011). Quality model for semantic IS standards. 8th International Conferanse "Standarization, Prototypes and Quality: a Means of Balkan Countries' Collaboration," Thessaloniki, October 2011.

IMS (2013) IMS Global Learning Consortium, Learning Measurement for Analytics Whitepaper, September 19, 2013. Retrieved September 1st, 2014, from http://www.imsglobal.org/IMSLearningAnalyticsWP.pdf

Hevner, A. (2007). A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems, 2007, 19(2):87-92, 1–6.

Hevner, A., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. Mis Quarterly, 28(1), 75–105.

Hoel, T. (2014a). LA on the agenda for international standardisation, blog post of 2014-07-4, online at http://www.laceproject.eu/blog/la-agenda-international-standardisation/, accessed 2014-08-02

Hoel, T. (2014b). Standards as enablers for innovation in education - the breakdown of European pre-standardisation (pp. 185–189). Presented at the 2014 ITU Kaleidoscope: Living in a Converged World - Impossible without Standards? (K-2014), IEEE. doi:10.1109/Kaleidoscope.2014.6858496

MacNeill, S., & Campbell, L. M. (2014). Analytics for Education. Journal of Interactive Media in Education. Online at http://jime.open.ac.uk/jime/article/viewArticle/2014-07/html

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., et al. (2011). Big Data: The next frontier for innovation, competition and productivity. McKinsey report online at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation, accessed 2014-08-01

Sales, B., Darmois, E., Papadimitriou, D., & Bourse, D. (2012). A Systematic Approach for Closing the Research to Standardization Gap. In The Future Internet (Vol. 7281, pp. 18–29). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-30241-1_3

Scanlon, E., Sharples, M., Fenton-O'Creevy, M., Fleck, J., Cooban, C., Ferguson, R., … & Waterhouse, P. (2013). Beyond Prototypes: Enabling Innovation in Technology-Enhanced Learning. Technology-Enhanced Learning Research Programme, London.

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., ... & Baker, R. S. J. D. (2011). Open Learning Analytics: an integrated & modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques. Online at http://www.solaresearch.org/OpenLearningAnalytics.pdf. Accessed 2014-08-01

Siemens, G. (2013). Learning Analytics The Emergence of a Discipline. American Behavioral Scientist, 57(10), 1380–1400. doi:10.1177/0002764213498851

Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. American Behavioral Scientist, 57(10), 1510–1529. doi:10.1177/0002764213479366

Slate future tense (2014). What the Failure of inBloom Means for the Student-Data Industry, Online at http://www.slate.com/blogs/future_tense/2014/04/24/what_the_failure_of_inbloom_means_for_the_student_data_industry.html, accessed 2014-08-01

Making Sense of Online Learning Behavior: A Research on Learning Styles and Collaborative Learning Data

Meng SUN[a], Jiu-Tong LUO[a], Dong-Ming QIAN[b] & Xiao-Qing GU[a*]
[a] *Department of Educational Information Technology, East China Normal University, China*
[b] *e-Educational System Engineering Research Center & Shanghai Engineering Research Center of Digital Education Equipment, East China Normal University, China*
*xqgu@ses.ecnu.edu.cn

Abstract: This study focused on the relationship between learning styles, online behaviors and group collaborations. Sixty junior students from a university in China were taken as research object. Index of Learning Styles was used as a measuring tool to test participants' learning styles. Relationships between variables were measured by using bi-variate correlations analysis and one-way analysis of variance respectively. The results revealed a meaningful relationship between learning styles and online collaborative behavior. In addition, groups' online collaborative performances could be significantly different. However, grouping by learning styles might not be the factor that make effects on group collaborations.

Keywords: Learning styles, online learning, online behaviors, group collaborations

1. Introduction

The emergence of "big data" in education holds promise for improving learning processes in formal education, and beyond as well (Siemens & Baker, 2012). Currently, the learning management system (LMS) has been widely used, and has stored a lot of data, which notably supporting researchers' studies. The analysis of learners' online behaviors is a research focus in educational technology area. In the field of cooperative learning, what factors can affect groups' performances is also a topic worthy to explore.

It has been proven that learners' personality can affect learner behaviors in online learning environment. Many educators consider learning styles as an important factor to influence students' learning process. Furthermore, education research has emphasized that collaborative learning could improve project quality and performance (Soliman & Okba, 2006). The way in which students are grouped may affect the group performance. One of the features that can be taken into consideration when grouping is students' learning styles.

The aim of this study was to examine the relationship between learning styles, online behaviors and group collaborations. For this purpose, the following research questions will be answered:

- Which dimensions of learning style have effect on learners' online behaviors?
- Which kinds of online behaviors could be affected by learning style?
- Is there a significant difference among groups' online performances?
- Is there a significant relationship between group members' learning styles and groups' online collaborative performances?

2. Theoretical framework

LMS such as Blackboard, Moodle, Sakai et al., and many other learning systems have been widely used in e-learning around the world, providing teachers and students with a great variety of features which can be included in the course such as learning material, quizzes, discussion forums, assignments, and so on (Graf & Liu, 2008). With a lot of e-learning behavioral data produced, learning analytics (LA) has emerged as a new technology aimed to make sense of these data. Learning analytics is a technology focused on measurement, collection, analysis and reporting data about learners and contexts，for purposes of understanding and optimizing learning and the environments in which it occurs (Siemens et al., 2011; Siemens, 2012). Research on learning analytics has found that learning behavior data can predict students' learning to some extent.

Learning style is a unique combination of primary forms of processing information as well as the way in which various techniques and personal idiosyncrasies are used, and it has been though as one of the factors which can affect person's behavior. Research on learning style started from 20th century, and many teachers used this theory while they were teaching. There are various definitions and classifications of learning style.

Learning style is the way in which each person absorbs and retains information and/or skills, regardless of how that process is described, it is dramatically different from each person (Dunn, 1984). Pask (1988) defined learning style as a kind of strategy that learners like to use when they were processing a specific information. And Kolb (1999), Lotas (1977) and Oxofrd (1993) et al. also put forward their own classifications of learning style from difference aspects. Soloman & Felder (1997) developed an index of learning styles and was widely used today. In this instrument, learning styles were divided into four dimension, such as sensing or intuitive, visual or verbal, active or reflective and sequential or global.

3. Methods

3.1 The Participants and Context

72 (14male, 58 female) junior students majoring in educational technology from a university in China took part in the research. They were all in the course on Instructional System Design which lasted for 20 weeks began in March and ended in June. The course task for the students was to work in groups of six on a project assignment. Students could choose their partners freely. During the course, they should do theme discussion at the course platform (Sakai) and upload their homework to the platform. In addition, every group member should discuss about their group projects after class via the instant communication tool (QQ).

3.2 Measuring Tools

This study used the Chinese version of Index of Learning Styles (ILS) developed by Soloman & Felder (1997), and it had been tested in many researches in China. This questionnaire contains 44 items divided into four dimensions: active-reflective, sensing-intuitive, visual-verbal and sequential-global. Each dimension has associated with 11 forced-choice items with each option (a or b). Taking active or reflective as an example, for statistical analysis, it used 'b' responses minus 'a' responses, and then the ranging of number will between -11 to +11 (Felder et al., 2005). In this study, +11, +9, +7, +5, +3, +1 represented the learning style were active, and -11, -9, -7, -5, -3, -1 represented the learning style were active on the active-reflective dimension. The higher the value, the stronger is the preference. And the rest three dimensions were following the same distinguishing method mentioned above.

4. Data sources and analysis

In the study, results of students' learning styles inventory and online behavior, including clickstreams of students' participating in platform activities and groups' chat logs after class were all expected to be recorded and analyzed. However, one of the 72 student did not finish the questions of the ILS, and one group did not submit their group chat logs. Finally, data from 10 groups (60 students) were evaluated in total. The population was composed of 83.3% female (N = 50) and 16.7% male (N = 10).

Adopting a quantitative method, SPSS 19.0 was used to analyze the data. In order to answers the research questions, bi-variate correlations analysis and one-way analysis of variance (ANOVA) test were used; other basically and descriptively statistical analysis techniques such as frequency, percentage and standard deviation of the distribution were also employed.

5. Results

5.1 Students' learning styles

The average scores of the four dimensions, as shown in Table 1, revealed that the leaning style of all the 60 students were a little more active, sensing, and sequential, but much more visual.

Table 1: The results of students' learning styles

Learning Styles	N	Mean	SD	Min	Max
active-reflective	60	0.33	3.77	-9	9
sensing-intuitive	60	1.10	4.11	-9	7
visual-verbal	60	5.67	3.73	-5	11
sequential-global	60	0.17	4.27	-7	11

5.2 The relationship between learning styles and online behaviors

To investigate the relationship between learning styles and online behavior, correlation analysis was performed. One kind of the online behavior was frequencies of participating in platform activities (FA) which mainly consisted of uploading, downloading and posting. The other was average frequencies of speech (FS) in all the six times of after-class group discussion via QQ.

The Table 2 showed that there was significant positive correlation between active-reflective dimension and TS ($r = 0.27$, $p < 0.05$). But the other dimensions of learning styles showed no significant correlations.

Table 2: Correlations analysis results of relationship between learning styles and online behaviors

variables	1	2	3	4	5
1.active-reflective					
2.sensing-intuitive	-0.14				
3.visual-verbal	-0.02	0.00			
4.sequential-global	-0.20	0.33**	-0.41		
5.FA	0.43	0.15	-0.13	-0.04	
6.FS	0.27*	-0.01	-1.00	-0.07	0.45**

** Correlation is significant at the 0.01 level;
* Correlation is significant at the 0.05 level.

5.3 Differences among groups' online performances

The online behavior differences among groups were analyzed in one-way ANOVA (see Table 3). It showed that among different groups, students' collaborative behaviors were significantly different ($F=3.260$, $p<0.01$). Students in Group2, 8 and 9 talked much, but students in Group 1, 11 and 12 talked little. However, significant differences were not found among groups at the level of platform activities ($F=0.694$, $P>0.05$), which was relatively independent.

Table 3: ANOVA results of groups' online performances

		Sum of Squares	df	Mean Square	F	Post Hoc Tests LSD
FS	Between Groups	9.010	9	1.001	3.260**	G2>G1 G8>G1 G8>G11 G8>G12

	Within Groups	15.355	50	0.307		G9>G12
	Total	24.365	59			G9>G1
FA	Between Groups	3.563	9	0.396	0.694	.711
	Within Groups	28.521	50	0.570		
	Total	32.083	59			

** Mean difference is significant at the 0.01 level.

5.4 Difference of group members' learning styles

It was clear that active-reflective learning style could make a significant influence on one's speech times, which were significantly different among groups. Thus, it should be analyzed that if it was the group formations that resulted in the difference, and whether the groups with more speech times had more active members or their group members got much higher cores at active-reflective dimension. Under the condition that each group contained both active and reflective members (see Table4), one-way ANOVA was used to investigate if there was a significant difference of group formation at active-reflective dimension. It could be seen from Table 5 that there was no significant group formation difference (F=1.436, p>0.05). All groups were heterogeneous and group formation were alike, but interaction within groups were significantly different (see Table 3).

Table 4: Group formation on active-reflective learning style

		G1	G2	G3	G4	G5	G7	G8	G9	G11	G12
active	N	4	4	1	3	4	5	5	5	5	2
reflective	N	2	2	5	3	2	1	1	1	1	4
Total	N	6	6	6	6	6	6	6	6	6	6

Table 5: ANOVA results of group members' learning styles

		Sum of Squares	df	Mean Square	F
active-reflective	Between Groups	172.000	9	19.111	1.436
	Within Groups	665.333	50	13.307	
	Total	837.333	59		

6. Discussion

This study explored a new and important issue on the relationship between learning styles, online behaviors, and group collaborations. It implemented a quantitative research methodology to analyze:

- Which dimensions of learning style that has effect on learners' online behaviors.
- Which kinds of online behaviors that could be affected by learning styles.
- Whether there is a significant difference among groups' online performances.
- Whether there is a significant relationship between groups' formation and groups' online collaborative performances.

This study highlighted the emergent themes such as learning analytics (LA), group collaborative in online environment. And this innovative research also got some meaningful findings from the perspective of LA. The main conclusions were listed as following. Firstly, learning styles could influence individuals' online collaborative learning behaviors rather than the non-collaborative learning behaviors. Among the four dimensions of learning styles, active-reflective dimension played a key role. This study showed that active learners talked much more than reflective learners in a group, and the more active they were, the more frequently the talked. The findings were found to be theoretically consistent with the predictions of the Felder-Silverman learning style model (Felder & Silverman, 1988). Active learners prefer to process information actively by doing something with the learned material, for example discussing, explaining, or testing it. On the other hand, reflective learners prefer to think about the material and work alone. Regarding group discussing, active learners are expected to post more often in order to ask, discuss, and explain something, while reflective learners are supposed to prefer to participate passively by rarely expressing themselves. As a result, active learners are fit for group work, and could be advised to play the leading role in the group discussion to create positive and active environment.

Secondly, some kinds of online behavior such as uploading and downloading activities were not influenced by learning styles. Students with different learning styles showed no significant difference in terms of participating frequencies. Such behaviors possibly affected by other learner characteristics like gender or motivation types (Lim & Kim, 2003).

Thirdly, group formations of heterogeneous leaning styles might not the main factor that affect online group collaborative performances. All groups contained both active and reflective learner, but their averages speech times were significantly different. In other words, active learners in some groups spoke more than which in other groups. It is similar with the reflective learners. Some researches maintain that people in heterogeneous groups have the opportunity to learn from each other and will have better outcomes. Students believe that heterogeneity has potential benefits for their group performance (Herrmann, 1987; Kyprianidou et al., 2012). It was interesting in this study that heterogeneous groups could also perform differently. Other factors rather than learning styles might play a key role, like “group members’ reliability, good mood, mutual respect and empathy, clearly agreed goals, willingness to help, flexibility and adaptability” (Kyprianidou et al., 2012).

7. Conclusions and future work

In this paper, the relationship between learning styles, online behaviors, and group collaborations are addressed. Students with different learning style preferences showed significantly different online behaviors in some patterns. These results seem to be important when instructors design course contents and set course tasks. However, not all group collaboration performances could be affected by learning styles. For example, group discussion could be rather different while the group formation were alike in terms of learning styles. As a result, instructors should consider more factors when grouping the students.

Future work will deal with correlations of learning styles and other online behaviors such as time spent on examples, exercises, self-assessment tests, content objects and so on. The Sakai platform should be re-developed to do the analytic work and give the students more suggestion on their study. Moreover, qualitative analysis methods should be used for group discussion contents in order to explore what are the main factors that affect non-collaborative behaviors on the platform and group performance.

References

Dunn, R. (1984). Learning Style: State of the Science. *Theory into Practice, 23*(1), 10-19.

Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education, *Engineering Education, 78,* 674–681.

Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education, 21*(1), 103-112.

Graf, S., & Liu, T. C. (2008, July). Identifying Learning Styles in Learning Management Systems by Using Indications from Students' Behaviour. *In Advanced Learning Technologies, 2008. ICALT'08*. Eighth IEEE International Conference on (pp. 482-486). IEEE.

Herrmann, N. (1987). Creativity, learning, and the specialized brain in the context of education for gifted and talented children. Adapted from an address to the7th world conference on gifted and talented children,Salt Lake City, Utah

Kolb, D. A. (1999). Learning style inventory. McBer and Company.

Kyprianidou, M., Demetriadis, S., Tsiatsos, T., & Pombortsis, A. (2012). Group formation based on learning styles: can it improve students’ teamwork? *Educational Technology Research and Development*, *60*(1), 83-110.

Lim, D. H., & Kim, H. (2003). Motivation and learner characteristics affecting online learning and learning application. *Journal of Educational Technology Systems*, *31*(4), 423-439.

Lotas, A. (1977). Learning styles and curriculum. *The Affective Educator*, *11*(3),3-27.

Oxford, R. (1993). Style analysis survey. Language Learning Styles in the ESL/EFL Classroom. Boston: Heinle & Heinle.

Pask, G. (1988). Learning strategies, teaching strategies, and conceptual or learning style. In Learning strategies and learning styles, 83-100. Springer US.

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., Duval E., Verbert, K. & Baker, R. S. J. D. (2011). Open Learning Analytics: an integrated & modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques. http://www.solaresearch.org/OpenLearningAnalytics.pdf.

Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. *In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM, pp. 4-8.

Siemens, G., & d Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. *In Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). ACM.

Soloman, B. A., Felder, R. M. (1997). Index of learning styles. www2.ncsu. edu/unity/lockers/users/f/felder/public/ILSdir/ILS-a. htm.

Soliman, M., & Okba, E. (2006). Teamwork as a new sustainable pedagogy for teaching architectural design. Ain Shams University international conference, Cairo, pp. 181–192.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

How can Learning Analytics fit into a General Evaluation Framework and already be addressed during Learning Design?

Christian M. STRACKE[a*]
[a]*TELIT Research Institute, University of Duisburg-Essen (UDE), Germany*
[a]*East China Normal University (ECNU) in Shanghai, China*
*christian.stracke@uni-due.de

Abstract: In this paper, the author describes how learning analytics can be included within a general evaluation framework and be defined from the beginning of learning design for a learning opportunity. It is analyzed what part of a general evaluation framework will be covered by learning analytics and how learning analytics can contribute to the impact assessment using the generic Evaluation Framework for Impact Assessment (EFI) as an example. It will thereby be discussed how learning analytics can be addressed at the start of the design phase of the learning opportunity based on the reference process model from the international quality standard ISO/IEC 19796-1. Finally it is indicated how an extension of the learning design specification like PAS 1032-2 can be helpful for the introduction and support of learning analytics and summarized which further research is required.

Keywords: Learning Analytics, Evaluation Framework, Learning Design, Impact Assessment, Impact Measurement

1. Introduction

Learning analytics is becoming more and more a hot topic and important question for organizations and policy makers: How to monitor learning processes and to measure the learning outcomes and results and their impact. Learning analytics are starting to be broadly applied today and raise many open questions and issues concerning privacy and data protection. The different legal situations in all countries and the lack of global agreements are current barriers and the public discussions are only beginning now. All these important issues cannot be discussed here as this paper focuses on the potential integration of learning analytics into a general evaluation framework to address learning analytics already in the learning design.

Thus, in the second section of this paper it is described first how learning analytics can be included within a general evaluation framework: It is analyzed what part of a general evaluation framework will be covered by learning analytics and how learning analytics can contribute to the impact assessment using the generic Evaluation Framework for Impact Assessment (EFI) as an example. In the third section, it is thereby discussed how learning analytics can be included in and addressed at the start of the design phase of the learning opportunity based on the reference process model from the international quality standard ISO/IEC 19796-1. An extension of the learning design specification can be helpful for the introduction and support of learning analytics. Finally a summary is given in the conclusions with a foresight for future research.

2. Learning Analytics in a General Evaluation Framework

In this section of this paper it is discussed how learning analytics can be included within a general evaluation framework: It is analyzed what part of a general evaluation framework will be covered by learning analytics and how learning analytics can contribute to the impact assessment. The guiding question is what relation exists between learning analytics and a general evaluation approach following the philosophy of Total Quality Management (TQM). For that, the generic Evaluation Framework for Impact Assessment (EFI) will be used as an example that will be introduced in brief first.

2.1 The generic Evaluation Framework for Impact Assessment (EFI)

The Evaluation Framework for Impact Measurement EFI was developed to close a gap for assessing and optimizing the holistic total quality development within learning, education and training. It combines the traditional (internal) evaluation of the processes and developed products with the (external) evaluation concerning the strategic objectives and impact that is becoming more and more crucial due to economic cost pressures and international competition. Through this connection, the Evaluation Framework for Impact Measurement EFI offers an adaptable model for the definition and specification of indicators for both, the internal lifecycle and the external relations.

The Evaluation Framework for Impact Measurement EFI is combining the measurement of two dimensions:

1. (Internal) Impact of (direct) Results as outputs and
2. (External) Impact of Outcomes as indirect results.

Using the Evaluation Framework for Impact Measurement EFI, the following theoretical procedure has to be applied in general:

First, the impact of the internal development and output as direct results will be measured by operative indicators. Within one given project or process the operative indicators will be related to the planned products of the project or process. The measurement of the operative indicators has to focus the two dimensions of the pilot implementations: (1) the internal processes and activities and (2) the internal results (to be tested).

Second, the impact of the external relations and outcomes as indirect results will be measured by strategic indicators. They will be related to the strategic objectives of a given project or process: The measurement of the strategic indicators has to focus the two dimensions of the given project or process: (1) the external processes and activities (within the whole organization and external relations) and (2) the outcomes and their impact and external relations.

The following figure presents the overview of the Evaluation Framework for Impact Measurement EFI and demonstrates its relations between the two dimensions of impact measurement (internal impact of pilots assessed by the operative indicators and external impact of outcomes assesses by the strategic objectives):

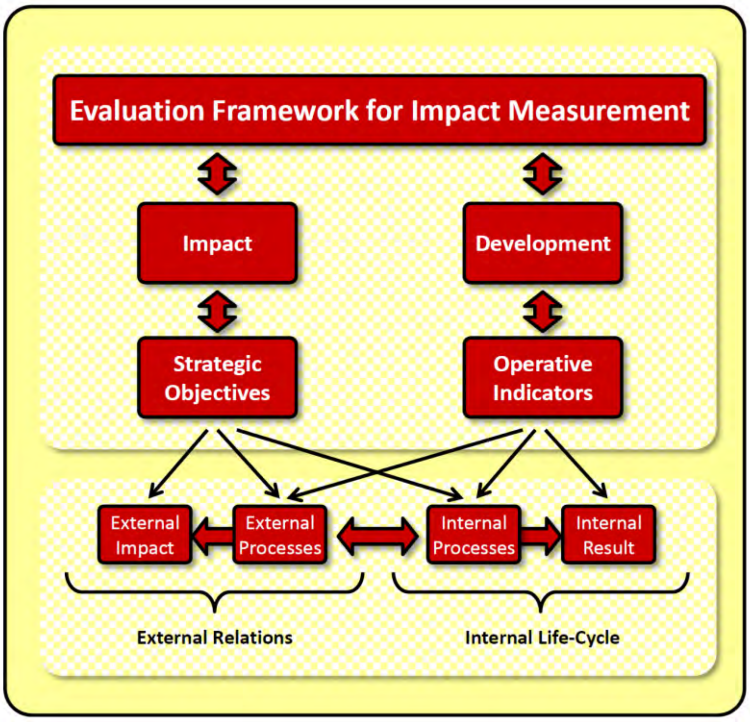


Figure 1. The Evaluation Framework for Impact Measurement EFI

2.2 Learning Analytics within the Evaluation Framework for Impact Assessment (EFI)

Learning Analytics is covering a broad range of different processes and results: It can start with monitoring the learning processes and providing feedback to the users as automatic recommendations for further progress as well as to the teachers, tutors or trainers as automatic indication which learners are progressing slowly or even failing and may require specific support and attention. Main objective of learning analytics is the measurement of the learning outcomes and results: Through their analysis it is expected that learning analytics are also contributing to the assessment of the impact. On the other hand, general evaluation frameworks have been developed during the last decades following the philosophy of Total Quality Management (TQM) and dealing with impact assessment, too. Thus, it is important to clarify their relationship and to combine and integrate them.

Therefore the guiding question is what relation exists between learning analytics and a general evaluation approach. General evaluation frameworks for a total quality management are addressing and covering all processes starting with the needs analysis whereas learning analytics is only starting with the learning process itself. Even though the concept and design of learning analytics should be discussed and defined from the beginning, the focus and scope of learning analytics is limited compared with holistic total quality management. Thus, learning analytics has to be part of a broader general evaluation framework like the Evaluation Framework for Impact Assessment (EFI).

Within EFI, learning analytics can directly support the impact assessment of the internal results, i. e. the learning outcomes of the learners achieved within the learning processes of the provided learning opportunity (e. g. by measuring the increase of knowledge, skills and competences in relation to the defined learning objectives). In addition learning analytics can also focus the assessment of the internal processes, i. e. the learning processes (e. g. by assessing the given answers and providing recommendations for further reading and in-depth learning).

For this purpose it is important that the operative indicators defined within EFI for the internal results and processes are also reflecting and aligned with the objectives of the learning analytics. On the other hand the operative indicators for the learning analytics has to be defined carefully that they can also contribute as input for the external impact assessment realized within EFI through the strategic objectives and related indicators. Then the indicators for the learning analytics can not only serve to monitor the learning processes (as internal products) and to measure the learning outcomes (as internal products) but also to support the assessment by EFI of the external impact on the organization, external stakeholders and the society.

3. Learning Analytics within Learning Design from the beginning

In this section, it is discussed how learning analytics can be addressed at the start of the design phase of the learning opportunity. The reference process model from the international quality standard ISO/IEC 19796-1 is an approach developed and approved in consensus as well as implemented worldwide. As it is covering the full life cycle of any learning opportunity, learning analytics is included even though it is not explicitly mentioned. In the following the role of learning analytics within this general reference process model should be identified: Therefore the international quality standard ISO/IEC 19796-1 will be introduced in brief first.

3.1 The international quality standard ISO/IEC 19796-1

The standard ISO/IEC 19796-1 is the first international quality standard for learning, education and training and provides a common reference framework for learning processes. It was developed in consensus by the Working Group 5 "Quality Assurance and Descriptive Frameworks" of the standardisation committee ISO/IEC JTC1 SC36 and issued by the International Organization for Standardization (ISO) in 2005. It contains the reference process model "Reference Framework for the Description of Quality Approaches" (RFDQ) to support stakeholders in learning, education, and training to document and (re-)define their daily business and processes. The reference process model of ISO/IEC 19796-1 is the integration of the following two main reference models (cf. ISO/IEC 2005):

- the generic process model and
- the generic descriptive model.

The reference process model covers the whole lifecycle of learning, education, and training in general including e-Learning and blended learning. Therefore it can be used to describe any learning scenarios as well as any educational and vocational training product and learning solution. It is important to note that the reference process model does not include any regulations about the sequence of the processes or interdependencies between them as well as it does not give any instructions on its specific implementation in detail as a prescription or regulation. The reference process model serves as an open descriptive framework that always needs the adaptation to the organisation, the learning context, and the given situation. The reference process model is based on the generic process model that is divided into seven process categories containing in total 38 processes. It is described by the following table:

Table 1: The process model of ISO/IEC 19796-1

ID	Category	Description	Processes
NA	Needs Analysis	Identification and description of requirements, demands, and constraints of an educational project	NA.1 Initiation NA.2 Stakeholder Identification NA.3 Definition of objectives NA.4 Demand analysis
FA	Framework Analysis	Identification of the framework and the context of an educational process	FA.1 Analysis of the external context FA.2 Analysis of staff resources FA.3 Analysis of target groups FA.4 Analysis of the institutional and organisational context FA.5 Time and budget planning FA.6 Environment analysis
CD	Conception / Design	Conception and Design of an educational process	CD.1 Learning objectives CD.2 Concept for contents CD.3 Didactical concept / methods CD.4 Roles and activities CD.5 Organisational concept CD.6 Technical concept CD.7 Concept for media and interaction design CD.8 Media concept CD.9 Communication concept CD.10 Concept for tests and evaluation CD.11 Concept for maintenance
DP	Development / Production	Realization of concepts	DP.1 Content realization DP.2 Design realization DP.3 Media realization DP.4 Technical realization DP.5 Maintenance
IM	Implementation	Description of the implementation of technological components	IM.1 Testing of learning resources IM.2 Adaptation of learning resources IM.3 Activation of learning resources IM.4 Organisation of use IM.5 Technical infrastructure
LP	Learning Process	Realization and use of the learning process	LP.1 Administration LP.2 Activities LP.3 Review of competency levels
EO	Evaluation/ Optimization	Description of the evaluation methods, principles, and procedures	EO.1 Planning EO.2 Realization EO.3 Analysis EO.4 Optimization/ Improvement

3.2 Learning Analytics within the Reference Process Model

The reference process model of ISO/IEC 19796-1 is a valuable and general instrument for the implementation and establishment of quality development in Learning, Education and Training (LET) and beneficial for the introduction of total quality management (cf. Stracke 2010). It has to be identified the role that learning analytics can play within it.

Several processes of the reference process model of ISO/IEC 19796-1 can be identified that are directly relevant for learning analytics and addressed by it: Learning analytics is measured during the process Activities (LP.2) of the process category Learning Process (LP) as part of the process Realization (EO.2) of the process category Evaluation/Optimization (EO). It has to be defined during the process planning (EO.1) of the same process category Evaluation/Optimization (EO) and is finally contributing to the process Analysis (EO.3). As already mentioned above, learning analytics should also be addressed from the beginning of the needs analysis and learning design: Thus, it would be necessary to include learning analytics into the definition of learning opportunities and their needs and design.

The identified processes from the reference process model of ISO/IEC 19796-1 can be used for the further refinement of indicators for learning analytics: As mentioned before, it is crucial that learning analytics are contributing to the general evaluation framework and that the indicators for assessing the impact of the learning opportunities have to be defined already during the learning design process in line with the overall evaluation as well as with the learning analytics.

An extension of the Learning Design (LD) specification developed by Rob Koper and his team at the Open University of the Netherlands in the year 2001 could be helpful for the introduction and support of learning analytics: The open Publicly Available Specification (PAS) DIN 1032-2 was developed by a working group of the German Standardization Body DIN based on the Learning Design: The main amendment is the introduction of the category context to define the environment and its conditions. This additional category is important for the learning analytics as it is providing the basic information for the definition of indicators measuring the learning outcomes.

Further research can reveal and transfer these conditions for the improvement of learning analytics within the evaluation planning and learning processes. And in particular learning analytics can support the enabling of new ways for the impact assessment of learning opportunities within a general evaluation framework that has to be discovered and discussed.

4. Conclusions

This paper presented how learning analytics can be included within a general evaluation framework and already be defined at the beginning of the learning design. For this specific purpose all open questions and issues concerning privacy and data protection that arise from a broad application of learning analytics were excluded: Such questions were outside of the scope of this paper, but must be addressed and are very important for a successful introduction of learning analytics. The generic Evaluation Framework for Impact Assessment (EFI) was introduced as an example of a general evaluation framework and could identify which part of it is covered by learning analytics as well as how learning analytics can contribute to the impact assessment. Based on the reference process model from the international quality standard ISO/IEC 19796-1, it was demonstrated that learning analytics can already be addressed during the start of the design phase for a learning opportunity. An extension of the Learning Design specification can contribute to the introduction and support of learning analytics. Future research should focus on the open question as to how the introduction of learning analytics can be harmonized in different systems and for different target groups and organizations by using standardized phases and processes such as the IDEAL reference framework. This would lead to comparable and hopefully interoperable learning analytics systems and data for the analysis and benchmarking across different systems, target groups and organizations.

References

Chermack, T. J.; Kasshanna, B. K. (2007). The Use and Misuse of SWOT Analysis and Implications for HRD Professionals. In: Human Resource Development International 10 (4), 383-399.

Chesbrough, H.W.; West, J.; Vanhaverbeke, W. (2006). Open Innovation: Researching a New Paradigm. Oxford: Oxford University Press.

Crosby, P. B. (1980). Quality is Free. The art of making quality certain. New York: McGraw-Hill.

Deming, W. E. (1986). Out of the Crisis. Cambridge, MA: MIT.

Deming, W. E. (1982). Quality, productivity and competitive position. Cambridge, MA: MIT.

Donabedian, A. (1980). The Definition of Quality and Approaches to Its Assessment [= Explorations in Quality Assessment and Monitoring, vol. 1]. Ann Arbor: Health Administration Press, 1980.

Drucker, P. F. (1999). Management Challenges of the 21st Century. Harper Business, New York.

Feigenbaum, A. (1986). Total Quality Control. Engineering and management. New York: McGraw-Hill.

Gassmann, O.; Enkel, E; Chesbrough, H.W. (2010). The future of open innovation. R&D Management, 40(3), 213–221.

Ishikawa, K. (1985). What is Total Quality Control? The Japanese Way. Englewood Cliffs, NJ: Prentice-Hall.

ISO 9000:2005 (2005). Quality Management Systems. Fundamentals and Vocabulary. Geneva: International Organisation of Standardization (ISO).

ISO 9001:2008 (2008). Quality Management Systems. Requirements. Geneva: International Organisation of Standardization (ISO).

ISO/IEC 19796-1:2005 (2005): Information Technology - Learning, Education, and Training — Quality Management, Assurance and Metrics — Part 1: General Approach. Geneva: International Organisation for Standardization (ISO).

Juran, J. M. (1992). Juran on quality by design. The new steps for planning quality into goods and services. New York: Free Press.

Juran, J. M. (Ed.) (1951). Quality Control Handbook. New York: McGraw-Hill.

Koper, Rob (2001): Modeling units of study from a pedagogical perspective - the pedagogical meta-model behind EML. http://eml.ou.nl/introduction/articles.htm , 2001-06, retrieved on 2004-05-17.

PAS 1032-2 (2004): Learning, Education and Training focussing on e-Learning – Part 2: Didactic Objects Model – Modelling and Description of Scenarios for Learning, Education and Training [in German].

Patton, M. Q. (2002). Quantitative evaluation and research methods. Newbury Park: Sage Publications.

Singh, K. (2007). Quantitative Social research methods. New Delhi: Sage Publication.

Soin, S. S. (1992). Total Quality Essentials. New York: McGraw-Hill.

Stracke, Christian M. (2014): "Evaluation Framework EFI for Measuring the Impact of Learning, Education and Training"; in: 华东师范大学学报（自然科学版）Journal of East China Normal University. Vol. 2014 (2). Shanghai: ECNU. p. 1-12. [DOI: 10.3969/j. ISSN 1000-5641. 2012.02.012]

Stracke, Christian M. (2013): "The Evaluation Framework for Impact Assessment"; in: Proceedings of 6th International Conference of Education, Research and Innovations 2013 [= ICERI 2013]. Madrid: IATED. p. 4654-4663 [ISBN: 978-84-616-3847-5].

Stracke, Christian M. (2010): "Quality development and standards in learning, education, and training: adaptation model and guidelines for implementations"; in: Информатизация образования и науки [= Информике (Informika), ISSN 2073-7572]; Vol. 7 (3), 2010. Moscow (Russian Federation): Informika. p. 136-146.

Thompson, J. L.; Martin, F. (2005). Strategic management: Awareness and change. London: Thomson learning.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Learning Analytics Data Items on Digital Textbooks

Yasuhisa Tamura[a*]
[a]*Sophia University, Japan*
*ytamura@sophia.ac.jp

Abstract: This paper proposes a set of data items to be collected in Digital Textbooks working on desktop/ laptop/ tablet PCs. Based on conventional LMS-based learning activity analytics, various types of data were proposed to use. In addition, modern tablet PC-based learning has advantage to collect more detailed learner's data with use of equipped sensors and material manipulation logging. This proposal is under discussion in IDPF EDUPUB community, which aims to specify ePub3-based Digital Textbook format and functions.

Keywords: Learning analytics, e-Textbooks, Learners' behavior, Analysis, Sensors, EDUPUB

1. Introduction

This article focuses on combination of two modern issues: Learning Analytics and Digital Textbooks. Some background information of these issues are introduced below.

Learning analytics (LA) has become a major area in learning science and learning technology research. From the end of 1990s, LMS (Learning Management System) based learning environments have emerged. Since then, many types of learning activities logs have been collected in these LMSs and analyzed. These data come from instruction-based activities, e.g. class participations, material views, and answers to quizzes. Also they include active learning-based ones, e.g. enrollments, utterances, interim and final products of group activities.

For LA researches, there are a series of International Conferences on Learning Analytics and Knowledge. These proceedings are available: Long et al. (2011), Dawson et al. (2012), Suthers et al. (2013) and Pistilli et al. (2014). As a general survey, Shum (2012) classifies 5 types of LA activities: (1) analysis dashboard of LMS or VLE, (2) predictive analysis, (3) adaptive learning analytics, (4) social network analysis, and (5) discourse analysis. Especially for active learning and collaborative learning, Shum and Ferguson (2012) shows some LA goal and future issues of these activities. Up to date discussion and information are available on Google Groups on Learning Analytics (2014).

As described below, the author intends to establish a basic and standard collection of data items to use LA activities with use of Digital Textbooks. This collection should include data items that are utilized in previous researches. In order to clarify these data items, the author investigated some of previous published papers and listed up the used data items. The summary is shown in Table 1 and Table 2. Table 1 shows 17 papers to focus on classroom and individual learning. Also, Table 2 shows 13 papers to focus on collaborative and active learning. These data items are referred in the proposal in Section 2.

Digital textbooks, also known as e-textbooks, are now investigated and planned to implement at several countries all over the world. KERIS (2014) in Korea started investigation and experiment in 2008, and lead to finish implementation throughout the country until the end of 2015. Also China, Singapore, Philippines, India and other Asian countries are proceeding investigation and experimental introduction. In Europe, England, France, Germany, Spain and other countries are under investigation and experiment. In United States of America, some states including California, Washington and Utah are planning to deliver open textbooks or complementary devices.

Table 1. Data items and Objectives of Learning Analytics Researches (classroom and individual learning).

Reference	Data Items	Goal of Analysis
Arnold and Pistilli (2012)	Posting of a traffic signal indicator on a student's LMS home page, E-mail messages or reminders, Text messages, Referral to academic advisor or academic resource center, Face to face meetings with the instructor	Relationship between items and achievement
Barber and Sharkey (2012)	Prior credits earned, Discussion post count/week, Late assignments, Orientation participation, Count of messages to instructor, Inactive time since last course	Prediction of class achievement
Clow (2013)	Visit, Registration, and contribution ratio of MOOCs	Drop rate analysis of MOOCs learners
Graf et al. (2011)	Templates, patterns, learning object, database connections of materials	Judgment of material difficulty
Holman et al. (2013)	Grade, Class standing, and badges of quizzes	Self prediction of achievement
Kizilcec et al. (2013)	Visiting, Enrollment, and assessment numbers in MOOCs courses	Number transition of MOOCs learners
Lonn et al. (2012)	Grade information every few weeks	Assistance necessity from mentors
Martin et al. (2013)	Answers of each sub-quiz	Visualization of learning process
Monroy et al. (2013)	Teacher's usage of teaching unit parts (overview, essentials, engage, explore, explain, evaluate, intervention, acceleration)	Heat map of unit parts usage
Niemann et al. (2012)	Learning object usage in a web portal	Similarity of learning objects
Pardos et al. (2013)	Quizzes and scaffolding help	Relationship between Scaffolding help and achievement
Raca and Dillenbourg (2013)	Video captured actions of learners	Learner behavior during classrooms
Santos et al. (2012)	Date and time range of learners	Visualization of learning status
Sao Pedro et al. (2012)	Quiz answers	Transition of problem solving skills
Tempelaar et al. (2013)	Achievements in various learning areas	Skill analysis (Self-belief, learning focus, planning, management, persistence)
Verbert and Duval (2011)	Dataset and functions of recommender system	Comparison of Recommender systems
Wolff and Zdrahal (2013)	Precision and recall of learning units	Comparison of TMA (Tutor-marked assessment) and VLE (Virtual learning environment)

Table 2. Data items and Objectives of Learning Analytics Researches
(Collaborative and active learning).

Reference	Data Items	Goal of Analysis
Ahn (2013)	Emails received, Emails sent, Friends, Friend Lists, Links, Member pages, Networks, Notes, Photos, Status messages, Videos, Wall posts	Factor analysis of media literacy (Negotiation, Networking, Judgment, Play, Multitasking, Appropriation, Transmedia navigation)
Cambridge and Perez-Lopez (2012)	Discussion post, blog, their narratives,	Analysis of discourse style to activate learner groups
Camilleri et al. (2013)	Pleases and numbers of utterances in virtual space	Behavior analysis
Cobo et al. (2012)	Reading and writing activities during online discussions	Clustering of learners
Ferguson and Shum (2011)	Keywords in text chat	Chat type (evaluation, explanation, reasoning, justification, perspective)
Koulocheri and Xenos (2013)	Bookmarks, blog posts, topics and files uploaded, bookmarks, comments on bookmarks/blog posts/topics/files in group	Visualization of member relationships
De Liddo et al. (2011)	Response type of utterances (respond, about, example, solution, support)	Relationship analysis of learners
Schneider et al. (2013)	Eye-tracking data	Estimation of collaborative learning skills
Schreurs et al. (2013)	Person, type of tie, topic	Visualization of learner relationship network
Shum and Crick (2012)	Quiz achievement and various activities	Relationship between individual learning achievement and meta-skills
Siadaty et al. (2012)	Vocabulary in shared Wiki and bookmark	Collaborative skills analysis of corporate learners
Suthers and Rosen (2011)	Chat, Discussion, File sharing	Multiple level visualization (Process, Domain, Event, Action, Mediation, Relationship, Tie)
Tempelaar et al. (2013)	Achievements in various learning areas	Analysis of necessary skills (Self-belief, learning focus, planning, management, persistence)

In Japan, MEXT (Ministry of Education, Culture, Sports, Science and Technology) (2011) published a roadmap called “The Vision for ICT in Education”, which were planning to introduce digital textbooks countrywide until 2020. Also, an experimental project was deployed from 2011 to 2013. It was a joint project between MEXT and MIC (Ministry of Internal Affairs and Communication) to introduce ICT and digital learning materials to selected 20 schools. Final report of this project (in Japanese) is available through MEXT (2014). At the same time, MEXT and MIC started experimental development projects of Digital Textbooks in 2013. In these projects, MEXT is focusing ePub3, while MIC is HTML5. These projects will continue in 2014.

On the other hand, various standardization organizations and communities are trying to specify standard file formats and specifications for Digital Textbooks. These projects and their timelines are shown in Figure 1. IEEE (2014) initiated Actionable Data Book Project in 2011, and published some research papers. Also, CEN (European Committee for Standardization) (2014) and IMS Global Learning Consortium (2014) began eTernity Project and ICE Project in 2012 and 2013, respectively.

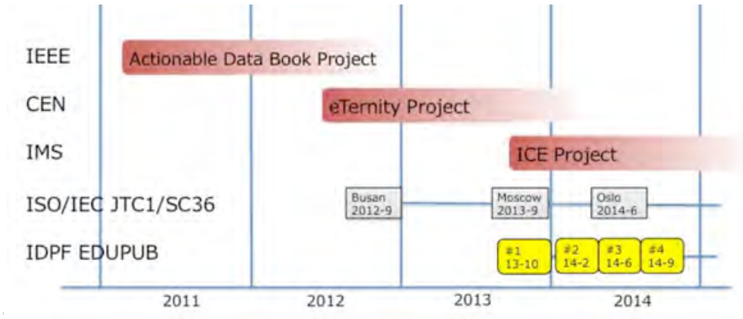


Figure 1. Standardization projects for Digital Textbooks

Among them, ISO/IEC JTC1/SC36 (2014), a subcommittee of ISO dedicated to e-learning technical specifications, started e-Textbook Project in September 2012 meeting at Busan, Korea. It is investigating related standardization activities, issued a set of questionnaires of Digital Textbooks to standardization communities, and arranged future issues in a document in 2014 meeting.

The latest and the most active one is called EDUPUB project. It is lead by IDPF (International Digital Publishing Forum) (2014), which specified ePub3 format for Digital Books. The first workshop of EDUPUB was held in October 2013 at Boston, USA, while the second in February 2014 at Salt Lake City, USA. The third workshop was held in June 2014 at Oslo, Norway, and the fourth is scheduled in September 2014 at Tokyo, Japan. Through these workshops, these outlines below are discussed.

- Core file format is ePub3.
- In order to add textbooks specific structural semantics, Pearson and Benesse staffs proposed their textbook and material descriptive tags. It is under online discussion.
- In order to attach learner note into the textbooks, “Open Annotation in ePub” specification is under discussion.
- For quiz data format, IMS QTI (2014) (Question and Test Interoperability) specification is a major candidate.
- For calling scheme of outer applications or resources, IMS LTI (2014) (Learning Tools Interoperability) specification is a major candidate.
- Textbook specific metadata items are under discussion.

The author is a member of MEXT Digital Textbook project, ISO/IEC JTC1/SC36 e-Textbook project, and IDPF EDUPUB Project. In the EDUPUB project, there was a proposal to specify a set of data items to be collected with use of LA. For this proposal, the author started to survey conventional research papers in order to specify commonly used data items, and also proposed a new set of items which are able to collect with use of tablet PCs. Section 2 shows this proposal.

2. Data Items Acquired with use of Digital Textbooks

2.1 Characteristics of Tablet PCs

There are many types of PCs utilized in classroom and individual learning. Both desktop PCs and laptop PCs have been common. In addition, tablet PCs have become popular in these years. Apple launched a first iPad in 2010. Also, Google and China/Korean hardware companies began to launch Android based tablet PCs in 2010. Nowadays, worldwide shipments of desktop / laptop PCs and tablet PCs are almost equal in 2014. In 2015, shipments of tablet PCs will be 20% more than desktop / laptop (Gartner 2014).

Even traditional laptop PCs are able to connect to computer network, download information from a certain server, upload it to a server, or communicate each other with use of e-mail and SNS. Also they have some sensors: brightness sensor, camera, and microphone. With use of these functions, they are able to generate data to be used in LA:

- Enroll to a class in LMS
- Access materials in LMS
- Upload quiz answer / assignment / reaction
- Show hint / advise
- Send, receive and read messages from / to instructors
- Enroll to a group in LMS
- Send / receive text / audio / video messages from / to instructors / another learner
- Access to shared whiteboard / file
- Timestamp of these activities

In conventional way, all of these information are collected in LMS. With use of additional functions attached with learners' Web browsers, some information can be collected in client (=learner) PCs, but this approach is not common.

However, modern tablet PCs equip many other types of sensors: screen touch sensor, GPS, digital compass, gyroscope, acceleration sensor, etc. With use of these sensors, a tablet PC is able to collect various information about learning activities and their environment. For example:

- View / flip one's textbook, reference or dictionary
- Insert highlights or underlines in one's textbook, reference or dictionary
- Write notes or annotations on one's textbook, reference or dictionary
- Refer reference or dictionary by specifying a certain part of textbook
- In addition to timestamps, places of these activities
- Environmental voice and noise of these activities
- Learner's face, expression, and visual environment of these activities

Some of these data are collected with use of equipped sensors directly, others should be analyzed by Digital Textbook viewer software or related application software. Also, the data collection implies privacy violation. This issue will be discussed in the later section.

2.2 Proposed Data Items

Based on investigation and consideration stated above, the author proposes data items below to be collected with use of Digital Textbooks. Figure 1 shows a framework. One atomic data includes "who" (Subject), "when" (Date & Time), "where" (Geographic point location, optional), and "what". A detail of lower right side table of Figure 2 is shown in Table 3.

Data items in Table 3 consist of two categories: (1) commonly used in conventional LA researches, shown in Table 1 and Table 2, and (2) assumed to be collected on Tablet PCs, Digital Textbook viewer and related software mentioned in Section 2.1.

This proposal is now disclosed to EDUPUB community, and weekly discussion is ongoing.

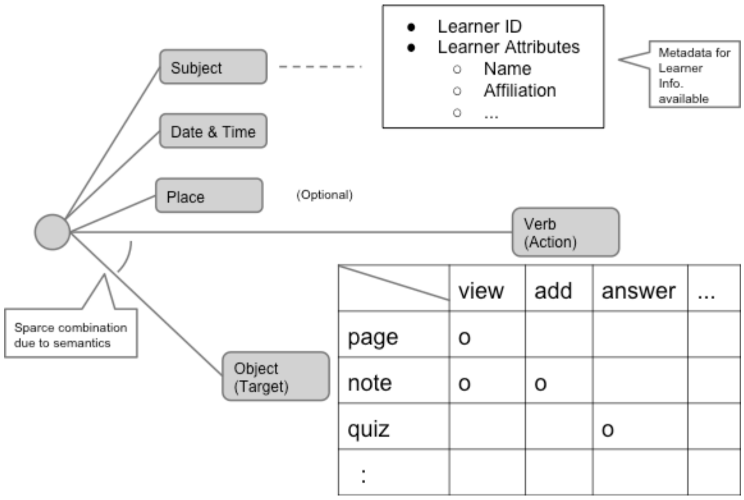


Figure 2. Framework of proposed data items.

Table 3. Detail of Proposed Data Items: Verbs and Objects.

Category	Verb (action)/ Object (target)	attend/quit	flip/view	add	modify	delete	answer	send	receive
Classroom / individual Learning	Class	o							
	Page of e-textbook or reference		o						
	Highlight / underline			o	o	o			
	Note (annotation)		o	o	o	o			
	Link		o	o	o	o			
	Quiz		o				o		
	Assignment		o				o	~~o~~	~~o~~
	Feedback		o				o		
	Message							o	o
Collaborative and active learning	Group	o							
	Shared whiteboard		o	o	o	o			
	Shared file		o	o	o			o	o

3. Discussion and Conclusion

During development process of the draft proposal above, there were some discussions whether tablet PC based fine-grained data should be included, for example:

- Face expression of a learner,
- Attitude of a learner,
- Voice of a learner and environmental sound,
- Acceleration data,
- Digital compass data, and
- Gyroscope data.

Also, it is emerging to utilize so called “wearable devices” to collect biological and environmental data, for example:

- Temperature of learner’s body and environment,
- Humidity of environment,
- Body sweat of a learner,
- Heart race of a learner,
- Blood pressure of a learner,
- Eye-tracking data of a learner, and
- Brain waves of a learner.

Currently, it is not clear that these data are useful to identify learner’s status or not, at least from preceding research results. Therefore, the author thinks that the data listed above are still early to include as “standard” data items for learning analytics. It’s why the listed data is omitted from Table 3. However, as far as the author knows, the proposal appeared in Figure 2 and Table 3 is the first appearance for “standard” data items for learning analytics activities. This “standard” means that major stakeholders support to adapt them as useful ones. There are many goals in LA, shown as samples in Table 1 and Table 2. This proposed data items cover these, and similar LA goals.

A major discussion point for the proposed data items is a risk of privacy violation of learners. This proposal includes geographical data and timestamp. So, an analyst or an instructor is able to grasp when and where a learner is. Also, a tablet PC is able to collect visual and audio data during learning activities. It might clarify a scene and accompanying friends during learning. Currently it is not clear what data violates learner’s privacy and doesn’t. We should clarify a threshold of private data, and make broad consensus. From this viewpoint, Table 3 does not include visual and audio data not intentionally recorded by learners.

One of the future issues is comprehensiveness of the proposal. Currently there is no major argument for the proposal. However, there are many other existing researches of LA. They should be investigated in order to guarantee comprehensive of this proposal. Also, IMS proposes Caliper specification. It also specifies a data set for LA. The document is not opened, but needs investigation.

The other one is to verify usefulness of these data items for LA activities. The proposed data items are listed from viewpoint of technical feasibility with use of tablet PCs and Digital Textbook software. However, it is not clear what characteristics can be analyzed with use of these items. Some them are already clear based on the conventional researches in Table 1 and Table 2. However, especially Digital Textbook specific items should be verified in this usefulness.

To conclude, the author proposes a set of data items to be collected with use of tablet PCs and Digital Textbook viewer software, shown in Figure 1 and Table 3. They are far more detailed than conventional LMS based data collection. However, it should be enhanced and brushed up in order to guarantee comprehensiveness and learner’s privacy.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 26282059. The author would like to appreciate Markus Gylling and all EDUPUB project members for dedicated discussion and continuous encouragement for this work.

References

Ahn, J. (2013). What Can We Learn from Facebook Activity? Using Social Learning Analytics to Observe New Media Literacy Skills, *Proc. 3rd Int’l Conf. Learning Analytics and Knowledge*, pp.135-144.

Arnold, K.E. and Pistilli, M.D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success, *Proc. 2nd Int’l Conf. Learning Analytics and Knowledge*, pp.267-270.

Barber, R. and Sharkey, M. (2012). Course Correction: Using Analytics to Predict Course Success, *Proc. 2nd Int’l Conf. Learning Analytics and Knowledge*, pp.259-262.

Cambridge, D. and Perez-Lopez, K. (2012). First Steps Towards a Social Learning Analytics for Online Communities of Practice for Educators, *Proc. 2nd Int’l Conf. Learning Analytics and Knowledge*, pp.69-72.

Camilleri, V., de Freitas, S. et al. (2013). A case study inside Virtual Worlds: use of analytics for immersive spaces, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.230-234.

Clow, D. (2013). MOOCs and the Funnel of Participation, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.185-189.

Cobo, G., García-Solórzano, D. et al. (2012). Using agglomerative hierarchical clustering to model learner participation profiles in online discussion forums, *Proc. 2nd Int'l Conf. Learning Analytics and Knowledge*, pp.248-251.

CEN (2014). eTernity Project, retrieved on May 20, 2014 from http://www.eternityproject.eu/

Dawson, S. et al. (eds) (2012). Proc. 2nd International Conference on Learning Analytics and Knowledge 2012. ACM.

EDUPUB1 (2013). IDPF First EDUPUB Workshop, retrieved on May 20, 2014 from http://idpf.org/edupub-2013

EDUPUB2 (2014). IMS Global Learning Consortium, Second EDUPUB Workshop, retrieved on May 20, 2014 from http://www.imsglobal.org/edupub/

EDUPUB3 (2014). IDPF, EDUPUB Europe 2014, retrieved on May 20, 2014 from http://idpf.org/edupub-europe-2014

Ferguson, R. and Shum, B. (2011). Learning Analytics To Identify Exploratory Dialogue within Synchronous Text Chat, *Proc. 1st Int'l Conf. Learning Analytics and Knowledge*, pp.99-103.

Gartner (2014). Gartner Says Worldwide Traditional PC, Tablet, Ultramobile and Mobile Phone Shipments On Pace to Grow 7.6 Percent in 2014, retrieved on May 20, 2014 from http://www.gartner.com/newsroom/id/2645115

Google groups on learning analytics (2014). retrieved on May 20, 2014, from https://groups.google.com/forum/#!forum/learninganalytics

Graf, S., Ives, C. et al. (2011). AAT – A Tool for Accessing and Analysing Students' Behavior Data in Learning Systems, *Proc. 1st Int'l Conf. Learning Analytics and Knowledge*, pp.174-179.

Holman, C., Aguilar, S. et al. (2013). GradeCraft: What Can We Learn From a Game-Inspired Learning Management System?, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.260-264.

IDPF (2014). retrieved on May 20, 2014 from http://idpf.org/

IEEE (2014). Actionable Data Book Project, retrieved on May 20, 2014 from http://standards.ieee.org/about/sasb/iccom/IC12-006-02_IEEE_Actionable_Data_ Book_for_STEM_Education.pdf

IMS Global Learning Consortium (2014). ICE Project, retrieved on May 20, 2014 from http://www.imsglobal.org/icecallforParticipation.html

IMS LTI (2014). IMS Global Learning Consortium, Learning Tools Interoperability (LTI), retrieved on May 20, 2014 from http://www.imsglobal.org/toolsinteroperability2.cfm

IMS QTI (2014). IMS Global Learning Consortium, Question and Test Interoperability (QTI), retrieved on May 20, 2014 from http://www.imsglobal.org/question/

ISO/IEC JTC1/SC36 (2014). retrieved on May 20, 2014 from http://www.sc36.org:8080/

KERIS (2014). retrieved on May 20, 2014 from http://english.keris.or.kr/

Kizilcec, R. F., Piech, C. et al. (2013). Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.170-179.

Koulocheri, E., and Xenos, M. (2013). Considering Formal Assessment in Learning Analytics within a PLE: The HOU2LEARN Case, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.28-32.

Long, P. et al. (eds) (2011). Proc. 1st International Conference on Learning Analytics and Knowledge 2011. ACM.

Lonn, S., Krumm, A. E. et al. (2012). Bridging the Gap from Knowledge to Action: Putting Analytics in the Hands of Academic Advisors, *Proc. 2nd Int'l Conf. Learning Analytics and Knowledge*, pp.184-187.

Martin, T., Aghababyan, A. et al. (2013). Nanogenetic Learning Analytics: Illuminating Student Learning Pathways in an Online Fraction Game, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.165-169.

MEXT (2011). The Vision for ICT in Education, retrieved on May 20, 2014 from http://www.mext.go.jp/b_menu/houdou/23/04/__icsFiles/afieldfile/2012/08/03/1305484_14_1.pdf

MEXT (2014). Final Report on Experimental Study on Learning Innovation Project (in Japanese), retrieved on May 20, 2014 from http://www.mext.go.jp/b_menu/shingi/chousa/shougai/030/toushin/1346504.htm

Monroy, C., Rangel, V. S.et al. (2013). STEMscopes: Contextualizing Learning Analytics in a K-12 Science Curriculum, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp. 210-219.

Niemann, K., Schmitz, H. C. et al. (2012). Clustering by Usage: Higher Order Co-occurrences of Learning Objects, *Proc. 2nd Int'l Conf. Learning Analytics and Knowledge*, pp.238-247.

De Liddo, A., Shum, B. et al. (2011). Discourse-Centric Learning Analytics, Proc. *1st Int'l Conf. Learning Analytics and Knowledge*, pp.23-33.

Pardos, Z. A., Baker, R. S. et al. (2013). Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes, Proc. *3rd Int'l Conf. Learning Analytics and Knowledge*, pp.117-124.

Pistilli, M. et al. (eds) (2014). Proc. 4th International Conference on Learning Analytics and Knowledge. ACM.

Raca, M., and Dillenbourg, P. (2013). System for Assessing Classroom Attention, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.265-269.

Santos, J. L., Govaerts, S. et al. (2012). Goal-oriented visualizations of activity tracking: a case study with engineering students, *Proc. 2nd Int'l Conf. Learning Analytics and Knowledge*, pp.143-152.

Sao Pedro, M. A., Baker, R. S. et al. (2013). What Different Kinds of Stratification Can Reveal about the Generalizability of Data-Mined Skill Assessment Models, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.190-194.

Schneider, B., Abu-El-Haija, S. et al. (2013). Toward Collaboration Sensing: Applying Network Analysis Techniques to Collaborative Eye-tracking Data, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.107-111.

Schreurs, B., Teplovs, C. e al. (2013). Visualizing Social Learning Ties by Type and Topic: Rationale and Concept Demonstrator, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.33-37.

Shum, S. B. (2012). Learning Analytics: Policy Brief, UNESCO Institute for Information Technologies in Education, retrieved on May 20, 2014 from http://iite.unesco.org/pics/publications/en/files/3214711.pdf

Shum, S. B. and Ferguson, R. (2012). Social Learning Analytics, *Educational Technology & Society*, Vol.15, No,3, pp.3-26.

Shum, S.B. and Crick, R.D. (2012). Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics, *Proc. 2nd Int'l Conf. Learning Analytics and Knowledge*, pp.92-101.

Siadaty, M., Gašević, D. et al. (2012). Learn-B: A Social Analytics-enabled Tool for Self-regulated Workplace Learning, *Proc. 2nd Int'l Conf. Learning Analytics and Knowledge*, pp.115-119.

Suthers, D. and Rosen, D. (2011). A Unified Framework for Multi-Level Analysis of Distributed Learning, *Proc. 1st Int'l Conf. Learning Analytics and Knowledge*, pp.64-74.

Suthers, D. et al. (eds) (2013). Proc. 3rd International Conference on Learning Analytics and Knowledge 2013. ACM.

Tempelaar, D. T., Heck, A. et al. (2013). Formative Assessment and Learning Analytics, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.205-209.

Verbert, K. and Duval, E. (2011). Dataset-driven Research for Improving Recommender Systems for Learning, *Proc. 1st Int'l Conf. Learning Analytics and Knowledge*, pp.44-53.

Wolff, A. and Zdrahal, Z. (2013). Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment, *Proc. 3rd Int'l Conf. Learning Analytics and Knowledge*, pp.145-149.

Tseng, S.-F. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Learning Analytics: An Enabler for Dropout Prediction

Shu-Fen TSENG[ad]*, Chih-Yueh CHOU[bd], Zhi-Hong CHEN[cd], Po-Yao CHAO[cd]
[a] *Department of Information Management, Yuan Ze University, Taiwan*
[b]*Department of Information Science & Engineering, Yuan Ze University, Taiwan*
[c]*Department of Information Communication, Yuan Ze University, Taiwan*
[d]*Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taiwan*
*gssftseng@saturn.yzu.edu.tw

Abstract: A key application of learning analytics is predicting students' learning performances and risks of dropping out. Heterogeneous data were collected from selected school to yield a model for predicting student's dropout. Results from this exploratory study conclude dropout prediction by learning analytics may provide more precise information on identifying at-risk students and factors causing them to be at risk.

Keywords: Learning analytics, dropout, predictive model

1. Introduction

It is gradually recognized the effective use of data and learning analytics are both critical components of digital learning strategies to personalize learning for needed students to increase student retention and achievement in schools (Bienkowski, et al., 2012). Analytics applied to education data can help schools and school systems to better understand how students learn and succeed. Significant improvements in technology tools and resources, and the focus on meeting the needs of individual students through personalized and digital learning have together provided an emerging context in which education systems have the opportunity to advance the values of learning analytics to truly inform teaching and learning.

A key application of learning analytics is monitoring and predicting students' learning performance and addressing potential issues early so that interventions can be provided to identify students who are at risk of failing a course of program of study (Bienkowski, et al., 2012). In particular, one of the key successes of learning analytics is predicting which students are at risk of dropping out. Tobin and Sugai (1999) demonstrated that using data on disciplinary referrals during middle school can predict those who are likely to drop out of high school. Their research supports the use of school records of discipline referrals as a screening device. The work by Bowers (2010) shows that by analyzing student grade history from k-8 or k-12, it can correctly identify over 80% of dropout students.

One of noteworthy cases is that IBM worked with Mobile County Public Schools in Alabama to apply analytics to education data to help school system identify which students were at risk of dropping out. Rapid access to information and analytics means school administrators and principals can make more informed decisions and take appropriate and timely action to develop an individualized response to each student's problems and monitor their progress. Timely intervention based on real-time information is helping Mobile County to keep students on the right path and to lower dropout rates (Centre for Information Policy Leadership, 2013; IBM, 2009).

It has long been recognized by school authorities in Taiwan that dropout students mostly come from disadvantage family (e.g. single parent, minority), who lack of self-discipline, are not interested in school work, and easily befriend juvenile delinquents. There were a total of 5,379 dropouts among k1-k12 students in the year of 2012 in Taiwan. Personal profiling outlined mostly of them dropped out during their middle school periods (89%). Among these dropouts, 58% of them came from single parent household. Major causes of their dropout were categorized into individual reasons (47%), followed by family (24%) and social (17%) problems (MOE, 2012).

School administrators and educators can only analyze these students when they had already dropped out of school. In contrast of taking remedial actions, the newly development of learning analytics could help predict in advance and provide the tools to calculate a student's risk of dropping out of school. In this exploratory study, we collect students' data from heterogeneous sources (ie. grade history, disciplinary referral, class attendance) and employ logistic regression model to identify risk factors of dropping out in one selected school. Then, we compare the findings to profiles of dropout that outlined by conventional method. We further discuss the potential benefits of learning analytics on dropout prediction.

2. Method

Taoyuan county ranks the second highest dropout rate in Taiwan. In this study, we collected data from one middle school in Taoyuan county. The school selected was among one of the highest dropout rate schools. Heterogeneous sources of administrative, academic, and disciplinary data of every student in this school were collected and organized into a database. Descriptive and explanatory statistics were used to determine the relationship between a dependent variable (whether a student was dropped out in current term) and independent variables (e.g. students' family background, academic grade, class absence, and disciplinary referral records in previous term). In particular, logistic regression analysis was employed to create a model for predicting whether a student was at risk of leaving school.

3. Findings

In the year of 2013, there were a total of 712 seventh and eighth grade students in this middle school. Among them, 11 students dropped out of school. Five of them were males and four out of 11 students were minorities. Correlation coefficients showed that dropout was positively related to number of class absence ($r=.363$), number of disciplinary referral ($r=.253$) in previous term, and it was negatively related to academic grade ($r=-.06$) of previous term. Despite the limited number of dropout cases, logistic regression model (table 1) still showed prediction power of disciplinary referral (odd ratio=1.136) and academic grade (odd ratio=.932) of previous term on a students' risk of dropping out of school. Number of class absence in previous term and gender showed insignificant impacts on predicting dropout.

Table 1: Logistic regression of dropping out of school by risk factors.

	Dropout		
	B	Exp(B)	Effect on dropout risk
Constant	.332	1.394	
Male	-.409	.664	
Class Absence (previous term)	.013	1.013	
Disciplinary referral (previous term)	.127	1.136*	Increase risk by 13.6%
Academic grade (previous term)	-.07	.932**	Decrease risk by 6.8%
Cox & Snell R2	.045		
N	712		

4. Discussion

The conventional way of outlining high risk student of dropouts is to analyze demographic and family characteristics of those who had already dropped out. In the year of 2013, among those dropout students in Taoyuan county, 47% of dropouts were from single parent household, 24% of them were minorities, and 6% of them were raised by grandparents. Based on this conventional method, individual and family reasons were ranked the highest causes of student to drop out of school.

By employing method of learning analytic, we collected and organized data from personal and family composition, academic performance and disciplinary records; our finding suggests that we can predict a student's risk of dropping out of school mainly based on his/her school related activity records of previous term. In particular, high number of disciplinary referral and poor academic grade were major factors that help to predict student's risk of dropping out of school. The results suggested instead of targeting individual reasons and family background, student's activities in school can be effectively used to predict a student's risk of dropping out of school. With more longitudinal data of students are collected and more middle schools join the project, we could build a better fit model for dropout prediction and calculate risk of dropout for students. This exploratory study concludes that dropout prediction from learning analytics may provide more precise information on identifying at-risk students and factors causing them to be at risk.

References

Alliance for Excellent Education (2014). Capacity enablers and barriers for learning analytics: Implications for policy and practice. Retrieved from http://all4ed.org/reports-factsheets/capacity-enablers-and-barriers-for-learning-analytics-implications-for-policy-and-practice/

Baker, R.J.d. (2013). Learning, schooling, and data analytics. In M. Murphy, S. Redding, & J. Twyman (eds.), Handbook on Innovations in Learning (pp.179-109). PA: Temple University.

Bienkowski, M., Feng, M., and Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Retrieved from http://www.ed.gov/edblogs/technology/files/2012/03/edm-la-brief.pdf

Bowers, A.J. (2010). Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. Practical Assessment, Research & Evaluation, 15(7), 1-18.

Center for Information Policy Leadership (2013). Big data and analytics: Seeking foundations for effective privacy guidance. Retrieved from http://www.hunton.com/files/Uploads/Documents/News_files/Big_Data_and_Analytics_February_2013.pdf

IBM (2009). Mobile county schools applied analytics to transform student performance. Retrieved from http://www.ibm.com/smarterplanet/us/en/leadership/mobilecounty/

Ministry of Education (2012). Dropout statistic in Taiwan. Executive Yuan, Taiwan.

Romero, C. and Ventura, S. (2010).Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, 40(6),601-618.

Siemens, G. and Baker, R.J.d. (2012). Learning analytics and educational data mining: Towards communication and collaboration. Retrieved from http://users.wpi.edu/~rsbaker/LAKs%20reformatting%20v2.pdf

Tobin, T.J. and Sugai, G. M. (1999). Using sixth-grade school records to predict violence, chronic discipline problems, and high school outcomes. Journal of Emotional and Behavioral Disorders, 7(1), 40-53.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Motivation and Engagement in MOOC – Teachers' Perspective

Li Fern TAN[a*], Kai Song GOH[b] & Emile SABASTIAN[b]
[a]*Office of Learning Technologies, Temasek Polytechnic, Singapore*
[b]*School of Informatics & IT, Temasek Polytechnic, Singapore*
*lifern@tp.edu.sg

Abstract: In recent years, the growth of online educational programs has been stimulated by the advancement of the Internet and learning technologies which have transformed the educational landscape. Massive Open Online Courses (MOOCs) have gained much popularity over the past few years and have changed the way people learn. In this qualitative study, we interviewed 14 academic staff at a polytechnic in Singapore and examined the key factors that motivated them to sign up for learning via MOOCs as well as factors that affected their choice of MOOC subjects. Our participants consisted of two distinct groups of academic staff, those who completed the MOOCs they signed up for and those who did not complete. We discovered that the two groups of participants were motivated differently when signing up for MOOCs. We also investigated the factors that led to successful completion of MOOCs by one group of academic staff and the challenges faced by the other group that cause their incompletion. Finally, we asked academic staff for their recommendations on how they think MOOCs could be made more engaging and adaptive to learners with different learning needs and styles. Their recommendations on how MOOCs should be administered, delivered and assessed are presented in this study.

Keywords: MOOC, motivation, engagement, adaptive

1. Introduction

Massive Open Online Courses (MOOCs) cast education in a new paradigm, leveraging on the pervasiveness of Web access and the potential of social learning. By making education available to everyone with access to the web, MOOCs break the boundaries of economic access, time and geographical location faced by traditional education. What was only available to a privileged group of learners is now available to tens of thousands of students. A MOOC generally carries no charges, no prerequisites, no predefined expectations for participation and no formal accreditation. It aims to reach out to anyone who is interested to learn something new. Today, in addition to taking some MOOCs without charge, participants may pay a fee for courses that lead to a certificate (Waldrop, 2013).

Online learning through MOOC is one of the emerging technologies for learning in education. According to Wikipedia (2014) emerging technologies learning is defined as "technical innovations that brings out new territory in some significant way". MOOC has evolved as an emerging technology into new pedagogy that will have an impact on teaching and learning. NMC Horizon Report (2013) stated that higher education will see widespread adoption of MOOC and tablet computing.

The pedagogy that MOOCs employ is different from traditional online learning. MOOCs can be delivered synchronously on a predefined schedule allowing learners to do his or her lessons without geographical boundaries and at his or her convenience. Instead of making a 45-60 minute recorded lecture online for learning, short lecture modules, each lasting 12-15 minutes are used. The shorter duration provides learners with convenient blocks to complete the learning that could be easily fitted in their schedule. Assessments and grading are automated within the platform. MOOCs are used for continuing education objectives by learners who wish to supplement their previous education with skills-enhancement or personal challenge purposes.

MOOCs began in 2011 when Stanford professor, Sabastian Thrun and Director of Research at Google, Peter Norvig taught "Introduction to Artificial Intelligence" online with 160,000 students enrolled and 20,000 completing it. Professor Thrun later resigned and founded Udacity. Thrun's colleagues at Stanford, Daphne Koller and Andrew Ng established their own platform, Coursera, in February 2012. Harvard's and MIT's EdX started in May 2013. Other MOOCs soon followed in the wake of these MOOCs (Nanfito, 2014). Today, institutions are offering a variety of courses on MOOC platforms such as Coursera, Canvas Network, Udacity and EduKart. There are other MOOC-like online platforms like Fathom, Sunoikisis and Connexions.

Enrolments in online education are increasing substantially although retention and completion rates remain low in the face of declining enrollment in higher education (Allen & Seaman, 2013). The 2013 Babson report by Allen and Seaman (2013) shows that more than six million students in public, private and for-profit educational institutes in the United States took at least one online course in the fall 2011 term. Other key findings from the report include low completion rates is barrier for the growth of online learning and 88.8 percent of academic leaders surveyed believe that student lack of discipline in online courses is an obstacle to growth. Cognitive, psychological and emotional connections to feel, think and behave are required for the online environment (Lehman & Conceicao, 2010). Concerns for motivation and factors for engaging learners have to be taken into consideration to help students stay motivated online. It is understandable that a few common reasons for student dropout are related to feelings of isolation, technology disruption, lack of support from faculty, lack of clarity in instructional direction, lack of social interaction and so on.

> Learners in the twenty-first century have been Web consumers for much of their lives, and are now demanding online instruction that supports participation and interaction. They want learning experiences that are social and that will connect them with their peers.
>
> West & West, 2009, p.2

It is also important to understand learners' characteristics, their learning behaviors and skills in the 21st century. The ECAR Study of Undergraduate Students and Information Technology, 2012 (EDUCAUSE, 2012) indicated almost 9 in 10 students own laptops, more than 60% of the students own smartphones and 15% of the students possess a tablet in United States. In a survey conducted by Infocomm Development Authority of Singapore (2014) in the year 2012, 99% of the individuals in the age group between 15 - 24years old who have used a computer and used the internet for the past 12 months. This reflects the behaviour of the generation of learners born in this digital era. Changes in student behavior due to technology usage bring new demands for learning and teaching. Students in the 21st century are IT savvy and are comfortable with technology. The use of Internet has become a norm and may be a way of life for students. They demand greater autonomy of their own learning and the addition of technologies has met their learning needs and preferences (Prensky, 2005). The infusion of information and communication technologies in teaching and learning has open up a wide range of opportunities for creating new kinds of learning activities and experiences. Technology is no longer the problem, but what to do with them to succeed in the new learning environment in this digital era (Carr, 2011). The advancement and adoption of technology for teaching had also transformed the role of the teacher. Kearsley (2000) wrote that the role of the instructor in online classes is to ensure high degree of interactivity and participation through careful design of learning activities which result in engagement with the subject matter and with the students. We felt that there is a need to explore ways to motivate and engage students to help them succeed in the online classroom.

2. Purpose of the Study

Motivation is one of the critical success factors leading to course completion in MOOCs. Motivated learners are more likely to engage in learning activities, participate in online discussion and ultimately, succeed in the course. Thus it is important to understand what motivates online learners, especially academic staff. In this study, key factors that influenced academic staff in their motivation and engagement in MOOCs and how these factors were embedded in the design of learning elements in MOOCs would be investigated. Furthermore, the contributing factors that have motivated academic staff to complete or have hindered their learning progress in MOOCs would be discussed as well. With these in mind, the research questions explored in this study are as follows:

1. What were the contributing factors that led the academic staff to sign up for MOOCs?
2. What were the contributing factors that led to successful completion of MOOCs among the academic staff?

3. What were the contributing factors that caused academic staff to drop out of MOOCs they had signed up for?
4. How had the instructors of the MOOC adapted to the different learning needs of participants?

3. Methodology

3.1 Method

Semi-structured interviews were used to collect data for this research. This qualitative inquiry is well suited for educational research as it enables deep exploration. Interviewees have the freedom to share their experiences and the interviewer retains control of the interview at the same time (Drever, 1995). It also provides the interviewer the freedom to explore general views or opinions in detail (Robson, 2002). Prior to the interview, pilot interviews were conducted to ensure that the set of questions used was effective in fulfilling the purpose of this study.

3.2 Participants and Settings

This study took place at a polytechnic, post-secondary institute, in Singapore. This is one of the five polytechnics in Singapore. There are six schools of study in this polytechnic each offering several diplomas in their specific domain of study. Purposive sampling was used to choose the sample consisting of participants who were appropriate for the study to provide rich information for researcher to develop a detailed analysis on the central phenomenon under study. The research participants were categorized into two groups; those that obtained statements of completion for the MOOCs they signed up for, we call this group of participants the "Completed" group; and those who went through at least 10 hours of the MOOC they signed up for but did not obtain statements of completion, we term this group of participants the "Attended" group. All participants were academic staff from the same school at the polytechnic, each teaching different subjects offered by the school.

Sampling was done according to the matrix as shown in Figure 1. We selected at least three participants from each category to ensure that each category was represented. Each category consists of a fair mix of appointment and non-appointment holders. Appointment holders refer to academic staff who hold an academic appointment like course manager, course coordinator or section head for the course. Non-appointment holders refer to academic staff whose main focus is on teaching. Both appointment and non-appointment holders are heavily involved in academic related matters, from course planning, course design to delivery, we believe that their inputs would be beneficial to this study. In all, 14 staff members were interviewed for our research.

Staff Type	Completed Group (Obtained Statement of Completion of any MOOC)	Attended Group (Completed more than 10 hours of any MOOC)
Appointment Holder (Managers/Coordinators/ Section Heads)	At least 3 participants	At least 3 participants
Non-appointment Holder	At least 3 participants	At least 3 participants

Figure 1. Quota sampling criteria of research participants.

4. Findings

Our participants enrolled in MOOC for a variety of reasons, some have more initial interest in upgrading themselves than others. A key principle to the framework of self-determination theory (Ryan & Deci, 2000) is that individuals enjoy activities when they believe they have autonomy over some aspects of them. Individuals who are self-determined perceive they have ability to make choices over their actions and have been shown to have augmented conceptual learning, positive attitude towards challenging tasks and increased motivation to attend lessons which resulted better performance (Filak & Sheldon, 2008). When it came to the motivation for signing up for learning via MOOC, there were noticeable differences between the "Completed" group and the "Attended" group.

Many of the academic staff who completed their MOOCs cited professional curiosity as one of the factors that prompted them to sign up for MOOCs. They were keen to find out how a course could be delivered in a fully online mode, to a large number of participants; operational details such as how lessons were organized and delivered and how assessments were administered and graded, how queries could be responded to were also on their lists. In short, they showed enthusiasm to find out how MOOCs worked.

While most academic staff from the "Attended" group signed up for MOOCs with the initial intention of completing it, all eventually did not complete. Only 1 interviewee from this group had foreseen that he would be unable to complete the course when he signed up for the MOOC, but did so anyway with the intention to explore the subject matter. The group of academic staff who did not complete the course mainly took on learning via MOOC because they believe this mode of learning suited their own work schedule. They also saw MOOCs as a rich source of up-to-date materials for subject area of their interest.

Some common motivational factors among the two groups of academic staff include fulfilling work-related goals, taking MOOCs as a personal challenge and as a form of self-enrichment. Only 1 of the interviewees mentioned that he signed up for learning via MOOC because it was free. This suggests to us that the flexibility of choice and convenience of having learning delivered via the Internet outweighs the draw that MOOC was delivered at no costs. Studies have shown that students' motivation is affected by their perception of the usefulness of what they would have been taught (Tabachnick et al., 2008). It was also found that students with long term goals or involved with long term projects who are able to see the significance or bearing in their learning with their future are more motivated as compared to those having short term goals.

The Venn diagram below summarizes the factors that prompted academic staff to sign up for learning via MOOC. There were no noticeable difference between appointment holders and non-appointment holders.

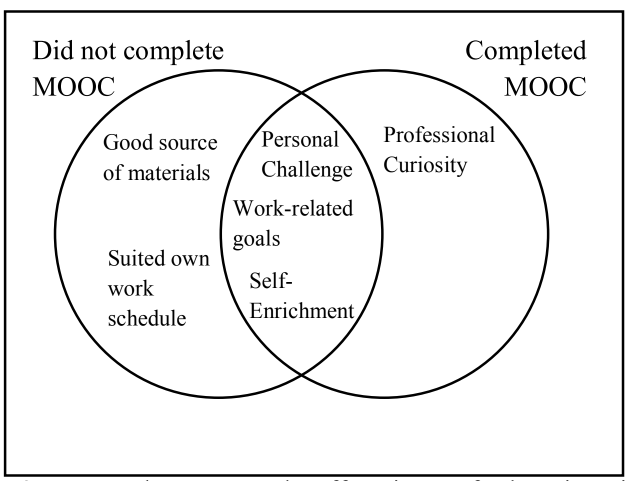


Figure 2. Factors that prompted staff to sign up for learning via MOOC.

We asked our interviewees what motivated them to sign up for the particular MOOCs they did. Even though the topic areas of the MOOCs they took ranged quite widely, the reasons why they chose what they did were fairly consistent when we fed the captured interview transcript into a word cloud generator.

Word cloud shown in Figure 3 was used to identify the prominent terms that were gathered from the interviews. Featured prominently were terms like "reputable", "university", "interest", "subject matter" and "relevance" which suggested that our participants preferred MOOCs offered by reputable institutions on subject matter that were related to their area of work or area of interest.

Figure 3. Word cloud showing why our participants chose the MOOCs they did.

For the group of academic staff who completed their MOOCs, we asked them what helped kept them going, we classify these into intrinsic and extrinsic factors. A summary of the factors, both intrinsic and extrinsic, that helped academic staff we interviewed complete the MOOCs they took can be found in Figure 4.

Intrinsic Motivation Factors	Extrinsic Motivation Factors
- Sense of achievement upon completion. - Relevance and applicability of subject matter to own work area. - Interest in subject matter. - Sense of being challenged by the MOOC. - Curiosity of how the MOOC will end. - Peer and social support.	- Certificate or badge received upon completion. - Regular email reminders of upcoming tasks. - Availability of course calendar showing important dates. - Organization and design of materials. - Sequencing of topics. - Right pitching of demands of assessments. - Timeliness of feedback on how to improve. - Alignment of course schedule to own work schedule. - Passion of instructors / tutors.

Figure 4. Factors that helped academic staff to complete MOOCs they signed up.

The most cited intrinsic factors were the anticipated sense of achievement upon completion of the MOOC and the relevance and applicability of subject matter to own work area.

The most cited extrinsic factors that kept them going were the alignment of the MOOCs' schedules to their own work schedule as well as the course design elements such as the organization and design of materials, sequencing of topics, right pitching of demands of assessments. Course administrative tools such as regular email reminders and availability of course calendar also helped our interviewees plan their schedule ahead to accommodate important submissions and to manage workload for their MOOCs.

For the group that did not complete their MOOCs, we found out what were some challenges they faced. These are classified into 2 categories, personal and course-related as shown in Figure 5.

Personal challenges	Course-related challenges
- Work commitment got in the way.	- Course had too many assessments. - Demands of MOOC were misrepresented. - Coverage of MOOC was misrepresented. - MOOC assumed prior knowledge which staff did not have. - MOOC turned out to be uninteresting. - Assessments were too complex. - Could not understand instructor's accent.

Figure 5. Challenges faced by academic staff.

Unsurprisingly, many of our interviewees who did not complete their MOOCs failed to do so because of work commitments, this was particularly apparent for the appointment holders among the interviewees. Recall that all our interviewees were academic staff, which meant that there was less flexibility in making time for MOOCs since they were bounded by their own teaching time-tables and the general academic calendar. Some interviewees also made conscious decisions to give up on MOOCs which they felt were misrepresented in terms of time commitment, assessment demands and/or topics covered.

Based on their experience of learning via MOOCs, both categories of our interviewees agreed that at present, MOOCs are not able to adapt to the different needs and learning styles of learners. We then asked our interviewees for suggestions on how they would design MOOCs such that different learning needs and styles could be catered for. Their recommendations were summarized in the next section.

5. Recommendations

In this section, some recommendations based on the academic staff experience as a learner of MOOCs are collated from the research findings.

5.1 Provide flexible start dates for MOOCs

"Most of our students today are older, are working and need more flexible schedules" (Palloff & Pratt, 2001, p.109). At present, MOOCs are either self-paced or institution-paced. In the former, students can choose to start the course at any time, there are often no deadlines for assessments and all materials are available to student once he begins the course. The drawback of this is that students would be at different stages of completion compared to his peers. The social learning aspect of MOOC would diminish. There will also be no reminders on upcoming tasks or deadlines. In the latter arrangement, institution offering the MOOC will decide when the course will start. The problem with this design is that students' busy period may coincide with the submission deadlines of the MOOC. Our research participants recommends for MOOC providers to consider providing courses with flexible start dates but not on a self-paced mode. This meant that students would have the flexibility to choose when they will start the MOOCs, once started the platform will work out a personalized calendar based on the MOOC's original design. This would give students greater autonomy in deciding when they would embark on their learning. This would not only increase learner's motivation through autonomy-supportive practice (Reeve & Jang, 2006) but also mitigating the problem of schedule clashes.

5.2 Provide flexible duration for MOOCs

An alternative to having flexible start dates for MOOCs is for MOOCs to have flexible durations. Presently, all institution-paced MOOCs also have institution determined duration. Our research participants recommend that MOOC providers allow different students to have different course duration for the same MOOC. The duration could be self-determined, based on learner's assessment of the demands of the MOOC and their own aptitude; or it could be determined based on students' performance for a diagnostic test, administered in the early weeks of the course. This provision of perceived control over the duration of the MOOC they are taking would be beneficial for students to stay motivated in their learning (Ryan & La Guardia, 1999).

5.3 Provide different track of study

Our research participants recommend that MOOC providers allow learners to determine their desired track of study with differing levels of difficulty. A learner who wishes to have a gentle introduction to the subject matter could opt for an introductory track while another who wishes for in-depth knowledge of the subject matter could opt for an advanced track of study. Learners have autonomy or sense of choice and feel controlled over their actions are more self-determined (Reeve et al., 2003). Some aspects of the materials and assessments could be over-lapping; students of different tracks could interact and learn from each other through the existing collaboration platforms such as discussion forums.

5.4 Standardize instructional and presentation format

Many of our research participants had taken more than one MOOC. Feedback received indicated that different instructors organize their materials in vastly different manner. Our research participants recommend for MOOC instructors to adopt a standardized format by which materials are organized. Our participants believed that this consistency in structuring how course information and materials are presented would have a positive impact on their learning. Having a standardized and consistent instructional or presentation format will make it easier for learners to create a mental image of what to expect from the course and help them manage course workloads. We recognize that this would require extra organization effort by the instructor to rework and restructure instructional style but believe that this would be a worthwhile endeavor since an organized learning environment that provides relevant, consistent, practical and timely materials to meet learners' needs, following the principle of easy to use and simple to use are important aspects to keep learners motivated.

5.5 Provide transcript for video lectures or audio lectures

Aside from recommending standardization of presentation and instructional format, our research participants also recommend that MOOC instructors consider providing a variety of learning materials. At present, materials for MOOCs mostly take the form of video lectures and lecture notes. Our research participants suggest for audio files and transcripts of video lectures to be made available as well. The former would be more suitable for people who prefer to learn on the go, using their mobile devices while the latter would cater to the group of learners who prefer to read rather than watch videos. Certainly, more effort would be needed to prepare the materials and it is more difficult to make any changes to the materials in future.

5.6 Create more opportunities for collaborative learning

> The learning community is the vehicle through which learning occurs online. Members depend on each other to achieve the learning outcomes for the course.
>
> Palloff & Pratt, 2007, p.40

Instructors could consider incorporating more opportunities for collaborative learning in the course design. In the constructivist perspective, learning is being viewed as an active process whereby construction of knowledge takes place through social interactions and collaborative work with each other (Vygotsky, 1978). Students grasp their own understandings and construct knowledge through interactions based on what they already know and believe (Richardson, 2003). Moreover, students should be able to choose their collaborative learning partners. They were more motivated if they have the freedom to choose their working partners as compared to group assigned by the instructor (Ciani et al., 2008).

5.7 Provide intelligent progress tracking

Presently, most MOOCs do not track the progress of individual students. Progress bars are typically associated to the course schedule rather than students' progress and email reminders are generally time-based rather than activity based. Our participants recommend for more intelligent progress tracking so that personalized reminders which based on individuals' completion of task could be delivered by the system. While we foresee this leading to greater administrative challenges for MOOC instructors, we also see the potential for the same set of triggers to be used for adaptive delivery or adaptive assessment.

5.8 Leverage on M-Learning

Finally, our research participants recommend that MOOC providers leverage on the potential of mobile learning and the pervasiveness of mobile devices among students. This could be the provision of mobile application linking learners to the MOOC platform and courses, syncing course dates to the calendar in learner's mobile devices; or taking advantage of existing collaboration tools such as group messages.

6. Limitations of Research

A limitation of this study is its possible lack of generalizability. Though the sample size is large enough for such qualitative study, the findings are typically relevant to the specific group of learners under investigation with its own characteristics. While this study is academic staff-specific, our goal is to share recommendations on the development of MOOCs for those who are interested to offer courses via this mode. We recommend for further studies to be done on more samples of MOOC learners to gain more objective inputs.

7. Conclusion

There are many factors that motivated learners to sign up for learning via MOOCs and successful completion for their choice of MOOC subjects. Instructors could promote individual interest by (1) providing learners with opportunities to have control over their learning (2) relating the usefulness of content to achieve their goals (3) creating a warm and personalized presence to help learners feel connected and engaged (Osborne et al., 2007).

The growth of online learning options continues to increase and will have an impact on the shape of higher education. Learners in the 21st century want learning experiences that support participation and social interaction that will connect them with their peers.

In this study, strategies on how MOOCs could be made more engaging and adaptive to learners with different learning needs and styles were made. We hope that these strategies and methods could help instructors design an online learning environment that meets the needs and learning behavior of students in the 21st century. By integrating support, instructor could help learners to have greater insights about effective time management, prioritizing and stay motivated throughout the course. In addition, they could help learners with different learning needs to identify a pathway for successful online learning.

References

Allen, I., & Seaman, J. (2013). Changing Course: Ten Years of Tracking Online Education in the United States. *Babson Survey Research Group*. Retrieved July 7, 2014, from http://www.onlinelearningsurvey.com/reports/changingcourse.pdf

Carr, N. (2011). The shallows: What the Internet is doing to our brains. New York: W. W. Norton & Company.

Ciani, K. D., Summers, J. J., Easter, M. A., & Sheldon, K. M. (2008). Collaborative learning and positive experiences: Does letting students choose their own groups matter? *Educational Psychology*, *28*(6), 627-641.

Drever, E. (1995). Using semi- structured interviews in small-scale research: A teacher's guide. Edinburgh: SCRE.

EDUCAUSE (2012). *ECAR Study of Undergraduate Students and Information Technology, 2012*. Retrieved Aug 4, 2014, from https://net.educause.edu/ir/library/pdf/ERS1208/ERS1208.pdf

Filak, V. F., & Sheldon, K. M. (2008). Teacher support, student motivation, student need satisfaction, and college teacher course evaluations: Testing a sequential path model. *Educational Psychology*, *28*(6), 711-724.

Infocomm Development Authority of Singapore. (2014). *Infocomm usage - Households and individuals.* Retrieved July 10, 2014, from http://www.ida.gov.sg/Infocomm-Landscape/Facts-and-Figures/Infocomm-Usage-Households-and-Individuals#4

Kearsley, G. (2000). Online education: Learning and teaching in cyberspace. Belmont, CA: Wadsworth/ Thomson Learning.

Lehman, R. M., & Conceicao, S. C. O. (2010). Creating a sense of presence in online teaching: How to "be there" for distance learners. San Francisco, CA: Jossey-Bass.

Nanfito, M. (2014). MOOCs: Opportunities, Impacts and Challenges. USA: Lexingto, KY.CreateSpace Independent Publishing.

NMC Horizon Report. (2013). Higher Education Edition. Retrieved September 4, 2014, from http://net.educause.edu/ir/library/pdf/HR2013.pdf

Osborne, J. W., Kellow, J. T., & Jones, B. D. (2007). Identification with acedamics, stereotype threat, and motivation to achieve in school. In D. M. McInerney, S. Van Etten, & M. Dowson (Eds.), *Standards in education: Sociocultural influences on motivation and learning* (pp. 119-146). Charlotte, NC: Information Age Publishing.

Palloff, R. M., & Pratt, K. (2001). Lessons from the cyberspace classroom: The realities of online teaching. San Francisco:Jossey-Bass.

Palloff, R. M., & Pratt, K. (2007). Building online learning communities: Effective strategies for the virtual classroom. San Francisco:Jossey-Bass.

Prensky, M. (2005). Engage me or enrage me: What today's learners demand. *Educause Review, 40*(5), 60-65.

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Psychology*, *98*(1), 209-218.

Reeve, J., Nix, G., & Hamm, D. (2003). The experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of Educational Psychology*, *95*, 375-392.

Richardson, V. (2003). Constructivist pedagogy. *The Teachers College Record*, *105*(9), 1623-1640.

Robson, C. (2002). Real world research: A resource for social scientists and practitioner researchers (2nd ed.). Oxford: Blackwell.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78.

Ryan, R. M., & La Guardia, J. (1999). Achievement motivation within a pressured society: Intrinsic and extrinsic motivations to learn and the politics of school reform. In T. Urban (Ed.), *Advances in motivation and achievement*, *11*, (pp. 45-85). Greenwich, CT: JAI Press.

Tabachnick, S. E., Miller, R. B., & Telyea, G. E. (2008). The relationships among students' future-oriented goals and subgoals, perceived task instrumentality, and task-oriented self-regulation strategies in an academic environment. *Journal of Educational Psychology*, *100*(3), 629-642.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Waldrop, M. (2013). Massive Open Online Courses, aka MOOCs, transform higher education and science. *Scientific American*. Retrieved July 7, 2014, from http://www.scientificamerican.com/article/massive-open-online-courses-transform-higher-education-and-science/

West, J., & West, M. (2009). Using wikis for online collaboration. San Francisco: Jossey-Bass.

Wikipedia. (2014). *Emerging technologies*. Retrieved September 4, 2014, from http://en.wikipedia.org/wiki/Emerging_technologies

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Effects of Gender Differences and Learning Performance within Residence Energy Saving Game-based Inquiry Playing

Ugyen DORJI[a], Patcharin PANJABUREE[b*] & Niwat SRISAWASDI[c]
[a]*Gelephu Higher Secondary School, Bhutan*
[b]*Institute for Innovative Learning, Mahidol University, Thailand*
[c]*Faculty of Education, Khon Kaen University, Thailand*
*panjaburee_p@hotmail.com

Abstract: Energy Saving in Residence is one of the important topics in Energy Education, whereby students are required to understand the factors of energy consumption and conservation. Several researchers have studied on how to use digital games to improve understand the topic. However, not every student understands is improved by using these kinds of learning assistance. Therefore, a study on effect of gender differences and learning performance within energy saving games have become a need to address the issue. This paper has attempted to develop a digital game based on inquiry-based learning called Residence Energy Saving Battle (RES-battle). Further, the paper examines the effect of gender difference in students' learning performance and the attitude towards the RES-battle. The result shows that students' learning performances on energy consumption and conservation that significantly improved after they participated in the RES-battle. The result also revealed that the RES-battle can decrease the difference between female and male learning performances reasonably and attitude toward the RES-battle. From the findings, this paper has implication for the development of students' residence energy consumption learning within the digital game.

Keywords: Game-based learning, pedagogy, gender study, energy education

1. Introduction

Electricity, a one of longest-used demand, is widely used for lighting or heating purposes and to make electrical equipment works. Due to widespread use, most of consumers are careless in its management. The amount of electrical energy usage occurs not only from residence budget and size but also from residents' behavior in using electrical equipment (Moll, Noorman, Kok, Engström, Throne-Holst, and Clark, 2005). In the past decade, several researchers attempted to provide the way to decrease the electricity usage in each residence. For example, Abrahamse, Steg, Vlek, and Rothengatter (2005) focused on altering the residents' behavior in the use of electrical equipment in appropriate way. Maharaj-Sharma (2012) suggested that student energy learning is need to apply and link to outside school and daily life. Consequently, learning about energy consumption and conservation in school science has become important in current research.

Although energy consumption and conservation is one of important physics concept in school science. Generally, this topic is a part of physics teaching and learning about electricity. In the conventional physics class, student learned electricity as attribute of electronic and consumption and conservation of energy in a form of employed energy calculation, but not as procedures as it is recognized in contemporary physics. As such, it is an instructional challenge to motivate student learning the concept meaningfully. Several methods and strategies have been used to make the concept of energy consumption and conservation more meaningful for understanding to the students and for encouraging them linking the concept to daily life. For example, Chiu, Chou, and Liu (2002) suggested that although using analogies or metacognition might help the learning process more permanent, but there is no significant conceptual development by these approaches. Slotta and Chi (2006) recommended that students can insight what concepts from their own understanding. Thus, teaching students to use electric energy efficiently in their daily life to encourage them to construct conceptual understanding of energy consumption and conservation by themselves might be useful way for saving energy in residence. In other words, digital interactive learning could be a useful way for conceptual learning of energy consumption and conservation and for simulating practices of saving energy in residence.

On the other hand, digital interactive learning is generally developed as multimedia learning units with simulations and games in many disciplines such as in mathematics, science, engineering, humanities, and social sciences (Cai, Lu, Zheng, and Li, 2006; Eck, 2006). Recent research attempted to purpose educational computer game which has advantage in students' habits and interests (Gee, 2006; Prensky, 2007) and indicated that the use of digital game is another useful method to improve energy consumption and conservation learning (Yang, Chien, and Liu, 2012). Playing educational computer game requires prior knowledge or pre-existing learning experience that helps students to apply the knowledge to make decision related to realistic situations (Papastergiou, 2009) and can be considered as a learning tool for teaching the factual information as well as worksheet activities (Spraggins and Rowsey, 1986). Therefore, developing a digital learning of energy consumption and conservation that integrates game-based learning technologies and pedagogy into learning of energy consumption and conservation may be a useful way to improve students' learning performance on the topic. However, the successful usage of pedagogy-driven digital game depends on the digital game, the learning strategies, and human factors. Among various human factors, gender difference play an important role when playing digital game affecting on learning performance (Paraskeva, Mysirlaki, and Papagianni, 2010).

In summary, pedagogy-driven digital game, gender differences, and conceptual learning performance are critical to learning saving energy in residence in which students are required to understand the energy consumption and conservation meaningfully. In this vein, the aim of this study is to examine the effects of gender differences on students' conceptual learning performance when playing the *Residence Energy Saving Battle (RES-battle)* digital game and attitudes toward the RES-battle.

2. Background and Motivation

In Bhutanese education system, science subject is introduced from primary level onwards and when it reaches in higher secondary level it is segregated into three disciplines; namely Physics, Chemistry and Biology. The aim of having this is to make the curriculum more relevant to the need of the society or localization and to bring a shift in the teaching style from teacher centered to learner-centered. However, in many schools in Bhutan teacher-centered instruction is a common practice. Most students in Bhutan lack understanding of concepts and they learn by memorization. Many science teachers feel and our personal experiences from the past has found that the answers that students write in the exams are just regurgitation of what they have learnt by memorization. Any twist or rephrasing of the questions made the students difficult or unable to answer the questions.

In recent years, the education system supplied hundreds of computers all over the nation under the project "Chiphen Rigphel Project" (Information Communication Technology-ICT Project) with grant from government of India. The objective of the project is to equip every student with IT literacy and improve science education. Now the system is stressing more on the practical oriented learning that enable students to acquire the skills that can be applied in the practical context. So, the system has observed that digital learning as one of the ways to achieve that goal. Therefore, there is strong need for teachers to change the instructional practices from teacher centeredness to child-centeredness by practicing effective teaching strategies with integration of digital learning like computer games, animations, simulations, and more, particularly in learning abstract concepts in science curriculum, especially on topic energy consumption and conservation.

Energy education has become very important element to educate students on the basic energy concept on daily electrical energy consumption and conservation. It help students to identify basic factors on which energy consumption depend and provide useful ways for reducing daily energy consumption both in and outside of their home and school. Integrating appropriate learning approaches for teaching the concept is very important for any effective learning process. Thus, introducing the concept of energy education into the ways the students understand is crucial for any educator (Gustafson and Branch, 2002). Moreover, one of effective learning strategies for making students as active leaner is inquiry-based learning approach (Kubicek, 2005). It requires teachers to provide the opportunity for students to observe, gather, analyze, and interpret data which students learn from their findings, explanation, predictions, and communications with peers, which led students to construct their own conceptual knowledge (Krajcik and Blumenfeld, 2006; Kuhn, Black, Keselman, and Kaplan, 2000).

Therefore, in this study, the inquiry-based learning approach was chosen to drive a digital game-based learning to assist students exploring factors related energy consumption, by simulating using electrical appliances for certain durations led to saving energy in residences. These features of the game might help students to improve conceptual learning performance on the topic of energy consumption and conservation. Moreover, when playing digital game, students' gender difference factors play might effect on their learning performance. Consequently, there is a need to examine how the RES-battle affect students' learning performance, and how gender differences affect students' learning performance of conceptual knowledge on energy consumption and conservation. To this end, this study emphasizes on whether gender differences affect students' learning performance by taking the RES-battle.

3. The Digital Game based on Inquiry-based Learning Approach: Residence Energy Saving Battle (RES-battle)

With the benefit of game-based learning in motivating students learning and inquiry-based learning approach in encouraging students explore and construct conceptual knowledge on energy consumption and conservation, this study developed *Residence Energy Saving Battle (RES-battle)* in which the inquiry-based learning approach drives activities in computer-based game playing. The RES-battle is designed in line with practical situations of energy consumption in our daily life. To save money and minimize energy consumption, these help the students in developing sense of awareness on energy conservation. The procedure of RES-battle is following steps:

- **Step 1:** The objectives of the RES-battle are introduced to the students, such as the students should be able to know how energy consumption is calculated in term of money and also how can they apply those factors in monitoring power bill by playing the RES-battle.
- **Step 2:** The rules and basic functions of the RES-battle are demonstrated to the students, such as the RES-battle is designed with scenarios of home comprising electrical appliances that are commonly operated. It is divided into two different-difficulty levels in which the students can go to the second level by passing the first level. Each level has 30 playing-time second. To get into another level, the students should have enough income/budget earned from current stage. In doing, the students should hunt/collect hidden coins which depend upon wattage of those appliances in that room. The students have to take cautious decisions on selecting/clicking on appliances to get those coins because every time they collect coins, it activates the appliance which consumes energy, and they have to pay at the end.

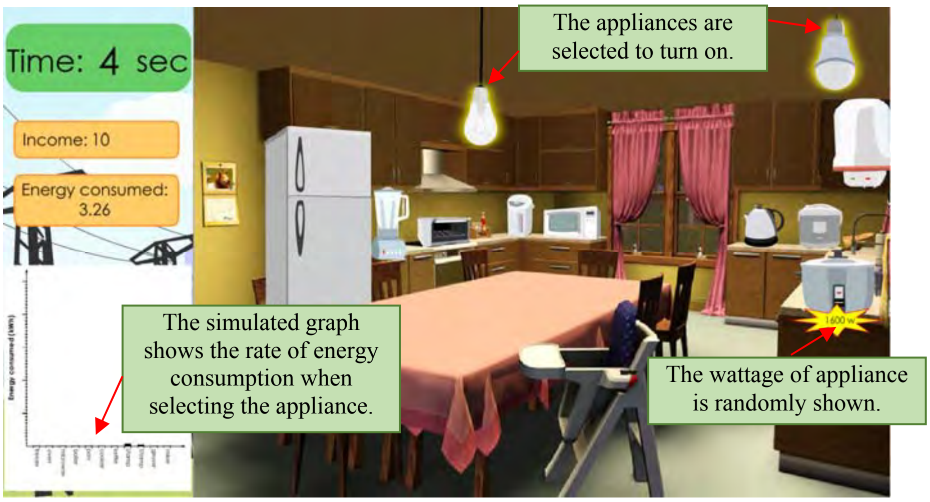


Figure 1. Illustrate playing the Residence Energy Saving Battle (RES-battle) in the first level

- **Step 3:** While playing the RES-battle, whatever coins are accumulated and how much energy has consumed are automatically calculated and displayed in graphical format at the side of the game screen as shown in Figure 1. These help students to apply their theoretical knowledge, skills and strategies to make decisions in their assigned role which promotes the understanding of concept. The wattage of each appliance will be shown randomly at certain interval. The value of hidden coins in each appliance remains constant whether it is chosen at the beginning or at last. In this step, the students are encouraged to inquire factors of energy consumption which are wattage of the appliance and duration of usage.

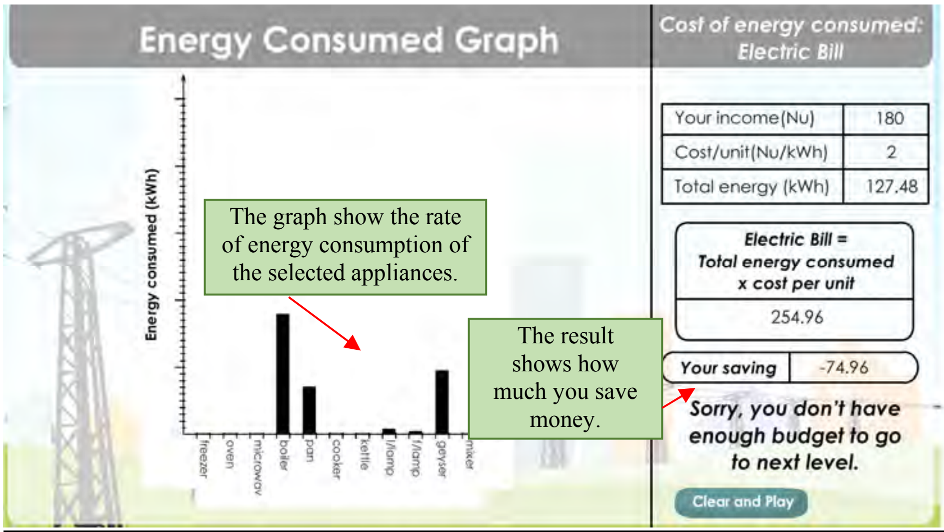


Figure 2. Illustrate results of playing the Residence Energy Saving Battle (RES-battle) in the first level

- **Step 4:** At the end of 30 second, the RES-battle automatically is stopped and displays "Time Up", after that simulated graph for energy consumption will be displayed for those appliances that the students operated during the game as shown in Figure 2. Based on that, cost for energy consumed is calculated and then saving of the students is shown *(your saving = income – cost of energy)*. If the saving is positive, the player can continue the game in the second level with more difficulty as shown in Figure 3, otherwise they cannot go further or can play again as shown in Figure 2.

4. Research Methodology

To investigate the effectiveness of the RES-battle, we investigate whether the RES-battle improve students' learning performance on energy consumption and conservation; result on this will led us to examine whether gender differences affect students' learning performance by taking the RES-battle and whether gender differences affect students' attitudes toward the RES-battle.

This study was designed by one group pre-posttest design. A total of 68 tenth graders of secondary school students in eastern Bhutan were recruited in this study. They participated in the RES-battle individually, and discussed the situations in the RES-battle with their peer to construct their own knowledge of factors of energy consumption and conservation. Before playing in the RES-battle, the students were asked to take pre-conceptual test of the topic. After finishing learning activities on the RES-battle, they were asked to take post-conceptual test. Moreover, they are asked to respond the attitude questionnaire to clarify the degree of their attitude toward the RES-battle.

To examine whether the RES-battle improve students' learning performance on energy consumption and conservation, pre- and post-conceptual tests were used as the research tools. The validity of the tests were determined by three experienced teachers teaching same subject. Each test contained 20 multiple-choice items, and one point was scored for each correct answer; therefore, the total score of the each test was 20.

To investigate students' attitudes toward the RES-battle, the attitude question adopted from Subba (2011) was used in this study. It consisted of 18 items, which were categorized into three categories: interest, participation, and satisfaction. This questionnaire was measuring using a 5-points Likert scale in which for "Interest" and "Participation" category, levels 5 represents for "Strongly agree", 4 for "Agree", 3 for "Neutral", 2 for "Disagree" and 1 for "Strongly disagree"; whereas in "Satisfaction" category, 5 represents for "Extremely Satisfaction", 4 for "High Satisfaction", 3 for "Moderate Satisfaction", 2 for "Low Satisfaction" and 1 for "Least Satisfaction".

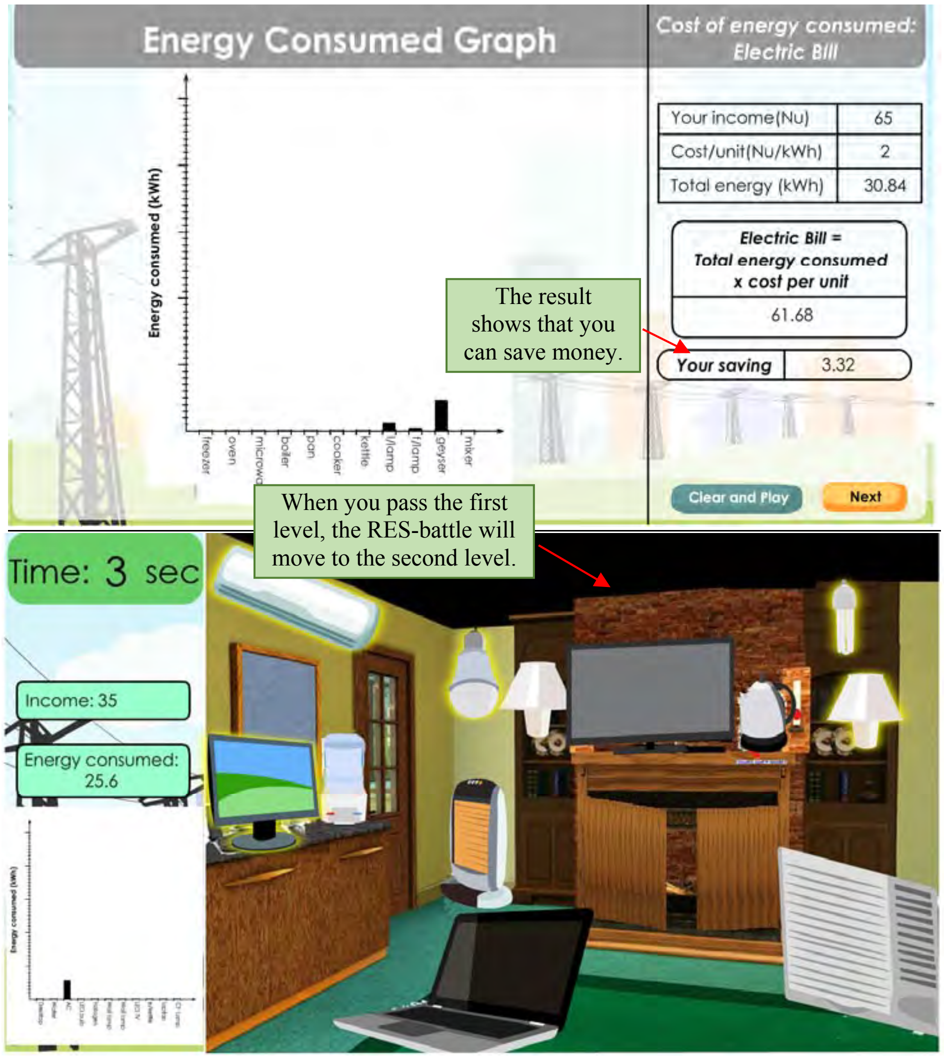


Figure 3. Illustrate success of the first level and move to the second level with more difficulty

5. Results

5.1 Students' conceptual learning performance

5.1.1 Overall conceptual learning performance

Table 1 shows the results of the conceptual test, in term of Mean and Standard Deviation (SD), between pre- and pot-test. The result shows that there was a significant difference between the pre- and post-test on students' conceptual learning of energy consumption and conservation (t = 13.887, p = .000). It indicates that the students performed better in the post-test than in the pre-test; implying that students' conceptual learning were significantly improved after participating in the RES-battle. As shown in Table 1, it seems obviously that the SD in the post-test was little increased from the SD in the pre-test. This result indicates that there is different amount of conceptual knowledge gain after taking the RES-battle; some student has high gain conceptual knowledge and some student has low gain conceptual knowledge. Thus, another data analysis was conducted to clarify the conceptual knowledge gain in the next section.

Table 1: Conceptual learning performance between pre- and post-test

Test	N	Mean	SD	*t*
Pre-test	68	7.56	2.27	13.887**
Post-test	68	13.35	3.03	

$^{**}p < 0.01$

5.1.2 Conceptual learning progression

To clarify the conceptual knowledge gain of the students after taking the RES-battle, this study employed normalized gain of Hake (1998) which defined the <g> as “High gain, <g> ≥ 0.7”, “Medium gain, 0.7 > <g> ≤ 0.3”, and “Low gain, <g> < 0.3”. From sixty-eight students’ conceptual pre- and conceptual post-test scores, the results show that there are 13, 36, and 19 students for high, medium, and low gains respectively. Thus, this results indicates that most of students gains conceptual knowledge of energy consumption and conservation moderately after taking the RES-battle. For overall result, as shown in Table 2, the conceptual score of pre- and post-test, the <g> is 0.47 indicating that the students have conceptual learning progression of their learning by gaining better conceptual knowledge after playing the RES-battle.

Table 2: Conceptual learning progression by the average normalized gain <g>

Conceptual test (Total score = 20)	N	Mean	S.D.
Pre-test	68	7.56	2.27
Post-test	68	13.35	3.03
<g>	0.47		

5.2 Students’ attitudes toward the Residence Energy Saving Battle (RES-battle)

Table 3 shows the descriptive statistics of the students’ attitude toward the RES-battle. Overall analysis from the questionnaires reveals that the students have rated “High Agree” for the RES-battle for being able to develop interest in learning the concept of energy consumption and conservation in residence. Moreover, the students respond that they were “Highest Agree” towards the RES-battle in which using the inquiry-based learning drives learning activities on the topic. Because the RES-battle encouraged them to explore and construct conceptual knowledge. In addition, the students rated that they have “High Satisfaction” in learning concept of energy consumption and conservation through the RES-battle.

Table 3: Students’ attitude degree after taking the RES-battle

Attitude Aspect (N=65)		**Mean**	**SD**	**Interpretation**
Interest	1. I enjoy learning very much with the RES-battle.	4.33	0.63	High Agree
	2. I become more curious and observant in the class when the lesson is integrated with the RES-battle.			
Participation	3. I enjoy participating in class activities when the lessons are taught using the RES-battle.	4.45	0.57	High Agree
	4. The RES-battle make me more attentive in the class.			

Attitude Aspect (N=65)		Mean	SD	Interpretation
	5. Integration of the RES-battle in the lesson promotes better interaction amongst friends and teachers.			
Satisfaction	6. It is easier for me to understand the content with the RES-battle.	4.40	0.42	High Satisfaction
	7. I get learning satisfaction when I learn the lesson with the RES-battle.			
	8. The RES-battle in the lesson helps me to develop confidence in learning electrical energy calculation.			
	9. I found the RES-battle useful in visualizing the concepts.			
	10. The RES-battle in learning helps me to think and analyze the real things in world.			
	11. It helps me to develop the relevance between the course and real world situations.			
	12. The RES-battle allows me to develop skills needed in the real world.			
	13. I like the way the teacher uses the RES-battle to teach energy consumption by various household appliance lesson.			
	14. The use of the RES-battle in the lesson helps me to build confidence in understanding the concept of energy consumption clearly.			
	15. I like the RES-battle in learning electrical energy consumption and conservation sessions because it enables me to learn faster.			
	16. I gain confidence when I learn the lesson using the RES-battle.			
	17. I like electrical energy lessons with the RES-battle because the lessons are interesting, informative and help to visualize the abstract concepts of energy better.			
	18. I like the RES-battle integration in all the subjects to help enhance our critical thinking.			

5.3 Gender Differences

5.3.1 Gender differences in overall conceptual learning performance

Table 4 shows the results of students' conceptual learning performance, in term of mean and Standard Deviation (SD), between females and males for conceptual learning performance. The results of the pre-test and post-test show that there were no significant difference between females and males for their conceptual learning performance. In other words, conceptual learning performance between females and males had no difference in the post-test. It indicates that females and males improved at a faster rate and were not statistically different in the end of the RES-battle.

Table 4: Conceptual learning performance between females and males in the pre- and post-test

Test	Gender	N	Mean	SD	*t*	*p*
Pre-test	Female	38	7.32	2.28	.994	.162
	Male	30	7.87	2.25		
Post-test	Female	38	12.95	2.94	1.249	.108
	Male	30	13.87	3.10		

5.3.2 *Gender differences in different sizes of conceptual learning gain*

To examine the gender difference for the high, medium, and low gain size of conceptual knowledge, analyzes were undertaken between boys and girls in the pre- and post-test for the three group in different sizes of conceptual learning gain. Table 5 shows the results of these analyzes. The results show that there were no significant difference between females and males in the pre- and post-test of the three gain size groups. These findings reveal that this females and males improved but were not statistically different in the end of the RES-battle.

Table 5: Performance of the high, medium, low gain size between females and males in the pre- and post-test

Size	Test	Gender	N	Mean	SD	*t*	*p*
High	Pre-test	Female	7	5.57	2.99	1.543	.075
		Male	6	8.17	3.06		
	Post-test	Female	7	17.00	.816	1.721	.056
		Male	6	17.67	.516		
Medium	Pre-test	Female	18	7.78	2.102	.156	.438
		Male	18	7.67	2.169		
	Post-test	Female	18	13.61	1.787	.494	.312
		Male	18	13.94	2.235		
Low	Pre-test	Female	13	7.62	1.758	.616	.573
		Male	6	8.17	1.941		
	Post-test	Female	13	9.85	1.068	.023	.491
		Male	6	9.83	1.329		

5.3.3 *Gender differences in attitude toward the Residence Energy Saving Battle (RES-battle)*

Table 6 shows the results of students' attitude toward the RES-battle, in term of mean and Standard Deviation (SD), between females and males. The results of the three attitude aspects show that there were no significant difference between females and males for their attitude. In other words, attitude toward the RES-battle between females and males were not difference. It indicates that females and males satisfied equally in the end of the RES-battle.

Table 6: Attitudes between females and males after taking the RES-battle

Attitude Aspect	Gender	N	Mean	SD	*t*	*p*
Interest	Female	37	4.38	.57	.655	.256
	Male	29	4.27	.70		
Participation	Female	37	4.50	.55	.748	.223
	Male	29	4.39	.61		
Satisfaction	Female	37	4.33	.40	1.566	.061
	Male	29	4.50	.43		

6. Conclusion and Implications

This study examined the effectiveness of a digital game in which the inquiry-based learning approach drives activities within the game. While playing the game, students were encouraged to explore elements in the game, discuss what they found with their peer, and construct their own knowledge. Although Gee (2007) and Unlusoy, de Haan, Leseman, and van Kruistum (2010) revealed that males interests in digital games is more than females do. But the findings from this study highlighted that there were no difference in the conceptual tests and attitude towards the developed game between females and males. Moreover, when the participants were categorized based on their conceptual knowledge gain size into high, medium, and low gain size, there were also no difference in the conceptual tests, and attitude toward the developed game between females and males. These findings could imply that the difference between females and males was reduced after taking the developed game. In other words, the developed game enabled the students who gained conceptual knowledge in different size to make significant conceptual learning improvement.

Therefore, to decrease the gap between females and males on conceptual learning performance when using the digital game-based learning, there is a need to develop the game that can provide the opportunities for interaction on game screen and also with other students. The road of game playing needs to accommodate the encoding and decoding of graphics through the game and the elements that promote students inquired evidence for constructing the knowledge.

Acknowledgement

This research project is supported by Thailand International Cooperation and Development Agency (TICA) scholarship.

References

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, *25*(3), 273–291.

Cai, Y., Lu, B., Zheng, J., & Li, L. (2006). Immersive protein gaming for bio edutainment. *Simulation & Gaming*, *37*(4), 466–475.

Chiu, M.-H., Chou, C.-C., & Liu, C. J. (2002). Dynamic processes of conceptual change: Analysis of constructing mental models of chemical equilibrium. *Journal of Research in Science Teaching*, *39*(8), 688–712.

Eck, R. V. (2006). Digital Game- Based Learning: It's Not Just the Digital Natives Who Are Restless. *EDUCAUSE Review*, *41*(2), 16–30.

Gee, J. P. (2006). *Are Video Games Good for Learning?* Presented at the Curriculum Corporation 13th National Conference, Adelaide.

Gee, J. P. (2007). *In What video games have to teach us about learning and literacy*. Basingstoke: Palgrave Macmillan.

Gustafson, K. L., & Branch, R. M. (2002). *Survey of instructional development models* (4th ed.). Syracuse, N.Y: ERIC Clearinghouse on Information & Technology.

Hake, R. (1998). Interactive-engagement vs traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Association of Physics Teachers. 66*, 64–74.

Krajcik, J. S., & Blumenfeld, P. C. (2006). Project-based learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 317–334). New York: Cambridge.

Kubicek, J. P. (2005). Inquiry-based learning, the nature of science, and computer technology: New possibilities in science education. *Canadian Journal of Learning and Technology, 31*(1).

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*(4), 495–523.

Maharaj-Sharma, R. (2012). What are upper primary school students' perceptions of science? The Trinidad and Tobago context. Retrieved from http://uwispace.sta.uwi.edu/dspace/handle/2139/13705

Moll, H. C., Noorman, K. J., Kok, R., Engström, R., Throne-Holst, H., & Clark, C. (2005). Pursuing more sustainable consumption by analyzing household metabolism in European countries and cities. *Journal of Industrial Ecology*, *9*(1-2), 259–275.

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education*, *52*(1), 1–12.

Paraskeva, F., Mysirlaki, S., & Papagianni, A. (2010). Multiplayer online games as educational tools: facing new challenges in learning. *Computers & Education, 54*(2), 498–505.

Prensky, M. (2007). *Digital game-based learning*. St. Paul, Minn.: Paragon House.

Slotta, J. D., & Chi, M. T. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, *24*(2), 261–289.

Spraggins, C. C., & Rowsey, R. E. (1986). The effect of simulation games and worksheets on learning of varying ability groups in a high school biology classroom. *Journal of Research in Science Teaching*, *23*(3), 219–229.

Unlusoy, A., de Haan, M., Leseman, P. M., & van Kruistum, C. (2010). Gender differences in adolescents' out-of-school literacy practices: a multifaceted approach. *Computers and Education, 55*(2), 742–751.

Yang, J. C., Chien, K. H., & Liu, T. C. (2012). A digital game-based learning system for energy education: an energy conservation pet. *The Turkish Online Journal of Educational Technology, 11*(2), 27-37.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

A Blended Learning Environment in Chemistry for Promoting Conceptual Comprehension: A Journey to Target Students' Misconceptions

Sumarin NIROJ[a] & Niwat SRISAWASDI[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Thailand*
*niwsri@kku.ac.th

Abstract: Blended learning has become a new form of learning and teaching in science education. Researchers have indicated that the blended learning could create more meaningful in learning, support deep-level understanding, and contribute to better learning achievement in the exam. The present article presents a blended learning environment, combination of computer-based interactive lecture demonstration (CBILD) and web-based inquiry science environment (WISE), for chemistry learning of state of matter and phase change. To create the blended environment, 50 twelve-graders were recruited in an investigation of prior knowledge by challenging their epistemological and ontological belief of the chemistry concepts. As such, a series of two-tier conceptual items were administered and then the students' unscientific conceptions about state of matter and phase change were extracted as an account. The finding showed that they hold various patterns of misconceptions covering the effects of pressure and temperature on arrangement of particle and phase change, and shape of molecule during phase change. Moreover, all of them have no scientific conception that plasma is a state of matter. Many misconceptions seem to due to the fact that they could not realize interaction between observable and unobservable level of chemical phenomena because of its complexity and abstraction. To make chemistry more accessible and meaningful for student learning, CBILD learning process was design in emphasizing macro-level representation and WISE was design to support sub-micro-level and symbolic representations in chemistry as face-to-face and on-line learning experience respectively. The learning process of blended learning environment was illustrated and described in a sequence for implementing in chemistry classroom. This could be an implication for researchers and teachers how to create blended learning environment which could improve teaching and learning strategy into a new form for science classroom, and it might enhance the change of student's misconceptions and their mental model development in science.

Keywords: Blended learning, interactive lecture demonstration**,** web-based learning, microcomputer-based laboratory, chemistry education

1. Introduction

Chemistry is a fundamental science which is abstract and complex by its nature. Due with its nature, students lack of transfer what they learned, e.g. concepts, to real-world problems and everyday life, and they give no meaning to what they have learned (Gilbert et al., 2002; Gilbert, 2006).Although, the chemistry learning activities attempted to link the subject matter with how the world works, the students still have numerous learning difficulties and misconceptions on the subject. Moreover, they merely link their own existing ideas to the new concepts leading to fragmented and fractured understanding (Gilbert and Boulter, 2000). Such misconceptions, especially on the topic state of matter and phase change, are due to the fact that the students could not distinguish between macroscopic and sub-microscopic explanations. The students have also difficulty linking observable phenomena to molecular level interaction (Chang and Linn, 2013). As such, many studies suggested that the molecular visualizations could help students integrate observable, molecular, and symbolic aspects of chemical change (Russell and Kozma, 2005).

With the benefit of computers and technology in science education community, the contemporary technology-based approach has been used to enhance students' conceptual understanding (Vreman-de Olde et al., 2013; Srisawasdi and Kroothkeaw, 2014). Moreover, the inquiry-based learning strategy supported by Information and Communication Technology (ICT) has known as an effective pedagogical approach in science learning. ICT can also enable new ways of education to deliver knowledge directly from teachers to students and the learning can take place anywhere and anytime. Recently, there are many computer-supported inquiry-based science learning environments such as WISE, Co-Lab, Inquiry Island, and nQuire (Sun et al., 2013). The instructions that teachers implement have been changed by conducting the new learning paradigm and innovative educational tools. Active learning and knowledge sharing have been replacing traditional teacher-centered lectures, due to their ability to integrate ICT into the curriculum and their emphasis on meaningful learning (Yen and Lee, 2011).

Blended (hybrid, mixed-mode) learning is referred as learning combination of the classical teaching in classroom (face-to-face learning) and teaching assisted by contemporary technologies. Thorn (2003) described blended learning as a way of meeting the challenges of tailoring learning and development to the needs of individuals by integrating the innovative and technological advances in online learning environment. Moreover, Yapici and Akbayin (2012) revealed that the students who participated in blended learning model could achieve learning outcomes in biology course than those who participated in the traditional teaching method significantly. In additions, they also have positive attitude toward science if the learning environment had incorporated by internet access (Yapici and Akbayin, 2012). As such, the blended learning could be used to overcome the difficulties of practical science.

With the importance of enhancing students' conceptual understanding in science concepts and considering benefits of blended learning environment as aformentioned, the researchers aim to develop an effective blended learning environment for chemistry learning of state of matter and phase change regarding students' specific misconceptions. As such, investigation of the misconceptions before designing of science learning environment could be a strategic way for research and development in science and technology education.

2. Blended Learning Environment

Lecture demonstration method is usually an integral part of a university education. This method helps students to visualize the material through students' understanding of the subject. Zimrot and Ashkenazi (2007) indicated that interactive lecture demonstrations (ILD) could be used to make more accessible of scientific conceptual understanding because of potential source of anomalous data that can trigger cognitive process of conceptual change. Moreover, ILDs are fun to do and provide concrete examples of abstract concepts. Nevertheless, using only the traditional approach to education, where the transfer of knowledge is achieved mostly by lecturing, has a number of shortcomings, because the students are not motivated enough to acquire knowledge actively (Hoic-Bozic, Mornar and Boticki, 2009; Burnham, 2001). To cope with this issue, the computer-based interactive lecture demonstration (CBILD) could be a novel of instruction to encouraging students learning and is one component of blended learning environment in this study.

Moreover, to enhance knowledge integration, the use of computer visualizations in inquiry activities can enhance students' conceptual understanding of molecular level interactions. web-based inquiry science environment (WISE), which is one of web-based learning environment and is one component of blended learning environment in this study, was developed based on the knowledge integration perspective that help students develop a more cohesive, coherent, and thoughtful account of scientific phenomena (Linn et al, 2003). WISE is a powerful, research-based online platform for designing, developing, and implementing science inquiry activities. Unique features and benefits of WISE consists library of free, assessments aligned with instruction, interactive visualizations and simulations, embedded prompts for reflection and collaboration, instructional support for diverse learners, teacher feedback and guidance tools, powerful authoring and customization tools library of free, and classroom-tested projects.

Therefore, in this study, the blended learning environment refers to a combination of computer-based interactive lecture demonstration (CBILD) and web-based inquiry science environment (WISE). In this blended learning environment, the learning sequence to facilitate the construction of conceptual understanding and induce the change of misconceptions on the topic state of matter and phase change has been designed and presented in this paper.

3. The Exploration of Students' Misconceptions

In this study, the researchers conducted an exploration to identify students' common misconceptions on chemistry topic of state of matter and phase change. The findings of the exploration provided us as a basis in order to design and create the blended learning environment by combining CBILD and WISE as a novel learning experience for chemistry learning.

3.1 Participants

To explore misconceptions on the topic, fifty twelve-grade students, age ranging from 17 to 18 years in a local public school at the Northeastern region of Thailand were recruited in to the study. They already completed a regular chemistry class and they were taught about the state of matter and phase change before participated in this study.

3.2 Research Instrument and Data Analysis

As aforementioned, 20-item two-tier conceptual test targeting chemistry concepts of state of matter and phase change was used to investigate students' common misconceptions. The conceptual test was adapted from a published research instrument, Particulate Nature of Matter (ParNoMa) (Bridle and Yezierski, 2011). Some of the two-tier conceptual items were constructed by the researchers, and the questions were developed specifically for the concept of phase change of gas to plasma concepts. This topic was chosen after an extensive literature review, which reported that there was no study about student's conceptual understanding in plasma state of matter before. The two-tier conceptual item contains a first tier of multiple choices associated with the main question and a second tier of open-ended reasoning. The two-tier conceptual test was reviewed by three experts for identifying construct and communication validity. The respondents were asked to complete the test within 60 minutes. In this study, the researchers have analyzed and interpreted the respondents' answers into four categories of conception: scientific conception (SC), incomplete conception (IC), misconception (MC), and no conception (NC). The misconceptions identified by this study were used in the development of blended learning environment for chemistry learning of state of matter and phase change.

3.3 Results of Students' Misconceptions on State of Matter and Phase change

The two main topics in this study were state of matter and phase change. These topics consist of many fundamental concepts in chemistry. The concepts were: (i) effect of pressure on the arrangement of particle; (ii) particle movement; and (iii) physical properties of matter, i.e. solid, liquid and gas, and plasma, a new-defined state of matter. After the survey, the researchers found that students hold various patterns of misconceptions in the first topics. In this paper, the researchers reported only students' misconceptions on the effect of pressure on arrangement of particle that most students illustrated especially with many misconceptions. In the second topic, most of students' misconceptions were in place with chemistry phenomena of bubble formed during boiling, phase change of gas to plasma, shape of molecule and the phase change, and the relationship between temperature and phase change. Table 1 displays percentage of students' misconceptions on state of matter and phase change.

Table 1: Students’ misconceptions about state of matter and phase change.

Topic	Concept	Specific types of misconception (MC)	Frequency (%)
State of matter	Effect of pressure on particle arrangement	When pressure of carbon dioxide decreased:	
		▪ No effect on space and size of its molecules	12
		▪ Decreasing of molecule attachment and getting smaller in size	8
		▪ More diffusion of substances	2
		▪ Increasing of space among molecules and getting larger in size	20
		▪ Increasing of numbers of molecule	12
		▪ Decreasing of space among molecules but no effect on size of molecule	24
		Total	78
Phase Change	Bubbles formed during boiling	The bubbles were produced because:	
		▪ Making change by heat	20
		The bubbles come from:	
		▪ Air	8
		▪ Decomposing of water into hydrogen atoms and oxygen atoms	22
		▪ Liquid pressure	2
		Total	52
	Phase change of gas to plasma	When gas reach to high temperature around 10,000 Celsius:	
		▪ Decomposing into hydrogen molecules and oxygen molecules	36
		▪ Decomposing into H^{+} *and* OH^{-}	14
		▪ Decomposing to the end of nothing	14
		▪ Nothing happened	14
		Total	78
	Shape of molecule and phase change	The condensation causes:	
		▪ Larger size of water molecule	13
		▪ Larger size of water molecule because of its combination	12
		The freezing causes:	
		▪ Larger size of water molecule because of its combination	17
		The melting causes:	4
		▪ Larger size of water molecule because of its expansion	
		Total	46
	The relationship between temperature and phase change	When ice was heated:	66
		▪ Changing to be water and stream state respectively	
		Total	66

4. Discussion

This pilot study reports students’ misconceptions on state of matter and phase change and the design of blended learning environment combining CBILD with WISE based on the findings of students' misconceptions. The result showed some patterns of students’ misconceptions about state of matter and phase change. This could be discussed that students hold with various misconceptions in the same content, it shows difference of understanding although they were provided in the same learning context in classroom. Many misconceptions are due to the fact that students do not distinguish between macroscopic and submicroscopic explanations and students have also difficulty linking observable phenomena to molecular level interaction (Chang and Linn, 2013). Almost all students could not explain in term of molecular level, some can answer correctly in multiple choice part of two-tier test but cannot give a reason of their answer or give it by not corresponding. Most chemistry courses require improving students' understanding by adding scientific concepts; however, students often add concepts without integrating them with their other related concepts, which leads to fragmented and fractured understanding (Gilbert and Boulter, 2000). Hence, the researchers can conclude that knowledge integration affects conceptual understanding.

To enhance scientific conceptual understanding the instructor needs to provide instruction for supporting the knowledge integration. The blended combination of CBILD and WISE was designed to address students' misconceptions by supporting knowledge integration and conceptual understanding to make science accessible.

5. The Design of Blended Learning Environment to Address Students' Misconceptions about State of Matter and Phase Change

The blended learning was considered learning environment to address students' misconceptions about state of matter and phase change. Face-to face learning takes place in the conventional science classroom using the computer-based demonstration provided by teacher in front of classroom via projector. The computer-based demonstration may be science experiment or video presentation that dependent on the content and accessible learning materials. For example, the experiment of plasma cannot be demonstrated or performed its experiment in classroom because it is very harmful and instrument also rare in laboratory. During learning period, students took one hour for CBILD in conventional classroom and then two hours for WISE in computer room. Teacher provided student CBILD strategy to make science accessible by empirical evident that is observable phenomena and lead them to distinguish macroscopic explanation. Student were further conducted to investigate unobservable phenomena in the molecular level by using molecular visualization supporting sub-microscopic explanation related to macroscopic and symbolic explanation. When they gain knowledge by related concepts, they will can link three explanations together and result in knowledge integration that becomes conceptual understanding, as shown in Figure 1.

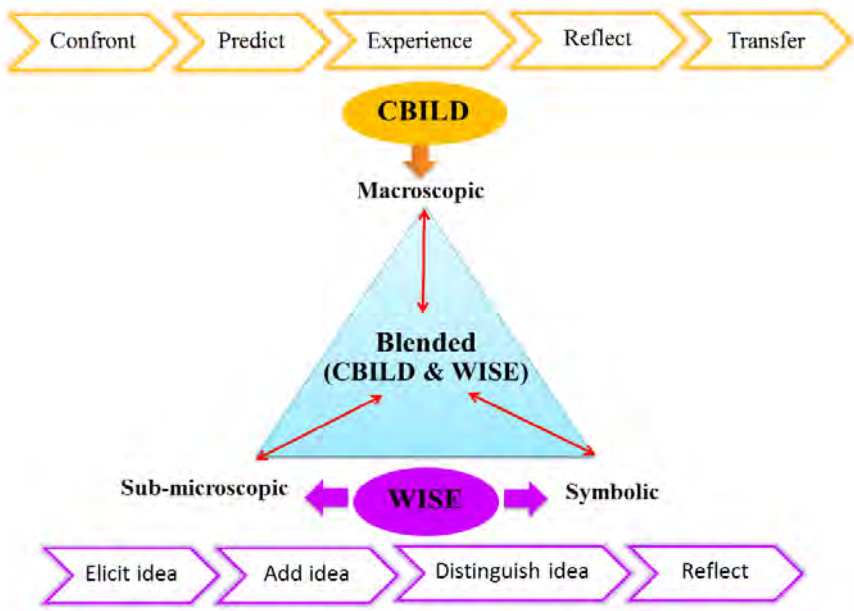


Figure 1. Features of blended learning environment to facilitate students' conceptual understanding in chemistry.

5.1 An Example of CBILD on the Relationship between Temperature and Phase Change

In this part, the researchers would like to present emerging the new technology into science classroom to make science to support students' conceptual understanding and meaningful learning about scientific content from empirical evident. The Vernier Software and Technology is conducted to demonstration of experiment, us sensors such as temperature probe to detect temperature during phase change process of water, and also use a gas pressure sensor to measure vapor pressure of liquid for study the effects of temperature and type of liquid on vapor pressure of liquid. These sensors can connect with computer supported by LabQuest Mini and Logger *Pro* software to acquire real time collection and analysis data, and directly display the result by giving a real-time graph. The researchers use interactive lecture demonstration (ILD) strategy to provide instruction. They help students to visualize the material that is being discussed in class and are frequently described by students as an important component of their understanding of the subject. In this method, the students are asked to predict the outcome of a demonstration and write down their prediction and explanation, and therefore commit to an explicit model. Peer discussion follows, with the students discussing their predictions in small groups– again, they have to address their existing models explicitly (see Table 2).

Table 2: An example of learning process in CBILD classroom.

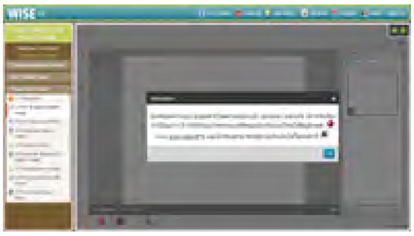
Components	Description of learning process	Examples of learning activity
1. Confront	● Teacher engages student by thinking and giving examples about phase change phenomena in daily life and then provides questions about relationship between temperature and phase change.	
2. Predict	● Teacher conducts students to predict how does temperature change if ice is heated and become water and finally to vapor, respectively, and then water become ice after cooling.	
3. Experience	● Teacher demonstrates change of temperature during phase change by using the Vernier Software and Technology to real time data collection and analysis.	
4. Reflect	● Teacher allows students to compare the results of experiment and their previous prediction and engages them to discuss on the confront phenomena ● Teacher induces student into formative assessment question or statement related the phase change phenomena.	

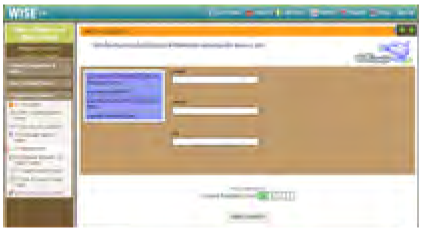
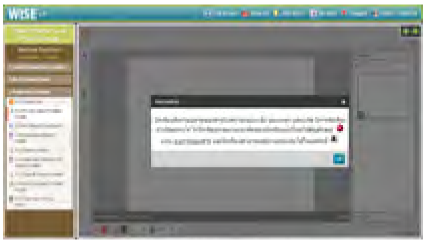
5. Transfer	● Teacher encourages student to explain another physical phenomenon based on the phase change concept.	

5.2 An Example of WISE on State of Matter

Students will use the visualization, resources from The Concord Consortium, PhET, Molecular Workbench and so on, as part of the WISE to explore unobservable scientific phenomena. The inquiry activity helps students to connect new concept to their existing concept. The knowledge integration pattern was implemented in synthesize of WISE. This pattern involves four main learning processes. (i) Eliciting student concept, consistent with the values of predictions; (ii) adding new concept, essential case or new scientific concepts; (iii) supporting students to distinguish among the new and existing concept by constructing arguments and using scientific evidence; and (iv) fostering reflection and self-monitoring to help students consolidate their understanding. Table 3 shows an example of inquiry activity based on knowledge integration (KI) processes. The researchers will assess students' conceptual understanding and knowledge integration by analyzing their answers the prompting questions embedded in WISE activity before and after investigation.

Table 3: Embedded prompts/assessment capturing the knowledge integration processes.

KI processes	Description and Examples of activity	Illustrations of WISE screenshot
1. Eliciting ideas	Using prompting questions to probe prior knowledge and existing ideas Example: ● Could you draw the model of arrangement of particle in each state, liquid solid and gas? ● Explain why did you draw the previous model.	
2. Adding new ideas	Providing necessary information to help students make sense of the topic and connect to existing ideas Example: ● The picture below represents the arrangement of particle in solid, liquid and gas state of matters, you can move the cursor anywhere inside the picture to see the difference of particle movement. ● For this part, you can control the molecular visualization to explore the structure of a solid, liquid and gas at the molecular level.	

3. Distinguishing ideas	Encouraging students to distinguish among ideas and realizing how existing ideas relate to, conflict with, or extend these new ideas Example: ● By the evident from previous page, recall the structure of a solid, liquid and gas at the molecular level in the model you investigated, Could you match the motion of particle with solid, liquid or gas?	
4. Sorting out ideas	Encouraging students to sort out and refine their knowledge based on these evaluations Example: ● Recall the previous questions. After your investigation, could you draw the model of arrangement of particle in each state, liquid solid and gas? And explain why did you draw the previous model.	

6. Conclusion and Future work

The blended learning environment designed based on misconceptions of state of matter and phase change finding in this pilot study will be used in eleventh-grade students to evaluate the effectiveness of the proposed approach. In further study, the combination of CBILD and WISE will be used to enhance eleventh-grade students' conceptual understanding, knowledge integration of the state of matter and phase change, and also scientific motivation. The simultaneous mixed methods strategy of non–equivalent control group design and phenomenological research design will be carried out. The participants will be separated into two group, control and experiment group. Only ILD will be implemented in a control group, on the other hand, ILD combined with WISE will be implemented in an experiment group. Student's knowledge integration will be addressed during learning through WISE, which is designed based on knowledge integration framework, by taking the questions before and after using WISE within class period. To compare difference the outcome between two groups, pre-test, post-test and embedded questions in WISE will be analyzed and interpreted to answer the research question; does blended learning combining CBILD and WISE effect eleventh-grade students' conceptual understanding, knowledge integration about state of matter and phase change, and scientific motivation differently from using only CBILD?.

Acknowledgments

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Bridle, C., & Yezierski, E. (2011). Evidence for the effectiveness of inquiry-based, particulate-level instruction on conceptions of the particulate nature of matter. *Journal of Chemical Education, 89*(2), 192-198.

Burnham, B. (2001). Managing technological change: strategies for college and university leaders. In A. W. (Tony) Bates. *International Review of Research in Open and Distance Learning,1*(2), 1-6.

Chang, H., & Linn, M. C. (2013). Scaffolding learning from molecular visualizations. *Journal of Research in Science Teaching, 50*(7), 858-886.

Gilbert, J. (2005). Visualization in Science Education (1st Edition). Dordrecht: Springer.

Gilbert, J. K. (2006). On the nature of 'context' in chemical education. *International Journal of Science Education*, 28(9), 957–976.

Gilbert, J., & Boulter, C. (2000). Developing Models in Science Education (1st Edition). Dordrecht: Kluwer Academic Publishers.

Gilbert, J., de Jong, O., Justi, R., Treagust, D. & Van Driel, J. (2002). Research and development for the future of chemical education. In J. Gilbert et al., (Eds.). *Chemical Education: Towards Research-based Practice* (pp. 391-408). Dordrecht/Boston: Kluwer Academic Publishers.

Hoic-Bozic, N., Mornar, V., & Boticki, I. (2009). A blended learning approach to course design and implementation. *Education, IEEE Transactions On, 52*(1), 19-30.

Linn, M. C., Clark, D. & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education, 87* (4), 517-538.

Srisawasdi, N., & Kroothkeaw, S. (2014). Supporting students' conceptual development of light refraction by simulation-based open inquiry with dual-situated learning model. *Journal of Computers in Education, 1*(1), 49-79.

Stojanovska, M. , Soptrajanov, B., & Petrusevski, V. (2012). Addressing misconceptions about the particulate nature of matter among secondary-school and high-school students in the republic of Macedonia. *Creative Education, 3*, 619-631.

Thorne, K. (2003). Blended Learning (1st Edition). London: Kogan Page.

Yapici, I., & Akbayin, H. (2012). The effect of blended learning model on high school students' biology achievement and on their attitudes towards the internet. *Turkish Online Journal of Educational Technology*, *11*(2), 228-237.

Yen, J., & Lee, C. (2011). Exploring problem solving patterns and their impact on learning achievement in a blended learning environment. *Computers & Education*, *56*(1), 138-145.

Zimrot, R., & Ashkenazi, G. (2007). Interactive lecture demonstrations: a tool for exploring and enhancing conceptual change. *Chemistry Education Research and Practice*, *8*(2), 197-211.

Sun D., Looi, C-K, & Phua, J.(2013). Design and evaluation of a collaborative inquiry environment to enhance science learning. *Proceedings of the 21st International Conference on Computers in Education*. Indonesia: Asia-Pacific Society for Computers in Education.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Investigating Correlation between Attitude toward Chemistry and Motivation within Educational Digital Game-based learning

NattidaNANTAKAEW[a]&Niwat SRISAWASDI[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
*niwsri@kku.ac.th

Abstract: Educational digital game-based learning, which is one of emerging pedagogies, could promote students' motivation. With the benefit of the educational digital game-based learning, in this study, the educational digital game was designed for learning chemistry. To investigate the effect of attitude toward chemistry lesson on students' motivation when they learn by using educational digital game, the study was conducted on eleventh grade students. The results show that attitude of interest in chemistry lesson makes students understanding and learning chemistry. The importance of chemistry in real-life, chemistry and occupational choice were not related to students' motivation in attention, relevance, confident and satisfaction. Moreover, the students have a positive effect motivation after playing the educational digital game. To this end, this study concludes that the use of educational digital game-based learning could support students' motivation even they have a positive or negative attitude in learning.

Keywords: Educational digital game, attitude toward chemistry, motivation

1. Introduction

Chemistry, which is the one of most important discipline, explains daily life phenomena. Chemistry concept related to other concepts in science such as the biology, physics and materials science. The nature of chemistry is abstract content which need to use imagination for connecting to real life situation. Chemistry requires three different levels of representation which are macroscopic, submicroscopic, and symbolic level. The topic of adhesive force in chemistry uses three levels of representation for explaining the phenomena. This topic related to understanding the basic phenomena in the science curriculum that student is incomprehensible (Eilam, 2004; Leite, Mendoza and Borsese, 2007).There are many factors related to difficulty in learning chemistry. Sirhan (2007) indicated the main factor of the learning difficulty in chemistry, curriculum content, and overload of students' working memory space, motivation, language and communication. Therefore, motivation and attitude to learn is challenging to study, because they are major factors that effect on the success of learning, achievement and willingness of students. The students do not have motivation to understand if they perceive that those contents are difficult. In other words, format of instruction affects students' attitude toward learning. If teacher could provide student-center instruction, students might be willing to learn.

In contemporary education, proper education needs to focus on students. Instruction based on interests, needs and abilities of students will persuade students willing to learn more, motivate them to learn led to get meaningful learning. Students in digital age interest in implementing technology. So, computer and communication technology, which is used widely today, can involve in education for supporting students' learning. It also provides opportunity for the students to understanding of both basic and in-depth content. Moreover, it may make them to understand the complex process and can apply knowledge into everyday life. Learning by using the educational digital game-based learning is a model that combines computer technology into the curriculum, so that the students gain both fun and knowledge (McNamara, Jackson and Graesser, 2010). So, the main challenge for educational research is combining digital game and instruction features in chemistry class. San Chee and Tan (2012) designed and develop an educational game named Legends of Alkhimia. They found that the game effectively fosters learning and supports conceptual understanding of chemistry. Moreover, Papastergiou (2009) found that students who learn via the computer game have more motivational than the non-gaming approach. So, educational computer games can be exploited as effective and motivational learning environments. The previous research has indicated that educational digital game in classroom can support learning and increase students' motivation. However, chemistry class in Thailand lack of combining game into the classroom. Before including game with appropriate chemistry instruction, the attitude toward chemistry should be concerned. Therefore, this study investigates correlation between attitude toward chemistry and motivation in learning via playing game.

2. Literature Review

2.1 Educational Digital Game

The new media and digital technology industries and digital gaming immerse several environments. Digital games consist of dazzling and sophisticated images and sounds, alongside textual communication. Players get engagement which is both pleasurable and challenging. The educational digital game keep players immersed in digital worlds, knowledge, information, and skill development become increasingly accessible outside confines of formal education (Castell, Jenson and Taylor, 2007). Currently, educators employed digital game that insert content of subject matter or information for educational purpose. Several research presented empirical evidences that the educational digital games have positive effect on student learning. It improved not only learning achievement but also learning attitude and motivation to learn (Giannakos, 2013; Pilli and Aksu, 2013; Sung and Hwang, 2013).

2.2 Game-based Learning

In the past, game produce only for entertainment but recently educational researchers have attempted to adapt games for learning which calls educational games or serious game (Sorensen and Meyer 2007; Stone 2008).The games that compose of challenge, control, curiosity and factasy can motivate persistence and enjoyment (Toro-Troconis and Partridge, 2010). The educators have developed games for three goals including: (i) students can learn from playing the game; (ii) the component of game can support learning; and (iii) students have motivation to learn when they learning by playing the game (McNamara, Jackson, & Graesser, 2010). Game-based learning is a kind of constrcutivist-based active learning. Based on the learning research, Watson, Mong and Harris (2011) found that using game in classroom made a shift of teaching from teacher-centered learning environment to student-centered learning environment.

2.3 Attitude and Motivation

The most important students' characteristic associated with successful studies is attitude, motivation, and genuine interest (Berg, 2005). Attitudes and motivation are both important factors for the learning process. Success in learning, positive attitudes to learning and motivation to learn are linked. The two major factors influencing attitudes towards a subject are teacher quality and curriculum quality. They strongly influenced by the perceived curriculum relevance, in the sense that the learner perceives what is taught being related to their lifestyle (Sirhan, 2007). Moreover, Hofstein and Mamlok-Naaman (2011) suggested the three key factors that should be considered for enhancing attitudes and interests are the methods used to present the content, instructional techniques, and gender issues.

3. Purpose

The goal of this study was to investigate correlation of attitude toward chemistry with motivation in learning after playing the educational digital game in the topic of adhesive force and to explore students' motivation after providing the game. Specifically, the research questions were answered:

- How were the influences of attitude toward chemistry on the students' motivation after providing the educational digital game?
- Is it suitable to implement the educational digital game in a Thai school?

4. Methods

4.1 Study Participants

This study recruited 37 students who are studying in eleventh grade, age ranging from 17 to 18 years in a local public school at the northeastern region of Thailand. They are enrolling program that emphasizes using science and technology in the classroom. They also studied about adhesive force which be contained in the topic of properties of liquid in the last semester.

4.2 Instruments

This research used two instruments for determining students' attitude towards the chemistry lesson and motivation in learning via the game. First, the attitude scale developed from Attitude Towards Chemistry Lesson Scale (ATCLS) of Ayyıldız and Tarhan (2013) consisting of 25 items. All items were classified into four scales, including interest in chemistry lessons (six items), understanding and learning chemistry (ten items), importance of chemistry in real-life (five items), and occupational choice related to chemistry (four items). Its Cronbach's alpha reliability coefficient of this instrument was 0.88, implying that it is reliable. Each scale of ATCLS has Cronbach's alpha reliability coefficient from 0.52 to 0.82. The sample item and description of each scale are provided in Table 1. Second, the motivation in learning via the game investigated by using Instructional materials motivational survey (IMMSS) developed from Huang (2011). This instrument consists of 18 items which are divided into attention (eleven items), relevance (four items), confident (three items) and satisfaction (one item). It Cronbach's alpha reliability coefficient was 0.86.Each scale of ATCLS has Cronbach's alpha reliability coefficient from 0.52 to 0.83.

Table 1: Scale descriptions and sample items for the ATCLS questionnaire.

Scale	Description	Sample item
Attention	Extent to which students' response to perceive instructional stimuli provided by the game.	The game has things that stimulated my curiosity.
Relevance	Extent to which student connect their prior learning experience with the game.	There were examples that showed me how the game could be important to some people in the learning setting.
Confidence	Extent to which student has positive expectation after finishing learning activity.	The game had so much information that it was hard to pick out and remember the important points.
Satisfaction	After learning via the game, student accept practice newly acquired knowledge	It felt good to successfully complete the game.

Table 2 shows the sample item and description for four scales. The answers of students in both ATCLS and MLG scale were labeled as strongly agree (5 point), agree (4 points), partly agree (3 point), disagree (2 point) and strongly disagree (1 point) from the positive to the negative.

Table 2: Scale descriptions and sample items for the IMMSS questionnaire.

Scale	Description	Sample item
Interest in chemistry lesson	Extent to which student preferred chemistry learning.	I would like the teaching period of the chemistry lesson more often.
Understanding and learning chemistry	Extent to which student developed themselves and implicated in chemistry easily.	I find using chemical symbols to be easy.
The importance of chemistry in real life	Extent to which student thought chemistry were appropriate to real-life.	I believe that chemical knowledge helps us interpret seriously events in our daily life.
Chemistry and occupational choice	Extent to which student use the information learned in the chemistry classroom for the futuristic work.	My career could be chemist/ chemistry teacher/ chemical engineer.

4.3 Learning Material

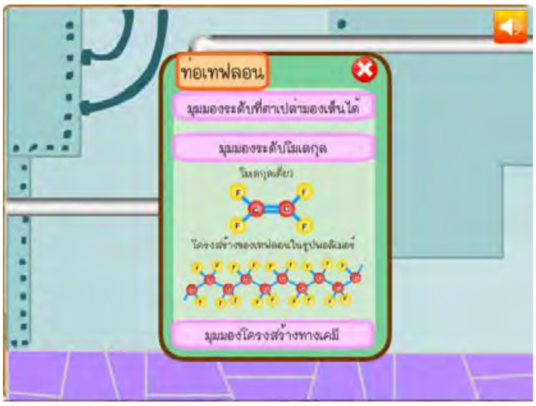
In this study, the design of the game, called "The Pipe", was related to content of adhesive force. The game provides problem situation to students. The game engages students to imagine macroscopic level linked to real life phenomena and also provides information which are three level of representation in chemistry including macroscopic, sub-microscopic and symbolic level. The students were asked to use those information to solve problem. Figure 1 illustrates procedures of learning activity.

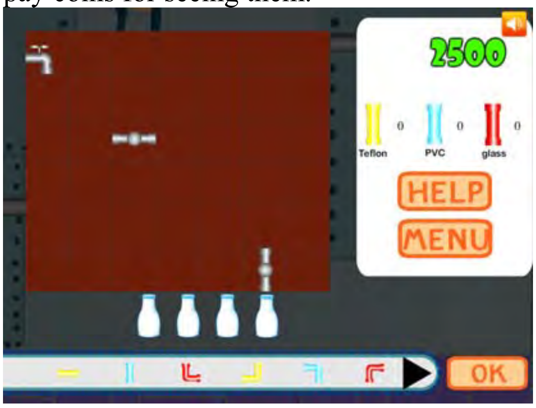
4.4 Data Collection and Analysis

The intervention class consists of 37 students. Before providing The Pipe game, students was surveyed the attitude toward chemistry lesson scale. The Instructional materials motivational survey was provided to the students after they interacted with the developed game. The data from two scales reflected the relation in each variable by using Pearson's correlation in SPSS. The result of Pearson's correlation describe relations of interest in chemistry lesson, understanding and learning chemistry, importance of chemistry in real-life, chemistry and occupational choice in attitude toward chemistry and each scale of motivation including attention, relevance, confident and satisfaction. The influence of attitude toward chemistry on learning motivation via playing the game was investigated.

(A) The game starts with problem situation in the factory. The problem is transferring water through pipes slowly. The player as chemist is asked to choose the proper pipe for the flow of water.

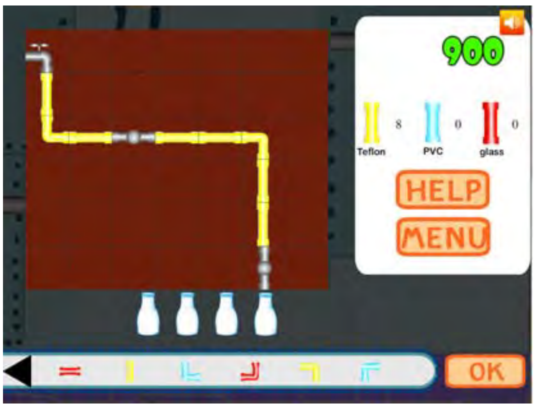
(B) In the game provides scaffolding for decision, the player can see molecular structure of water and each type of pipe. In this part, the player can observe experiment demonstration of the flow of water. However, the player needs to pay coins for seeing them.





(C) This is one sample of the molecules in scaffolding part.

(D) After students click to play a game, it goes in this part in which they pay coins for the various shapes of each pipe for connecting two fix pipes to each other. Students have to concern about budget that they have.



(E) After the player put the proper pipes, they need to click OK bottom for confirming decision.

(F) The last part of the game shows money left, the number of pipes and time used for flowing of water is shown.

Figure 1. Example of chemistry learning activity on properties of liquid.

5. Results and Discussion

5.1 Correlation between Attitude and Motivation

Table 3 shows Pearson's correlation of Interest in Chemistry Lesson (ICL), Understanding and Learning Chemistry (ULC), Importance of Chemistry in Real-life(ICR), Chemistry and Occupational Choice (COC) in ATCLS and Attention (A), Relevance (R), Confident (C) and Satisfaction (S) in IMMSS. Mean and standard deviation are also presented in Table 3.

Table 3: Descriptive and correlation for attitude toward chemistry lesson and motivation.

Scale	ICL	ULC	ICR	COC	A	R	C	S
ICL	1							
ULC	0.57**	1						
ICR	0.21	0.36*	1					
COC	0.66**	0.563*	0.61**	1				
A	0.13	0.10	0.19	0.16	1			
R	0.18	0.04	0.06	0.11	0.78**	1		
C	0.16	0.21	0.11	0.13	0.64**	0.61**	1	
S	0.14	0.05	0.02	-0.06	0.67**	0.57**	0.74**	1
Mean	16.84	29.71	17.05	12.89	36.87	9.90	13.64	3.74
SD	3.19	3.178	2.86	2.70	7.11	1.81	2.60	1.03

$^{**}p< 0.01$

$^{*}p < 0.05$

Regarding Pearson's correlation analysis of each scale from ATCLS, interest in chemistry lesson was positively related to Understanding and Learning Chemistry (ULC), and chemistry and occupational choice. Understanding and Learning Chemistry (ULC) was positively related to importance of chemistry in real-life (ICR) and chemistry and occupational choice (COC). In addition, importance of chemistry in real-life(ICR) was positively related to chemistry and occupational choice(COC).All scale positively related together except interest in chemistry lesson (ICL) scale that do not relate to importance of chemistry in real-life (ICR). These results imply that students attend to chemistry lesson because they understand chemistry concept and want to work in career related chemistry.

From the result of IMMSS, attention (A), relevance (R), confident (C) and satisfaction (S) are linked together. From the findings, it suggests that if students have only one scale of attention (A), relevance (R), confident (C) and satisfaction (S), they have motivation to learn via games. Consider Table 3, the interest in chemistry lesson (ICL), understanding and learning chemistry (ULC), importance of chemistry in real-life (ICR), chemistry and occupational choice (COC) was no related to attention (A), relevance (R), confident (C) and satisfaction (S) when provided the Pipe game to students. So, the educational digital game can use for all students even if they have a negative or positive attitude toward chemistry.

The findings from previous study revealed that the educational digital game improve both attitude toward learning and motivation, because it furnished more attractive of learning environment (Eseryel, Law, Ifenthaler, Ge and Miller, 2014; Sung and Hwang, 2013). But we do not know about the effect of attitude on motivation during interaction with the educational digital game. This study indicated that motivation in learning via game does not depend on attitude toward chemistry. Although students negative or positive attitude toward chemistry, they can learn chemistry by playing game.

5.2 *Students' Responses*

The scale of motivation in IMMSS was summarized in Table 4 which shows the means and standard deviation of each scale. We found that students had high level of motivation in attention, relevance, confident and satisfaction after participating in the developed game.

Table 4: Scale means and summary response.

Scale	Mean (SD)	Description
Attention	36.87(7.11)	High level
Relevance	9.90 (1.81)	High level
Confident	13.64 (2.60)	High level
Satisfaction	3.74 (1.03)	High level

Many previous studies presented that learning via playing game support perspective of cognition and affection. From evidence of achievement test, the game can develop students' performance. Students understand content easily and deeply. For affective domain, the game increase students' motivation. Students preferred to learn by playing game, because they had a positive perspective in learning. They also perceived that the game is useful, easy, interested and enjoyable. All of them can motivate students to learn. We assure that game support learning effectively (Cheng, Huang, and Chen, 2013; Sung and Hwang, 2013).

6. Conclusion and implementation

The result of this study provided a more understand on influences of attitude toward subject matter, for example, chemistry, on student's motivation to learn in a setting of educational game learning environment. The finding indicated that attitude toward chemistry have no correlation on motivation to learn chemistry via game. Thus, we can use the educational digital game for participants who have both a positive and negative effect. Although they like or dislike to learn chemistry lesson, they remain have a positive effect after learning via educational digital game.

Although there are many researches indicated that teaching and learning via game improve students' motivation, we should collect pre- and post-motivation for comparing motivation before and after learning. In an addition, the challenge is how to immerse the digital game into classroom instruction. A previous study by San Chee and Tan (2012) used educational game to support students' inquiry learning process, and they found that the students can effectively inquire to learn science through digital game. Based on the findings of this study, we will design educational digital game about properties of liquid use Student-Associated Game-based Inquiry (SAGOI) approach for improving chemistry learning in quasi-experimental design that include two different-intervention groups of students. One group will provide SAGOI instruction and another acquire traditional instruction. The mixed research methodology combined quantitative method of non-equivalent control group design with qualitative method of phenomenological research design will carry out in future research.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Berg, C. (2005). Factors related to observed attitude change toward learning chemistry among university students. *Chemical Education Research Practice*, *6*(1), 1-18.

Sung, H., & Hwang, G. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, *63*, 43-51.

Cheng, M., Su, T., Huang, W., & Chen, J. (2013). An educational game for learning human immunology: What do students learn and how do they perceive?. *British Journal Of Educational Technology*. *45*(5), 820-833.

De Castell, S., Jenson, J., & Taylor, N. (2007). Digital games for education: When meanings play. *Intermedialities*, *9*, 45-54.

Eilam, B. (2004). Drops of water and of soap solution: Students' constraining mental models of the nature of matter. *Journal Of Research In Science Teaching*, *41*(10), 970-993.

Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An Investigation of the Interrelationships between Motivation, Engagement, and Complex Problem Solving in Game-based Learning. *Educational Technology & Society*, *17*(1), 42–53.

Giannakos, M. (2013). Enjoy and learn with educational games: Examining factors affecting learning performance. *Computers & Education*, *68*, 429-439.

Huang, W. (2011) Evaluating learners’ motivational and cognitive processing in an online game-based learning environment. *Computers in Human Behavior, 27*, 694-704.

Hofstein, A., & Mamlok-Naaman, R. (2011). High-school students’ attitudes toward and interest in learning chemistry. *Revista Educacion Quimica En Linea*, *22*, 90-102.

Leite, L., Mendoza, J. & Borsese, A. (2007). Teachers' and prospective teachers' explanations of liquid-state phenomena: A comparative study involving three European countries. *Journal Of Research In Science Teaching*, *44*(2), 349-374.

Mcnamara, D. S., Jackson, G. T., & Graesser. (2010). Intelligent Tutoring and Games (ItaG). *Gaming for Classroom-Based Learning: Digital Role Playing as a Motivator of Study,*44-57. doi: 10.4018/978-1-61520-713-8.ch003

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers \& Education*, *52*(1), 1-12.

Pilli, O., & Aksu, M. (2013). The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus. *Computers \& Education*, *62*, 62-71.

Sirhan, G. (2007). Learning Difficulties in Chemistry: An Overview. *Journal Of Turkish Science Education*, *4*(2), 2-20.

San Chee, Y., & Tan, K. (2012). Becoming Chemists through Game-based Inquiry Learning: The Case of Legends of Alkhimia. *Electronic Journal Of E-Learning*, *10*(2).

Sorensen, B., & Meyer, B. (2007). Serious Games in language learning and teaching-a theoretical perspective, 559-566.

Stone, R. (2009). Serious games: virtual reality’s second coming?. *Virtual Reality*, *13*(1), 1-2.

Sung, H., & Hwang, G. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, *63*, 43-51.

Toro-Troconis, M. & Partrigde, M. (2010). Designing Game-based learning activities in virtual worlds: Experiences from undergraduate medicine. *Gaming for Classroom-Based Learning: use of gaming in virtual worlds,* 270-289. doi: 10.4018/978-1-61520-713-8.ch016.

Development and preliminary evaluation of a knowledge management-based online teacher community platform for science fair instruction: A cluster analysis

Chiu-Ming, Hu[a], Chao-Shen Cheng[a], Li-Jen Wang[a], Huei-Tse Hou[b], Yi-Chun Kuo[a], Cheng Teng Yao[c] & Ying-Tien Wu[a*]

[a]*Graduate Institute of Network Learning Technology, National Central University, Taiwan*

[b]*Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan*

[c]*Central Taiwan Teaching/Learning Resource Center, Feng Chia University, Taiwan*

*ytwu@cl.ncu.edu.tw

Abstract: Inquiry is the core of modern science education. In science classes, conducting science fair projects is one of the most common inquiry activities. Through conducting these projects, learners have the chance to carry out open inquiry which may help them build deeper understanding of science knowledge, concepts, science skills, and positive attitudes toward science. However, previous research has revealed that teachers encountered many challenges when implementing inquiry-based instruction, such as insufficient time, sources, professional knowledge and experience of inquiry learning. To address the important issue, by integrating community-based knowledge management (KM) tools and personal-based knowledge management (PKM) tools, a KM-based online teacher community platform was developed and preliminarily evaluated in this study. Moreover, cluster analysis was also conducted to categorize the participants' attitude-behavior patterns into different clusters, and the attitudes and behaviors of these different clusters were further analyzed. A total of 103 volunteer Taiwanese elementary school science teachers participated the system evaluations of the Teacher Science Fair Instruction Knowledge Management System (TSFI-KMS) in this study. Their responses on the quantitative questionnaire designed for the system evaluation in this study showed that they expressed satisfactory perceived usefulness and ease of use of the TSFI-KMS. Also, they expressed high willingness to use the TSFI-KMS for professional development regarding inquiry-based instruction. Further analysis indicate that the teachers' perception and usage preference toward TSFI-KMS might be influenced by their academic background and experience of using social media. In this study, suggestions on teaching practices, improvement on the system design, and future research are also discussed.

Keywords: Inquiry; science fair; knowledge management; cluster analysis

1. Introduction

Inquiry is the core of modern science education. In science classes, conducting science fair projects is one of the most common inquiry activities. Through conducting these projects, learners have the chance to carry out open inquiry which may help them build deeper understanding of science knowledge, concepts, science skills, and positive attitudes toward science. However, the literature revealed that many teachers may lack professional knowledge, time, resources, and assistance when conducting science fair instruction (Anderson, 2002), and only a few science teachers know how to guide students to conduct such projects or inquiry activities effectively (Justi & Gilbert, 2002). Therefore, supporting science teachers to develop their pedagogical content knowledge (PCK) (i.e., professional knowledge) regarding science fair instruction, as well as providing them with resources and assistance for science fair instruction, is crucial.

To address the important issue mentioned above, online communities, which have been advocated as a potential tool for teachers to promote their professional development, would be helpful. However, an online platform for teacher communities focusing on science fair instruction is still not yet available. Therefore, this study aimed to develop such a platform. Moreover, the management of professional knowledge created by an online community is always an important issue for both the online community and its members. However, most online teacher community platforms are not formulated based on knowledge management perspectives. By integrating community-based knowledge management (KM) tools (Spector, 2002) and personal-based knowledge management (PKM) tools (Tsui, 2002), a KM-based online teacher community platform was developed and preliminarily evaluated in this study. Moreover, cluster analysis (Hou et al., 2011) was also conducted to categorize the participants' attitude-behavior patterns into different clusters, and the attitudes and behaviors of these different clusters were further analyzed.

2. System development

In this study, the Teacher Science Fair Instruction Knowledge Management System (TSFI-KMS) was developed based on the KM and PKM theoretical framework proposed by Spector (2002) and Tsui (2002). TSFI-KMN is an online knowledge management environment. To meet different teachers' preferences for using the KM system, there are two different portals in TSFI-KMS: the "personal-based portal" and the "community-based portal."

The system framework of TSFI-KMS consists of several main KM and PKM modules and a knowledge base which stores members' profiles, knowledge sharing process and knowledge documents. Examples of detailed functions of the modules are shown in Table 1.

Table 1: Portals, knowledge management tool phase, modules, and functions of TSFI-KMS

Portal	Knowledge management tool phase	Module	Function examples
KM-based (community-based portal)	Communication	Communication module	Asynchronous communication tools, project progress reports
	Coordination	Coordination module	Project calendar, coordination tools
	Collaboration	Collaboration module	Images, videos, and document sharing, science fair project collaboration
PKM-based (personal-based portal)	Indexing and information capturing/ management	Searching and Information capturing module	Information Indexing, Searching results combination, information capturing, information alert, information/documents uploading
	Personal communication management/analysis	Communication management module	E-mail management, communication message analysis
	Learning profile management	Personal profile management module	Learning process tracking, learning profile uploading

3. Methodology (System evaluation)

3.1 Participants

The participants of this study were 103 volunteer Taiwanese elementary school science teachers, including 41 male teachers and 62 female teachers. Regarding the teaching experience distribution of the participant teachers, 23 teachers (22.3%) had less than 5 years, 40 (38.9%) had between 5 and 9 years, 24 (23.3%) had between 10 and 14 years, and 16 (15.5%) had over 15 years. They also had various experience of science fair instruction and of using online social media, such as forums, Facebook, and blogs.

3.2 Evaluation procedures

When the participant teachers first logged into the TSFI-KMS platform, they were given a brief introduction to how to use the system. Then, the participant teachers' background information was collected through an online questionnaire before they started to explore the platform. The participants were asked to explore TSFI-KMS by themselves whenever they had free time during a period of two weeks. After the exploration task, the teachers were asked to evaluate TSFI-KMS by completing online questionnaires.

3.3 Instruments and data collection

In this study, an integrated system evaluation involving both attitude and behavior perspectives was conducted. With online questionnaires, this study collected the data regarding attitude evaluations:

1. *Teachers' acceptance of TSFI-KMS*:
   The 6-point Likert-scale questionnaire developed in Phang et al. (2009) was modified and used in this study. The modified instrument consists of two scales: usefulness (6 items) and usability (7 items). In this study, the overall alpha reliability value of the instrument is 0.93.
2. *Teachers' perceived usefulness of the KM and PKM tools in TSFI-KMS*:
   Two 6-point Likert-scale instruments consisting of 12 and 18 items for assessing teachers' perceived usefulness of the KM and PKM tools in TSFI-KMS were developed in this study. The alpha reliability values of the two instruments are 0.92 and 0.94, respectively.
3. *Teachers' perceived usefulness of TSFI-KMS for improving science fair instruction PCK*:
   A 6-point Likert-scale instrument with 5 items developed in this study was used to evaluate the teachers' perceived usefulness of TSFI-KMS for improving their science fair instruction PCK.

The current study also collected data regarding the *teachers' participation in evaluating TSFI-KMS*. The teachers' participation in evaluating the two portals of TSFI-KMS was evaluated respectively by counting their number of clicks when evaluating the two portals of the system.

4. Major findings and Conclusions

4.1 Teachers' attitudes toward TSFI-KMS and participation in evaluating TSFI-KMS

Table 2 shows that the participating teachers expressed high acceptance of TSFI-KMS (mean=5.03). Also, they highly recognized the usefulness of the KM tools (mean=5.22) and PKM tools (mean=5.13) provided by TSFI-KMS, and agreed with the usefulness of TSFI-KMS for improving their science fair instruction PCK (mean=5.19). Moreover, the teachers had relatively more participation in the community-based portal evaluation (mean=14.74) than in the personal-based portal evaluation (n=6.41), indicating that the teachers in this study might be more oriented towards using the community-based portal of TSFI-KMS.

Table 2: Teachers' attitudes toward TSFI-KMS and participation in platform evaluation

		mean	S.D.
Attitudes toward the platform	Acceptance of TSFI-KMS	5.03	0.59
	Perceived usefulness of the KM tools	5.22	0.49
	Perceived usefulness of the PKM tools	5.13	0.44
	Perceived usefulness for improving PCK	5.19	0.59
Participation in platform evaluation	Community-based portal evaluation	14.74	10.43
	Personal-based portal evaluation	6.41	6.06

4.2 Cluster analysis of the participants' attitude-behavior patterns

A cluster analysis using the attitude and behavior indicators (as revealed in Table 1) was further conducted in this study. We first conducted a hierarchical cluster analysis, then a Ward method's dendrogram to determine the appropriate number of clusters. The teachers' attitude-behavior patterns are divided into three clusters, as shown in Table 3.

According to Table 3, among the three teacher groups, the teachers in group B (about 3.9%) most actively participated in the community-based portal evaluation, while they had the lowest participation in the personal-based portal evaluation. Moreover, these teachers also expressed the most positive attitudes toward TSFI-KMS and the KM and PKM tools, indicating that they strongly recognized the usefulness of the KM and PKM tools. Consequently, they are more likely to continue to use TSFI-KMS, with an orientation towards using the community-based portal. These teachers with higher acceptance of the KM tools may more actively participate in an online community, and thus are likely to play significant roles in the social interactions within the online community. In other words, these teachers are usually the leaders of knowledge sharing or the mediums of knowledge exchange within online communities. Therefore, meeting these teachers' needs in improving their PCK is crucial for promoting the depth of knowledge sharing within TSFI-KMS. To this end, refinements to TSFI-KMS should be made in future research.

Table 3: The result of the cluster analysis

Attitude and behavior indicators	*Clusters*		
	A (n=64, 62.1%)	*B* (n=4, 3.9%)	*C* (n=35, 34%)
Acceptance of TSFI-KMS	5.01	5.23	5.06
Perceived usefulness of the KM tools	5.14	5.42	5.34
Perceived usefulness of the PKM tools	5.10	5.30	5.16
Perceived usefulness for improving PCK	5.12	5.10	5.32
Community-based portal evaluation	9.14	52.25	20.69
Personal-based portal evaluation	5.38	3.75	8.60

The teachers in group C (34%) most keenly participated in the personal-based portal evaluation, and also expressed the most positive perceptions of the usefulness of TSFI-KMS for improving their science fair instruction PCK. It seems that these teachers recognize the usefulness of TSFI-KMS for improving their PCK. Thus, they might also be willing to use TSFI-KMS in the future, with an orientation towards using the personal-based portal. Moreover, compared with most of the teachers in this study (i.e., Cluster A), they perceived higher usefulness and technology acceptance based on TSFI-KMS. However, they had relatively lower participation in the community-based portal, which might restrain the interaction and knowledge sharing with other teachers. Therefore, how to integrate the community-based and personal-based portals more effectively should also be an important issue in the refinement of TSFI-KMS in future studies.

Although the teachers in group A (62.1%), in general, had relatively lower average scores on the attitude indicators compared with the other two groups, their average scores for these indicators were still greater than 5, revealing that they still expressed positive attitudes toward TSFI-KMS. Nevertheless, they had relatively lower average scores for participation in the platform evaluation (i.e., the mean score of the community-based portal evaluation plus the mean score of the personal-based portal evaluation), indicating that they might not be as eager as the teachers in the other two groups to use TSFI-KMS. Besides, compared with the personal-based portal evaluation, the teachers in group A were likely to be more oriented towards using the community-based portal. Since the main objective of TSFI-KMS is to improve teachers' professional development in PCK, suitable scaffolding tools provided by the platform for motivating science teachers' online professional development should be addressed in future research.

In conclusion, the teachers in this study expressed high satisfaction with and acceptance of TSFI-KMS, and scored highly on the usefulness of the KM and PKM tools in TSFI-KMS, and the usefulness of TSFI-KMS for improving science fair instruction PCK. Therefore, the TSFI-KMS platform is suitable for teachers' professional development in science fair instruction. Besides, the cluster analysis in this study revealed that the teachers in the different clusters had their own preferences and attitudes toward the use of KM tools, providing implications for platform refinements, the formation and the management of online teacher communities, and teacher continuing professional development.

Acknowledgment

This study was funded by the National Science Council, Taiwan, ROC, under grant contract number NSC 101-2628-S-008-001-MY3, but the opinions expressed in this article do not reflect the position of the National Science Council.

References

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education, 13*(1), 1-12.

Bell, R., Smetana, L., & Binns, I. (2005). Simplifying inquiry instruction. *The Science Teacher*, 30-33.

Bencze, J. L. & Bowen, G. M., (2009). A national science fair: Exhibiting support for the knowledge economy. *International Journal of Science Education, 31*, 2459-2483.

Collins, A. & Bielaczyc, K. (1997) Dreams of technology-supported learning communities. In *Proceedings of the Sixth International Conference on Computer-Assisted Instruction*. Taipei, Taiwan.

Duncan-Howell, J. (2009). Teachers making connections: Online communities as a source of professional learning. *British Journal of Educational Technology, 41*(2), 324-340.

Hou, H. T., Chang, K. E., & Sung, Y. T. (2011). A longitudinal analysis of the behavioral patterns in teachers using blogs for knowledge interactions. *British Journal of Educational Technology, 42*, 2, e34-36.

Jonassen, D. H., Howland, J., Moore, J. C., & Marra, R. M. (2003). Learning to solve problems with technology - a constructivist perspective. Upper Saddle River, N.J.:

Justi, R. & Gilbert, J.K. (2002).Modeling, teachers' views on the nature of modeling, and implications for the education of modelers. *International Journal of Science Education, 24*(4), 369–387.

Phang C. W., Kankanhalli A., & Sabnerwal R. (2009). Usability and Sociability in Online Communities: A Comparative Study of Knowledge Seek and Contribution. *Journal of the Association for Information Systems*, 10(10), 724-747.

Quintana, C., Reiser, B. J., Davis, A., Krajcik, J., Fretz, E., & Duncan, R. G. et al. (2004). A Scaffolding design framework for software to support science education. *Journal of the Learning Science, 13*(3), 337-386.

Spector, J.M., (2002). Knowledge Management Tools for Instructional Design. *Journal of Educational Technology Research and Development, 50*(4), 37-46.

Tsui, E. (2002), "Technologies for personal and peer-to-peer (P2P) knowledge management", CSC Leading

Edge Forum (LEF) Technology Grant Report, available at www.csc.com/aboutus/lef/mds67_off/uploads/P2P_KM.pdf
Windschitl, M. (2004). Folk theories of “inquiry:” How preservice teachers reproduce the discourse and practices of an atheoretical scientific method. *Journal of Research in Science Teaching, 41*(5), 481–512.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

An Evaluation of Macro-Micro Representation-based Computer Simulation for Physics Learning in Liquid Pressure: Results on Students' Perceptions and Attitude

Jarunya BUYAI[a], Niwat SRISAWASDI[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
** niwsri@kku.ac.th*

Abstract: Computer simulation has been widely used to enhance teaching and learning for last decades and researchers mentioned that the use of computer-simulated experimentation can actively engage and enhance student's meaningful learning in subject contents. As such, a computer simulation for physics learning in liquid pressure has been created with regarding the interplay among macro- and micro representation of physical knowledge. To evaluate the developed simulation, 40 twelve-grade students were recruited to participate with the simulation. A Likert-scale perception and attitude questionnaires were administered to the students before the participation as pre-test. The perception questionnaire was, only, distributed to the student as post-test. The repeated-measures MANOVA results indicated that there was no significant main effect on gender and interaction effect between gender and time (pre-test/post-test), but there were significant main effect on time. That is, there was no effect of gender difference on students' perceptions towards physics learning through computer simulation. Moreover, Pearson's correlation indicated that the computer simulation could be used to promote physics learning experience for all students even if they have a negative or positive attitude toward physics lessons

Keywords: Computer simulation, interactive experience, physics learning, content representation

1. Introduction

Currently, innovative technologies in science teaching and learning is growing continuously. The use of learning technologies such as computer simulation offers students an interactive learning experience and allows students to learn on their own way (Vreman-de Olde et al. 2013). Computer simulations have become increasingly powerful and available to teachers in the past decade (Trundle and Bell, 2010). In community of science and technology education, computer simulation is recognized as a pedagogical tool to support conceptualization of science concepts and facilitate process of scientific inquiry or discovery by visualizing and interacting with dynamics models of natural phenomenon (Ton and Joolingen, 2012; Perkins et al., 2006; Wieman, Perkins and Adams, 2008). These technology offer idealized, dynamic and visual representations of invisible phenomena and experiments which would be dangerous, costly or otherwise not possible in school laboratories (Hennessy, 2007). In addition, researchers found that a learning environment with computer simulation has the advantages that students can systematically explore hypothetical situations, in a realistic environment without stress, in comparison with textbooks and lectures (Rutten, Joolingen, and Veen, 2012).

In physics education, pressure of liquid is a fundamental concept for student's comprehension about fluid mechanics concepts. A few studies have reported that students often encounter learning difficulties and hold unscientific understanding of this concept. There were many misconceptions about the liquid pressure among high school students (Kariotoglou and Psillos, 1993). For example, most of students believed that the shape of the container and amount of the liquid is effective on liquid pressure (Sahina, Çepni, and Ipek, 2010). According to this problem, this study aims to develop a macro-micro representation-based computer simulation for physics learning in liquid pressure to facilitate student's conceptual comprehension of these physics concepts. In this study, a preliminary findings on secondary school students' perceptions was reported.

2. Application of Computer Simulation in Science Education.

Computer simulations are a powerful instructional tools which has been recognized by the community of science education by presenting theoretical, experimentation, models of real-world components, phenomena, processes in science, for students in order to observe, explore, recreate, and receive immediate feedback about phenomena and processes. To support effective science learning activities, computer simulation was used to facilitate their learning difficulties and abstract and complicated content as conceptual learning tools. (Chen et al. ,2011; Colella, 2000; de Jong and Van Joolingen, 1998). In an addition, computer simulation helps learners to understand chemical or biological phenomena which are not able to observe directly (Cook 2006; Wu and Shah 2004). To create awareness through higher-order thinking skill, function of real-time data displays related to a dynamic phenomenon and information on how change parameters synchronously were employed into simulation-based teaching and learning. (de Jong and van Joolingen, 1998; Ronen and Eliahu, 2000).

Even student hold alternative conceptions about science-related phenomena, computer simulation could be used to support a more meaningful learning in science concepts through the process of conceptual change (Srisawasdi and Kroothkeaw, 2014). Computer simulation has significant potential as a supplementary tool for effective conceptual change learning based on the integration of technology and appropriate instructional strategies. There are several educational values that computer simulation adds into science learning activities (Hennessy et al. 2006). As such, successful concepts of simulation-based teaching and learning have been reported by means of discovery learning (de Jong and van Joolingen, 1998; de Jong, Linn, and Zacharia, 2013) and inquiry-based learning (Perkins et al., 2006; Srisawasdi and Kroothkeaw, 2014; Wieman, Perkins, and Adams, 2008). With the importance, computer simulations are effectively linked to pedagogy as well. (Flick and Bell 2000). Following these class types, student can discover the principles, rules, and characteristics of scientific phenomena through change variable values and observe effects to form scientific conclusions in computer simulation (Veemans et al., 2006). Therefore, the use of computer simulation with theses pedagogical approaches could be as instructional line for teaching and learning in school science.

3. Purpose

The goals of this study were to investigate students' perceptions towards computer-simulated physics experimentation and correlation between attitude toward physics and physics motivation after interacting with the computer-simulated physics experimentation. Specifically, the following questions were answered:

(1) Do the students engaged in computer-simulated physics experimentation perform significantly better by perceive learning, flow of learning experience, enjoyment, perceive ease of use, perceive of satisfaction, and perceive of usefulness?

(2) How were the influence of physics attitude on students' perceptions after interacting with computer-simulated physics experimentation?

4. Method

4.1 Participants

A total of 40 twelfth grade students (female = 25, male = 15), age ranging from 17 to 18 years, in a local public school at the northeastern region of Thailand participated in this study. They were attending a physics course for basic education level. Regarding to prior learning experience, they have no experience yet using computer-simulated experiment in physics. This implied that they are heterogeneous on perception towards computer-simulated physics experimentation before interacting with the present experimental study.

4.2 Instructional Materials

The design of computer-simulated experimentation in physics of liquid pressure provides a rich context of representations where macro- and micro-scale representations were employed coordinately to visualize how the physical phenomena works. In additions, the computer-simulated physics experiment on liquid pressure has been produced by examining the attributes of the physics concept to provide information in which essential mental sets are needed to construct a scientific view of the concepts. According to this step, a concept map associated to learning liquid pressure simulation was constructed in hierarchical order of attributes. After the attributes and essential underlying concepts were identified by the researchers and two experts. As such, it involved three essential parameters (i.e. depth of water, shape of container, and density of liquid) which related to liquid pressure phenomenon, as shows in Figure 1.

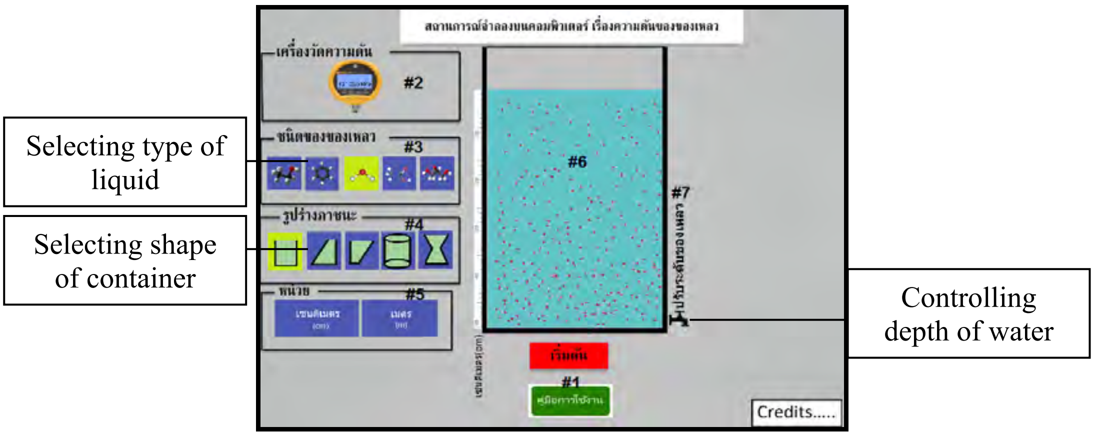


Figure 1. Example of screen interface of computer-simulated physics experiment on liquid pressure

For the parameter of depth of water, the computer-simulated physics experiment on liquid pressure prepares to build student's conceptual knowledge on relationship between the depth of water and pressure, and serves physical understanding on what would happen to pressure if height of water is changed. In this part, student could interact with controlling levels of liquid and then measure its pressure by using a pressure gauge. Moreover, they could also see whether the liquid molecules behave relating to the depth. Figure 2 displays an example of interactive features in the computer-simulated physics experiment on relationship between liquid pressure and depth.

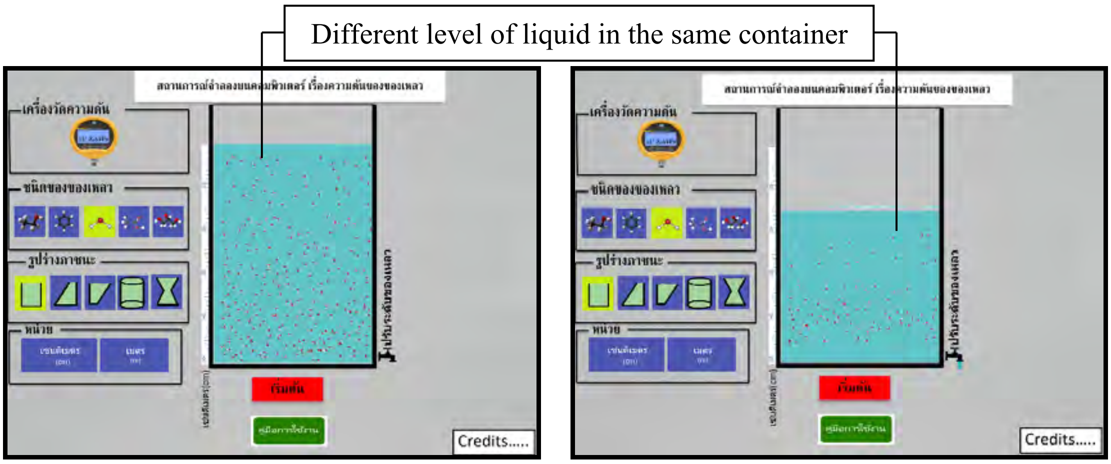


Figure 2. Interactive feature of different levels of liquid in the liquid pressure simulation on relationship between liquid pressure and depth

To facilitate student's conceptual knowledge on impact of liquid density on its pressure, the computer-simulated physics experiment on liquid pressure provides opportunity to select different types of liquid (i.e. ethanol, benzene, water, acetic acid, and glycerin) for investigating the concept. In this part, student could interact with selecting types of liquid and then measure its pressure. Moreover, they could also see whether the liquid molecules behave relating to the type of liquid. Figure 3 displays an example of interactive features in the computer-simulated physics experiment on relationship between density of liquid and its pressure.

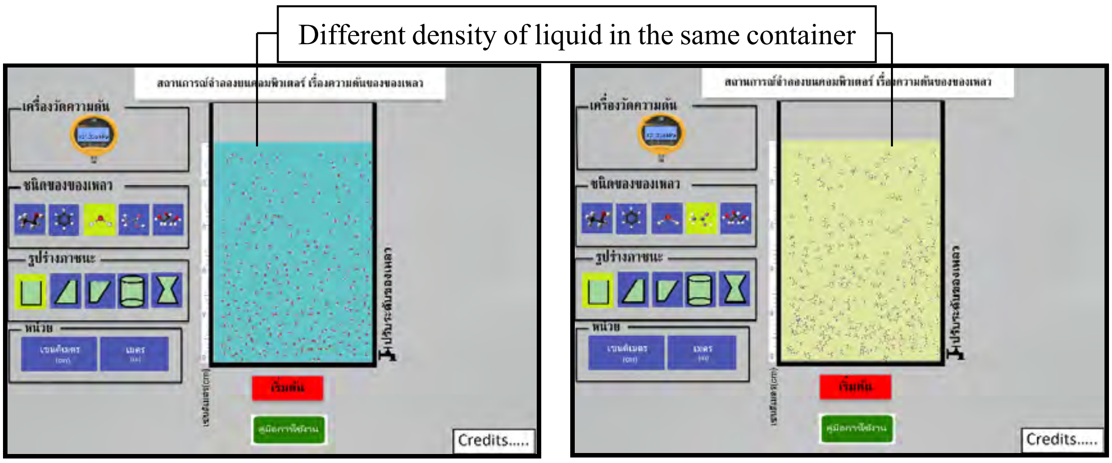


Figure 3. Interactive feature of different densities of liquid in the liquid pressure simulation for measuring the liquid pressure

Visualizing an effect of shape of container on liquid pressure was an aim for creating this computer-simulated experimentation. Figure 4 displays an example of interactive features in the computer-simulated physics experiment on relationship between shape of container and liquid pressure. Student could interact with the experiment by selecting a shape of container and then measure its pressure by using a pressure gauge. This part prepares to build student's conceptual knowledge on what would happen to pressure if shape of container is changed.

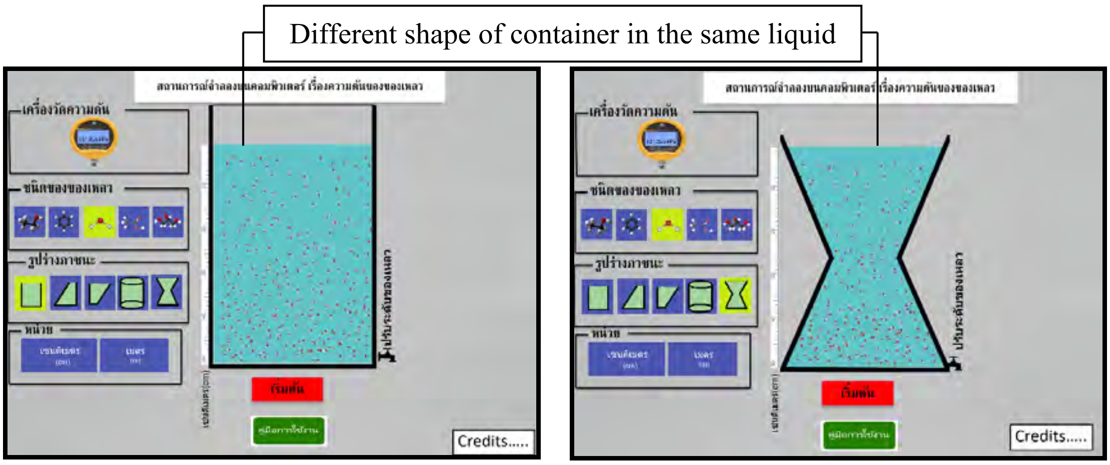


Figure 4. Interactive feature of different shapes of container in the liquid pressure simulation

4.3 Instrument

A 21-item Likert-scale questionnaire was developed to use in this study for examining students' perceptions towards the computer-simulated physics experiment on liquid pressure on six subscales: flow, enjoyment, perceived learning, perceive ease of use, perceive of usefulness, perceive of satisfaction. All of these 5-point Likert scale items obtained from (Cheng, 2014) and Barzilai and Blau (2014). From the English version, an identical version in Thai was constructed, and one expert was recruited to identify communication validity of the items. The respondents were required to consider each possible reason for computer simulation and rate how much the respondent agree with into five scale (1-strongly disagree; 2-disagree; 3-neutral; 4-agree; 5-strongly agree). The reliability for the overall questionnaire was 0.88 and for each subscale was presented in Table 1.

Table 1: Example items of perception questionnaire for each construct.

Dimension	Sample items	α
Perceive learning	▪ The simulation added to my knowledge. ▪ I learned new things from the simulation. ▪ The simulation will help me remember the things I learned.	0.803
Flow	▪ I lost track of time when I played. ▪ I really got into the simulation. ▪ I was very involved in the simulation.	0.822
Enjoyment	▪ I enjoyed the simulation. ▪ I had fun playing the game. ▪ Playing the simulation was pleasant.	0.745
Perceive ease of use	▪ It is easy for me to learn how to use simulation. ▪ The user interface of simulation is easy to use. ▪ I can easily accomplish what I need to do in simulation.	0.737
Perceive of usefulness	▪ Simulation can help me learn more effectively. ▪ Simulation can improve my course performance. ▪ It is useful to study the course content with simulation.	0.842
Perceive of satisfaction	▪ I feel comfortable to use simulation. ▪ I enjoy the experience of using simulation. ▪ I am willing to continue using simulation for learning in other courses	0.774

In an addition, a 25-item Likert-scale questionnaire obtained from Ayyıldız and Tarhan (2013) was applied to measure attitude towards physics lessons. All items were classified into four scales, including interest in chemistry lessons (6 items), understanding and learning physics (10 items), importance of physics in real-life (5 items), and occupational choice related to physics (4 items). Its cronbach's alpha reliability coefficient of this instrument was 0.88, implying that it is reliable. Each scale has cronbach's alpha reliability coefficient from 0.52 to 0.82. Table 2 shows the sample item and description for four scales.

Table 2: Scale descriptions and sample items for the physics attitude questionnaire

Scale	Description	Sample item
Interest in physics lesson	Extent to which student preferred physics learning.	I would like the teaching period of the physics lesson more often.
Understanding and learning physics	Extent to which student developed themselves and implicated in physics easily.	I find using symbols in physics to be easy.
The importance of physics in real life	Extent to which student thought physics were appropriate to real-life.	I believe that physics knowledge helps us interpret seriously events in our daily life.
physics and occupational choice	Extent to which student use the information learned in the physics classroom for the futuristic work.	My career could be physicist/ physics teacher/ engineer.

4.4 Data Collection and Analysis

The participants were asked to complete the perception questionnaire, to measure their pre-perceptions towards the computer-simulated physics experiment on liquid pressure, and the physics attitude questionnaire, to measure attitude towards physics lessons, for 15 minutes. After completing the instrument, they were exposed to interact dependently with the experiment for 25 minutes. After completing the experiment, the students' post-perceptions were examined by the same questionnaire for 10 minutes. The statistical data techniques selected for analyzing students' science motivation was repeated-measures MANOVA in SPSS to compare effect of intervention considering gender (female/male) and time (pre-test/post-test). In an addition, Pearson's correlation was used to investigate relationship between physics attitude (interest in physics lesson, understanding and learning physics, importance of physics in real-life, physics and occupational choice) and their perceptions (flow, enjoyment, perceived learning, perceive ease of use, perceive of usefulness, perceive of satisfaction).

5. Results

The MANOVA indicated no significant main effect for gender (Wilks' lambda=0.875, F (6, 33) =0.783, p=0.589, partial η^2 = 0.125). There was no significant difference on perceived towards learning science through computer simulation between females and males. The univariate results on gender revealed none of the six subscales on perceived towards learning science through computer simulation reached a statistical significance between females and males. That is both females and males performed indifferently with regard to perceive learning (PL), flow (Fl), enjoyment (Ej), perceive ease of use (PE), perceive of usefulness (PU), perceive of satisfaction (PS). Also, there was no significant interaction effect between gender and time (Wilks' lambda=0.915, F (6, 33) =0.795, p=0.085, partial η^2 = 0.090). This means that the learning module has similar effects on perceived towards learning science through computer simulation for females and males. However, there was a significant main effect for time (Wilks' lambda=0.475, F (5, 40) =5.201, p<0.005, partial η^2 = 0.525). The multivariate eta squared, η^2, indicates the effect size, and a value of 0.525 means that about 52.5% of multivariate variance of the dependable variables was associated with time. The results of the univariate test for females and males students are summarized in Table 3.

Table 3: The students' perceptions towards computer-simulated physics experiment by time and univariate MANOVA.

Dimension	Time		F	Sig.	η^2
	Pre-test	Post-test			
Perceive learning (PL)	10.82 (3.16)	13.85(2.09)	21.170	0.000^{***}	0.385
Flow (Fl)	12.86 (3.50)	15.62 (2.549)	12.322	0.001^{**}	0.245
Enjoyment (Ej)	8.54 (2.47)	9.82 (2.11)	4.647	0.038^{*}	0.109
Perceive ease of use (PE)	8.06 (2.23)	9.56 (1.93)	10.263	0.030^{*}	0.213
Perceive of usefulness (PU)	8.69 (2.40)	9.94 (1.71)	6.230	0.017^{*}	0.141
Perceive of satisfaction (PS)	8.74 (2.73)	10.79 (1.96)	14.158	0.000^{***}	0.227

Note. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

As can be seen in Table 3, The univariate MANOVA on the six dimension scores of perceived towards learning science through computer simulation were significant differences across time, from pre-test to post-test. The univariate results revealed a significant effect on PL ($F_{1,38} = 21.170$, $p < 0.05$, partial $\eta^2 = 0.385$), Fl ($F_{1,38} = 12.333$, $p < 0.05$, partial $\eta^2 < 0.245$), Ej($F_{1,38} = 4.647$, $p < 0.05$, partial $\eta^2 = 0.109$), PE ($F_{1,38} = 10.263$, $p < 0.05$, partial $\eta^2 = 0.213$), PU ($F_{1,38} = 9.940$, $p < 0.05$, partial $\eta^2 = 0.141$) and PS ($F_{1,38} = 14.158$, $p < 0.05$, partial $\eta^2 = 0.227$). According to aforementioned results, the overall result suggested that the increase of perceived towards learning science through computer simulation regarding perceive learning, flow, enjoyment, perceive ease of use, perceive of usefulness and perceive of satisfaction from the pre-test to post-test was homogeneous both females and males after participating with the computer simulation. That is, there was no effect of gender difference on perceived towards learning physics, through computer simulation learning.

For examining correlation between their physics attitudes and perceptions towards the experiment, Table 4 shows Pearson's correlation of interest in physics lesson (IPL), understanding and learning physics (ULP), importance of chemistry in real-life (IPR), physics and occupational choice (POC), and perceive learning (PL), flow (Fl), enjoyment (Ej), perceive ease of use (PE), perceive of usefulness (PU), perceive of satisfaction (PS). Mean and standard deviation are also presented in Table 3.

Regarding Pearson's correlation analysis of each variable, the result showed that there were no significant correlation between students' physics attitudes and perceptions towards the computer-simulated physics experiment. Thus, the result implied that the experiment could be used for all students even if they have a negative or positive attitude toward physics lessons.

Table 4: Descriptive and correlation for Attitude toward physics lesson and perceptions.

Variable	IPL	ULP	IPR	POC	PL	Fl	Ej	PE	PU	PS
IPL	-									
ULP	0.56^{**}	-								
IPR	0.02	0.26	-							
POC	0.25	0.30	0.09	-						
PL	-0.09	-0.22	0.07	-0.27	-					
FL	0.10	-0.07	0.09	-0.06	0.55^{**}	-				
EJ	-0.03	-0.12	-0.00	-0.16	0.64^{**}	0.79^{**}	-			
PE	0.01	-0.23	-0.10	-0.10	0.26	0.53^{**}	0.33^{*}	-		
PU	0.05	-0.16	0.23	-0.25	0.48^{**}	0.70^{**}	0.65^{**}	0.45^{**}	-	
PS	-0.12	-0.10	0.14	-0.06	0.53^{**}	0.52^{**}	0.59^{**}	0.30	0.70^{**}	-
Mean	15.38	28.03	15.65	10.43	13.58	15.48	9.80	9.58	9.93	10.73
SD	2.77	4.71	3.17	2.51	2.09	2.54	2.11	1.93	1.71	1.96
Note. $^{*}p < 0.05$, $^{**}p < 0.01$										

6. Discussions

This study reports an impact of an innovative technological tool for physics instruction, a computer-simulated experimentation on liquid pressure, for promoting students' perceptions and attitude towards physics lessons. This tool was developed by emphasizing the interplay between macro- and micro-scale representations to induce cognitive process on construction of conceptual understanding and mental model. The result shows an increasing of students' perceptions scores from pre-test to post-test reached a statistically significant effect across the time, and gender different had no significantly effect on the increasing of perceptions. This finding could be argued that in physics concept made progress throughout their experiencing with the computer-simulated experimentation on liquid pressure. This indicates that the experimentation successfully helped students getting better perceptions for physics learning of liquid fluid. The result is consistent with the research findings that students performed better achievements with learning from computer simulation (Tuan Soha, et. al., 2010). Based on attitude theory, Zimbardo and Leippe (1991) stated that attitudes can be either negative or positive and these attitudes can affect both perceptions and behavior. However, this study indicated contradict with the theory that there was no effect of students' prior physics attitudes on their perceptions towards computer-simulated experimentation.

7. Conclusion

A macro-micro representation-based computer simulation for physics learning in liquid pressure has been developed to promote students' conceptual comprehension and motivate them to learn physics more meaningfully. To preliminary evaluate its effectiveness for physics teaching and learning, this study employed the physics computer simulation to students and findings revealed that (i) gender difference has no effect on students' perceptions towards learning of physics through macro-micro representation-based computer simulation, and (ii) their attitudes towards physics lessons have no effect on the use of the computer simulation for physics learning in liquid pressure. As such, it is clear that both females and males increased their perceptions on perceive learning, flow, enjoyment, perceive ease of use, perceive of usefulness, and perceive of satisfaction after interacting with the simulation. Moreover, their attitudes towards physics lessons cannot intervene perceptions when they learn from the computer simulation. These findings could be used as a basis to develop an alternative computer simulation for promoting physics instruction by emphasizing an interplay of macro- and micro-scale representations.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Adesina, A.O., & Akinbobola, A.O. (2005). The attitude of students towards part-time degree programmed of the faculty of education, Obafemi Awolowo University, Ile-Ife. *Journal of Research of Education*, 2(1), 1-4.

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, 70, 65-79.

Bell, R. L., & Trundle, K. C. (2008). The use of a computer simulation to promote scientific conceptions of moon phases. *Journal of Research in Science Teaching*, 45(3), 346–372.

Bozkurt, E., & Ilik, A., (2010).The effect of computer simulations over students' beliefs on physics and physics success. *Procedia Social and Behavioral Sciences*, 2 (2010), 4587–4591.

Caspi, A., & Blau, I. (2011). Collaboration and psychological ownership: how does the tension between the two influence perceived learning. *Social Psychology of Education*, 14(2), 283-298.

Cheng, G. (2014). Exploring students' learning styles in relation to their acceptance and attitudes towards using Second Life in education: A case study in Hong Kong, *Computers & Education*, 70 (2014), 105-115

Colella, V. (2000). Participatory simulation: Building collaborative understanding through immersive dynamic modeling. *Journal of the Learning Science*s, 9(4), 471-500.

Chen, Y. L., Hong, Y. R., Sung, Y. T., & Chang, K. E. (2011). Efficacy of simulation-based learning of electronics using visualization and manipulation. *Educational Technology & Society*, 14(2), 269–277.

Chen, Y.-L., Pan, P.-R., Sung, Y.-T., & Chang, K.-E. (2013). Correcting misconceptions on electronics: Effects of a simulation-based learning environment backed by a conceptual change model. *Educational Technology & Society*, 16(2), 212–227.

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179–202.

Hendrickson, A. B. (1997). Predicting student success with the learning and study strategies 14. inventory (LASSI). *Unpublished Master's Thesis, Iowa State University*.

Hennessy, S., Deaney, R., & Ruthven, K. (2006). Situated expertise in integrating use of multimedia simulation into secondary science teaching. *International Journal of Science Education,* 28(7),701–732.

Jaakkola, T., & Nurmi, S. (2007). Fostering elementary school students'understanding of simple electricity by combining simulation and laboratory activities. *Journal of Computer Assisted Learning*, 24(2008), 271–283.

Kariotoglou, P., & Psillos, D. (1993). Pupils' pressure models and their implications for instruction. *Research in Science & Technological Education*, 11 (1), 95.

Redish, E., Saul, J., & Steinberg, R. (1998). Student Expectations in Introductory Physics. *American Journal of Physics*, 66(3), 212-224.

Ronen, M., & Eliahu, M. (2000). Simulation-a bridge between theory and reality: The case of electric circuits. *Journal of Computer Assisted learning*, 16(1), 14–26.

Sahina, Ç., Çepni, H., & Ipek, S., (2010). Computer supported conceptual change text: Fluid pressure. *Procedia Social and Behavioral Sciences*, 2 (2010), 922–927

Srisawasdi, N., & Kroothkeaw, S. (2014). Supporting students' conceptual development of light refraction by simulation-based open inquiry with dual-situated learning model. *Journal of Computers in Education,* 1(1), pp.49-79.

Tuan Soha, T., Mohamad Arsada, N., & Osmana, K. (2010). The Relationship of 21st Century Skills on Students' Attitude and Perception towards Physics, *Procedia Social and Behavioral Sciences* 7(C) (2010) 546–554.

Vreman-de Olde, C., de Jong, T., & Gijlers, H. (2013). Learning by designing instruction in the context of simulation-based inquiry learning. *Educational Technology & Society*, 16(4), 47–58.

Zacharia, Z., & Anderson, O. (2003). The effects of an interactive computer-based simulation prior to performing a laboratory inquiry-based experiment on students' conceptual understanding of physics. *American Journal of Physics*, 71, 618.

Zimbardo, P.G., & Leippe, M.R. (1991). The psychology of attitude change and social influence. *Philadelphia, Temple University Press*.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Promoting Students' Physics Motivation by Blended Combination of Physical and Virtual Laboratory Environment: A Result on Different Levels of Inquiry

Chakkrapan PIRAKSA[a], Niwat SRISAWASDI[b*]
[a]Science Education Program, Faculty of Education, Khon Kaen University, Thailand
[b]Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Thailand
*niwsri@kku.ac.th

Abstract: In science and technology education community, technology-based pedagogy in science learning has been mentioned its effectiveness for facilitating scientific inquiry in school science. As such, this study investigated an effect of inquiry-based learning process into a blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory on secondary school students' physics motivation. Study participants were 66 eleventh-grade students of diverse learning abilities in a public school in Northeastern region of Thailand. They were measured intrinsic motivation (IM), career motivation (CM), self-determination (SDT), self-efficacy (SEC), and grade motivation (GM) in physics learning by using a 25-item questionnaire both before and after participating the intervention. To evaluate the intervention, repeated-measures MANOVA was performed to examine its effects regarding type of inquiry (open- and guided inquiry) and time (pre- and post-test). The results showed that students' physics motivation for pre- and post- test were significantly different and their motivation were improved after participating with blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory for both types of inquiry. This evidence indicated that inquiry-based physics learning with the blended laboratory environment (physical and virtual lab) influenced students' progression of physics motivation. As such, blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory could be considered as a pedagogic technology-based laboratory environment for teaching and learning of science by inquiry.

Keywords: Sound wave interference, Microcomputer Based Laboratory (MBL), Computer simulation, and Motivation

1. Introduction

Several students have experienced difficulty in physics course due to misconceptions in many physics contents (Singh, Singh, Kumari, and Kumar, 2011), especially properties of sound wave. By the nature, properties of sound wave involve reflection, interference, reflection, diffraction, and also propagation of sound wave. Meanwhile, sound wave interference is invisible, complicated, and boring (Hola, 2007). Normally, teachers' teaching in a regular classroom can encourage students to succeed in school and unable to motivation to learn as interact in complex ways to lead learning (Schunk, 2005). Teacher's teaching approaches also depress motivation of students and decrease students' learning performance in physics. Recently, most researchers have been concentrated on the scientific conceptions. The issues of motivation to learn physics has been attended in respect of science achievement and scientific conceptions (Hamzah and Mdzain, 2010). However, it's difficult to achieve this ultimate goal because many learners are treated with a lack of motivation to learn science (Glynn, Brickman, Armstrong, and Taasoobshirazi, 2011).

In the recent year, students are educating a shift from passive sitting and listening to a more dynamic learning experience. Several active-teaching methods are introduced to solve those problems. Inquiry-based approach is a verity of instructional methods to apply with high school students. Science educators also confirmed methods of inquiry are more effective and valued to both teaching and learning (Guzey and Roehrig, 2009; Sadeh and Zion, 2011). The use of inquiry-based approaches is strongly subscribed to teaching and learning of science (Minstrell and VanZee, 2000), student-centered, providing students with opportunities to formulate and conduct their own scientific investigations (Singer, Marx, and Krajcik, 2000). Scientific inquiry tasks play an important role for students in the process of conceiving scientific problems and questions, formulating hypotheses, designing experiments, gathering and analyzing data, and drawing conclusions (Hofstein, Navon, Kipnis, and Naaman, 2005).The researchers also revealed that the cookbook-laboratory activities do not promote the development of students' higher order thinking skills. On the other hand, in inquiry-based laboratory students are more associated with, and usually have positive attitudes regarding their laboratory experience (Abd-El-Khalick and Akerson, 2004)

On top of that, this learning process is a wide range of efficient technological environments and applications including animations, simulations and modeling tools, microcomputer-based laboratories (MBL), intelligent tutoring systems, web resources and environments, spreadsheets, scientific databases, for instance, in the science education community of practice. Using as tools, MBL and computer simulation are subject to introduce students' cognitive development and result in students' positive response (Hola, 2007) because they facilitated to understand the scientific conceptions that confront them (Mulder, Lazonder, and Jong, 2011; Russell, Lucas, and McRobbie, 2003). It was not until third decade ago, MBL was reported to understand and integrate learners the sophisticated topics of physics including temperature probe, heat energy (Russell et al., 2003), and properties of sound wave (Gunhaart and Srisawasdi, 2012). Furthermore, the capacity of MBL enable learners to immediately transform data from each experiment into graph, the most powerful form of presentation. Learners will be engaged a construct and had conceptualized change after all. In the interval, computer simulations are examined to be the most technically complicated option for offering various benefits for the teaching and learning of science (Blake and Scanlon, 2007). For this reason, a well-designed computer simulation used within MBL as educational technology and inquiry learning as instructional process can be very effective in promoting meaningful learning in scientific concepts (Bell and Trundle, 2008)

This research utilized both tools to engage learners a meaningful learning of sound wave interference. Conceptualized change was expected to achieve by measuring five components involving Intrinsic motivation (IM), Career motivation (CM), Self-determination (SDT), Self-efficacy (SEC), and Grade motivation (GM). Inquiry types were examined as dependent variables for motivation.

2. Literature Review

2.1 Inquiry

In Thailand, instructors popularly recommend to use inquiry-based learning as one of many instructional strategies to implement in science education. Theoretically, inquiry-based learning is defined as the creation of a classroom where students are engaged in essentially open-ended, student-centered, hands-on activities involving asking questions about the world around them, gathering evidence, and providing explanations (Colburn, 2000). It is restricted that solely activities, e.g. building a model of an atom, cannot be referred to inquiry-based learning if they are conducted in the absence of research questions as a part of inquiry process. The inquiry-learning literature tends to be more closely associated with the acquisition of science process skills or the scientific thinking and reasoning patterns that scientists use to construct (Bunterm et al., 2014)

Researchers typically discriminate between different levels of inquiry-based learning depending on the amount of specific instructions given to students. (Buck, Bretz, and Towns, 2008). Buck (2008) proposed a fifth-level model. At the first level (Confirmation), the problem, procedure, analysis, and correct interpretations of the data are all provided to the students. At the second level (Structured inquiry), the laboratory manual provides the problem, procedures, and analysis by which students can discover relationships or reach conclusions that are not already known from the manual. At the third level (Guided inquiry), the laboratory manual provides the problem and procedures, but the methods of analysis, communication, and conclusions are for the student to design. At the fourth level (Open inquiry), the problem and background are provided, but the procedures/design/methodology are for the student to design, as are the analysis and conclusions. At the highest level (Authentic inquiry), the problem, procedures/design, analysis, communication, and conclusions are for the student to design.

This investigation compared two kinds of inquiry-based processes: guided versus open inquiry. Learners will be engaged to have a construct with providing the problem and procedures for guided inquiry but providing just the problem for open inquiry. Sadeh and Zion (2011) examined the influence of these two different inquiry learning processes on the attitudes of Israeli high school Biology learners toward their inquiry project. It is found that there were significant differences between the two groups. Learners were more satisfied and felt they gained benefits from implementing the project to a greater extent for open inquiry. On the other hand, they conducted more documentation for Guided inquiry. Bunterm et al. (2014) examined the effects of guided against structured inquiry on secondary students' learning of science with three schools in north-eastern Thailand. In comparison, students in the guided-inquiry condition showed greater improvement in both science content knowledge and science process skills. Any moment now, researchers have been subject inquiry-based learning using MBL and computer simulation as tools for conceptual understanding and change in physical science to middle and high school students, pre-service teachers to enhance learners' meaningful learning in the area of scientific concepts (Gunhaart and Srisawasdi, 2012).

2.2 Microcomputer based Laboratory(MBL) and Computer simulation for science instruction

At this moment in time, computational technologies are increasing attention among science educators because of their potentials to support new variety of science classroom (Srisawasdi, 2008). MBL and computer simulation are taking participant their own prominent rules in thinking and reflecting learning input for an instructor and a conceptual construct respectively. Serving as alternative software for teaching assistant tool, MBL is widely used for instructional activities to stimulate students' curiosity as a learning motivator, develop students' scientific skills, foster collaborative network, understand in scientific concepts, and establish students' cognitive construct (Srisawasdi & Kroothkaew, 2014).

Additionally, Redish, Saul, and Steinberg (1997) investigated that active-engagement tutorials using MBL equipment were replaced for traditional problem-solving recitations in introductory calculus-based mechanics classes for engineering students at the University of Maryland. Two specific tutorials, on the concept of instantaneous velocity and Newton's third law, were performed with eleven lecture classes taught by six different teachers with and without tutorials. Classroom achievement tests were probed by using standard multiple choice questions and a free-response final exam question. The result shows that the MBL tutorials originated in a remarkable improvement compared to the traditional recitations. Russell et al. (2003) designed and provided experiments with grade 11 physics classes of 29 students. The research distinguished the learners and illustrated the patterns of interactions in the MBL. Analysis of students' discourse and actions identified kinematics in multiple ways. The finding is that MBL activities likely catalyzed students' construction of understanding. Students were able to design the research questions, predictions, designing experiments, collecting data, and drawing conclusions.

In addition, Gunhaart and Srisawasdi (2012) used MBL as a tool for scientific thinking and computer simulation as a cognitive tool for conceptual learning to improve the construction of physics conceptual understanding on properties of sound wave at macroscopic (observable) level. The results show achievement caused them importantly obtaining a better conceptual score at the end of their learning. In addition, the qualitative analysis suggests the students had changed their conceptual understanding on physics of sound wave properties in three characteristics including differentiation, class extension, and reconceptualization. Srisawasdi (2012) introduced MBL and computer simulation to cover basic science concepts including three physical science activities; air resistance of falling objects, heat of fusion for ice and photosynthesis and respiration for 26 second year pre-service teachers in Thailand. Results indicate that all the groups did not perceive differently the goal and the support of computerized science laboratory. The highest attitude group realized the ease of use, self-learning, and value greater than the medium and the low attitude groups, but the medium attitude group possessed the most satisfaction with the laboratory.

2.3 Motivation

Motivation stands for an internal state that activates, guides, and maintains behavior, students' drive to learn and achieve to their potential at school. There are five components of motivation to learn science (Glynn et al., 2011) including; First, intrinsic motivation (IM) involves learning science for its own sake. Second, career motivation (CM) included the relevance of science to one's career is a leading theme in students' explanation. Third, self-determination (SDT) as dimension refers to the control students' believe that they have over their learning of science. Forth, self-efficacy (SEC) is an achievement by predisposing students to work harder, persist longer, and overcome barriers when pursuing academic goals and finally, grade motivation (GM) as the students' competition often associate with grade.

In term of science education, motivation to learn is the engagement related to the achieving goals, students' understanding of science and the activation of strategies for action (Lee & Brophy, 1996). Due to the relationship between motivation, cognitive engagement and conceptual change, motivation to learn is a particular issue to concern in science education. Pintrich, Marx, and Boyle (1993) have suggested that the construction of new knowledge in science is strongly influenced by prior knowledge, conceptions gained prior to formal learning. Consequentially, conceptual change is much tougher because it requires new information to engage at an adequately deep level to recognize conflicts between existing understanding and new information (DeBacker and Nelson, 2000). Confirmation persuades that decisions to engage in endeavored learning might be affected by individual students' motivation including engaging goals in an activity, beliefs in abilities and the nature of the task, and valuing of the task (Miller, Greene, and Montalvo, 1996; Nolen and Haladyna, 1990). In addition, learning environment, especially laboratory maybe the key factors affecting motivation differences.

3. Methodology

3.1 Research Objective

this study aims to investigate students' physics motivation delivered in blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory for physics learning of sound wave. Specifically, the main research questions for this study was that do the students engaged in blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory perform significantly better by students' physics motivation?

3.2 Study Participants

The total of 66 students-respondents in their eleventh grade (16-17 years old) were recruited in this present study. They were divided into two experimental groups which received different learning process in blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory: open-inquiry laboratory learning (N=31) and guided-inquiry laboratory learning (N=35) groups. Both groups were assigned to learn a physics lesson on sound wave. The researchers conducted an informal interview with physics teacher in two regular classes, and the results showed that all of students have basic skills on using computer. However, all of them have never experience yet using hands-on microcomputer-based laboratory and computer-simulated laboratory in physics class.

3.3 Instrument

In this study, a 25-item science motivation questionnaire was used to measure students' motivation to learn physics on five subscales: intrinsic motivation (IM), self-determination (SDT), self-efficacy (SEC), career motivation (CM), and grade motivation (GM) (Glynn et al., 2011). The questionnaire was originally developed by Glynn et al. (2011) and then adapted into Thai version to assess students' motivation to learn science. From 25 items English version, the translation an identical version in Thai was constructed and Cronbach's alpha of Thai version were 0.79, 0.81, 0.89, 0.81 and 0.85 for IM, SDT, SEC, CM and GM respectively (Srisawasdi, submitted).

3.4 Data Collection

For investigating students' physics motivation in whether they perform inquiry-based learning process with a blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory on sound wave phenomena. The study participants were asked to response the 25-item 5-point Likert-scale questionnaire for 10 minutes at both before and after interacting with the blended lab. On each item, respondents were assigned to rate how much the respondent agree with into five scale, from 1-strongly disagree to 5-strongly agree. In the blended lab class, both groups participated physics learning of sound wave through inquiry-based learning process for 480 minutes. Figure 1 and 2 illustrate example of blended lab activity for physics learning of sound wave.

Figure 1. Illustrative example of classroom learning activity through hands-on MBL guided (Left) and open (Right) inquiry laboratory

Fig. 2 Illustrative interface the bending interference of sound wave simulation (obtained from PhET) for computer-simulated guided and open inquiry laboratory

The statistical data techniques selected for analyzing students' perceptions was repeated-measures MANOVA in SPSS 21.0.

4. Results

The results for the repeated-measures MANOVA indicated significant main effect for different levels of inquiry (guide- and open inquiry) (Wilks' lambda=0.755, F (5, 60) = 3.887, p = 0.004, partial η^2 = 0.245). There was significant difference on students' physics motivation between guided- and open-inquiry learning process. According the significance, the univariate results was performed and it revealed that all of the five subscales on physics motivation reached a statistically significant difference between guided- and open-inquiry learning process. That is, both guided- and open inquiry in blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory performed differently with regard to IM, CM, SDT, SEC, and GM. In additions, there was a significant interaction effect between different levels of inquiry and different times measured (pre- and post-test) (Wilks' lambda = 0.717, $F(5, 60) = 4.738$, $p = 0.001$, partial $\eta^2 = 0.283$). This means that different levels of inquiry had similar effects on students' physics motivation in the blended lab.. In addition, there was a significant main effect for different time measured (Wilks' lambda = 0.483, $F(5, 60) = 12.855$, $p < 0.000$, partial $\eta^2 = 0.517$). This suggests that, on average, the students' physics motivation have changed over inquiry-based learning experience with blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory. Univariate analyses of variances on each subscale were conducted as follow-up tests to the one-way MANOVA. The results of the univariate test regarding different time measured are summarized in Table 1.

Table 1: The students' subscale means of physics motivation by time and univariate MANOVA

Subscale	Time		F	Sig.	η^2
	Pre-test	Post-test			
Intrinsic motivation (IM)	17.58 (3.123)	19.85(2.562)	29.920	0.000	0.319
Career motivation (CM)	17.17 (3.580)	19.92 (2.668)	39.803	0.000	0.383
Self-determination (SDT)	17.23 (2.971)	18.21 (2.551)	5.675	0.020	0.081
Self-efficacy (SEC)	14.55 (3.398)	15.91 (3.703)	11.467	0.001	0.152
Grade motivation (GM)	19.33 (2.879)	21.21 (2.551)	24.289	0.000	0.275

As displayed Table 1., The univariate MANOVA on the five subscale scores of physics motivation were significant differences across time, from pre-test to post-test. The univariate results revealed a significant effect on IM ($F_{1,64} = 29.920$, $p < 0.001$, partial $\eta 2 = 0.319$), CM ($F_{1,64} = 39.803$, $p < 0.001$, partial $\eta 2 < 0.383$), SDT ($F_{1,64} = 5.657$, $p < 0.05$, partial $\eta 2 = 0.081$), SEC ($F_{1,64} = 11.467$, $p < 0.01$, partial $\eta 2 = 0.152$), and GM ($F_{1,64} = 24.289$, $p < 0.001$, partial $\eta 2 = 0.275$). According to aforementioned results, the overall result suggested that the increase of physics motivation regarding intrinsic motivation, career motivation, self-determination, self-efficacy, and grade motivation from the pre-test to post-test was homogeneous for both guided- and open-inquiry learning process after participating with blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory. That is, there was effect of different levels of inquiry on students' physics motivation for learning with blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory.

5. Discussion

This research reports an effect of innovative teaching and learning of physics, inquiry-based learning process in a blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory, on promoting students' physics motivation. The result indicated an increasing of students' physics motivation scores considering from before and after participating with the intervention. This finding could be discussed that inquiry types subjecting to five subscales of physics motivation are shown in Table 1. In cases of guided inquiry with innovatively effective tools of MBL and computer simulation persuaded students to particularly focus on given content, interference of sound wave and delivered them opportunities to construct knowledge with team groups (Zion, Cohen, and Amir, 2007) and prospectively achieved conceptual change after students attended this learning process. This indicated that learners were more satisfied to explore a knowledge construct themselves with MBL. This study showed time as the main effect of the learning process to the students' motivation on sound wave interference. According to the results, there were statistical significant effect in all subscales of the students' physics motivation. This result implied that the learning process which were MBL and computer simulation could motivate the student to learn physics. Due to learning using both MBL and computer simulation, the students had higher scores of the motivation. Considering to MBL, this method allows students to learn trough actual laboratory using technology as tools supporting their learning process (Russel, 2003) and computer simulation by which they had background of sound wave interference that invisible in micro level(Srisawasdi, 2008). The result consistent with the research findings that students perform better in physics concepts with learning from integrating of MBL and computer simulation(Gunhaart and Srisawasdi, 2012) . A possible explanation for why learners made develop on physics motivation from before to after is that the teaching and learning could induce learners into the problem solving (Russell et al., 2003; Thornton and Sokoloff, 1989)

Considering different levels of inquiry (open- and guided inquiry), the findings introduced to acquire more effective process. One of the best findings was that open-inquiry laboratory learning was more effective learning process to motivate student in physics learning of sound wave than that of guided-inquiry laboratory learning. The results showed there was a significant difference for all of motivational subscales in both inquiry levels. This evidence is consistent with the claims that the inquiry are a well designed learning process for science learning and can engage mindful investigation in doing science (Bunterm et al., 2014; Sadeh a Zion, 2011). Moreover, open-inquiry learning through computer simulation affected students revising unscientific understanding and improving their physics outcomes (Srisawasdi, 2014). Also, Srisawasdi (2012) has mentioned that hands-on microcomputer-based laboratory support improvement of attitude and perception toward learning. Therefore, this implied that using inquiry-based could support the students' leaning in affective domain such as motivation. Especially, open-inquiry laboratory learning process where student have opportunity to design, collect and analysis data, discuss with peers, make conclusion and communicate findings by their own way delivered them a novel learning process of science can motivate to learn physics greater than prescribed physics experimentation.

6. Conclusion

This study investigated an effect of inquiry-based learning process into a blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory on secondary school students' physics motivation. After implementing the intervention, the results show that; (i) both guided- and open-inquiry learning process in blended combination of hands-on microcomputer-based laboratory and computer-simulated laboratory improved students' intrinsic motivation, career motivation, self-determination, self-efficacy, and grade motivation towards physics learning across time; and (ii) open-inquiry laboratory learning process was more effectively to enhance students' physics motivation than the guided inquiry. To this end, blended combination of physical, hands-on microcomputer-based laboratory, and virtual, computer-simulated laboratory, environment could be used to motivate student in learning of physics by inquiry. However, to address students' conceptual learning performance we are going to investigate how to use the inquiry-based learning process through blended lab for facilitate development of mental model and ability of scientific reasoning.

Acknowledgements

This work was financially supported by the National Research University (NRU) program of Khon Kaen University, and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Abd-El-Khalick, F. A., & Akerson, V. L. (2004). Learning as conceptual change: Factors mediating the development of preservice elementary teachers' views of nature of science. *Science Teacher Education, 88*(5), 785-810. doi: 10.1002/sce.10143

Bell, R. L., & Trundle, K. C. (2008). The use of a computer simulation to promote scientific conceptions of moon phases. *Journal of Research in Science Teaching, 45*(3 , March), 346-372. doi: 10.1002/tea.20227

Blake, C., & Scanlon, E. (2007). Reconsidering simulations in science education at a distance: features of effective use. *Journal of Computer Assisted Learning 23*, 491-502. doi: 10.1111/j.1365-2729.2007.00239.x

Buck, L. B., Bretz, S. L., & Towns, M. H. (2008). Characterizing the Level of Inquiry in the Undergraduate Laboratory. *Journal of College Science Teaching, n.p.*, 52-58.

Bunterm, T., Lee, K., Kong, J. N. L., Srikoon, S., Vangpoomyai, P., Rattanavongsa, J., & Rachahoon, G. (2014). Do different levels of inquiry lead to different learning outcomes? a comparison between guided and structured inquiry. *International Journal of Science Education*. doi: 10.1080/09500693.2014.886347

Cohen, L., Manion, L., & Morrison, K. (2000). *Research Method in Educaton*. USA and CANADA: Taylor&Francis Group.

Colburn, A. (2000). An inquiry primer. *Science Scope, 23*(6), 42-44.

DeBacker, T. K., & Nelson, R. M. (2000). Motivation to learn science: Differences related to gender, class type, and ability. Journal of Educational Research. *93*(4), 245-255.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire ii: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching, 2*(10), 1159-1176.

Gunhaart, A., & Srisawasdi, N. (2012). Effect of integrated compute-based laboratory environment on students' physics concepttual learning of sound wave properties. *Procedia - Social and Behavioral Sciences, 46*, 5750 – 5755. doi: 10.1016/j.sbspro.2012.06.510

Guzey, S. S., & Roehrig, G. H. (2009). Teaching science with technology: Case studies of science teacher development of technology pedagogy, and content knowledge. *Contemporary Issues in Technology and Teacher Education, 9*(1).

Hamzah, M. S., & Mdzain, A. N. (2010). The effect of cooperative learning with dslm on conceptual understanding and scientific reasoning among form four physics students with different motivation levels. *Bulgarian Journal of Science and Education Policy 4*(2).

Hofstein, A., Navon, O., Kipnis, M., & Naaman, R. M. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching 00*(00), 1-16.

Hola, O. (2007). Computer - based education in physics *Problems of Education in the 21st Century, 1*, 44-49.

Lee, O., & Brophy, J. (1996). Motivational patterns observed in Sixth-Grade science classrooms. *Journal of Research in Science Teaching, 33*(3), 303-318. doi: 10.1002/(SICI)1098-2736(199603)33:3<303::AID-TEA4>3.0.CO;2-X

Miller, R. B., Greene, B. A., & Montalvo, G. P. (1996). Contemporary educational psychology *Bhuvaneswari ravindran, 21*, 388-422.

Minstrell, J., & VanZee, E. H. (2000). Inquiring into inquiry learning and teaching in science

Mulder, Y. G., Lazonder, A. W., & Jong, T. d. (2011). Comparing two types of model progression in an inquiry learning environment with modelling facilities. *Learning and Instruction, 21*, 614-624.

Nolen, S. B., & Haladyna, T. M. (1990). Motivation and studying in high school science. *Journal of Research in Science Teaching, 27*(2), 115-126. doi: 10.1002/tea.3660270204

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change *Journal of Research in Science Teaching, 45*(3), 167-199.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1997). On the effectiveness of activeengagement microcomputer-based laboratories. *American Journal of Physics, 65*(1), 45-55.

Russell, D. W., Lucas, K. B., & McRobbie, C. J. (2003). The role of the Microcomputer-Based Labolatory display in supporting the construction of new understanding in kinematics. *Research in Science Education, 33*, 217-243.

Sadeh, I., & Zion, M. (2011). Which Type of Inquiry Project Do High School Biology Students Prefer: Open or Guided? *Research in Science Education*. doi: 10.1007/s11165-011-9222-9

Schunk, D. H. (2005). Self-regulated learning: the educational legacy of paul r. pintrich. *Educational Psychologist, 40*, 85-94.

Singer, J., Marx, R. W., & Krajcik, J. (2000). Constructing extended inquiry projects: Curriculum materials for science education reform. *Educational Psychologist, 35*, 165-178.

Singh, M. K., Singh, S., Kumari, A., & Kumar, P. (2011). Teach biology science using a computer simulation process. *International Transactions in Applied Sciences, 4*(2), 267-270.

Srisawasdi, N. (2008). *Development of computerized laboratory environments to promote active science learning: nose simulator and contact angle simulation.* (Ph.D. Dissertation), Mahidol University.

Srisawasdi, N. (2012). Student teachers' perceptions of computerized laboratory practice for science teaching: a comparative analysis. *Social and Behavioral Sciences, 46*, 4031- 4038. doi: 10.1016/j.sbspro.2012.06.192

Srisawasdi, N. (2014). Developing technological pedagogical content knowledge in using computerized science laboratory environment: an arrangement for science teacher. *Research and Practice in Technology Enhanced Learning, 9*(1), 123-143.

Srisawasdi, N., & Kroothkaew, S. (2014). Supporting students' conceptual development of light refraction by simulation-based open inquiry with dual-situated learning model. *Journal of Computers in Education*. doi: 10.1007/s40692-014-0005-y

Thornton, R. K., & Sokoloff, D. R. (1989). Learning motion concepts using real-time microcomputer-based laboratory tools. *Ameican Association of Physics Teachers, 58*(9), 858-867.

Zion, M., Cohen, S., & Amir, R. (2007). The spectrum of dynamic inquiry teaching practices. . *Research in Science Education, 37*(4), 423-447.

An Experimental Study on the Effects of an Online Student-Constructed Tests Learning Activity

Fu-Yun YU* & Chia-Ling SU
Institute of Education, National Cheng Kung University, Taiwan
*fuyun.ncku@gmail.com

Abstract: While the learning potential of student-constructed tests for the promotion of knowledge integration and elaboration has been suggested, its learning effects warrant further empirical examination. Three fifth-grade classes (N=76) participated in this study for nine consecutive weeks. A one-group pre-post experimental research design was used, and an online student-constructed tests learning system was adopted to support elementary students' science learning. The results from the paired *t*-tests found significant increase in students' attitudes toward science and science learning motivation as a result of the incorporated activity. Yet, no significant differences were found in students' use of cognitive and metacognitive strategies after the activity. Based on the collected data, suggestions for instructional implementations are provided.

Keywords: Experimental study, learning effects, online learning system, primary school settings

1. Introduction

The learning benefits of student-generated questions (hereafter name SQG) have been well established empirically. In general, empirical evidence accumulated since the 1960s provides a solid basis for its effects on enhancing understanding, academic achievement, motivation, question-generation abilities, the use of cognitive and metacognitive strategies, problem-solving abilities and attitudes toward the subject matter studied (Brown and Walter, 2005; Chin, Brown and Bruce, 2002; Dori and Herscovitz, 1999; English, 1997; Perez, 1985; Rosenshine, Meister and Chapman, 1996; Yu and Liu, 2008).

Recently, researchers have experimented the idea of engaging students in constructing a test and found promising evidence for its potential. Specifically, data on students' perceptions found that students' preference to and perceptions of student-constructed tests (SCT) and teacher-constructed tests (TCT) were statistically significant at $p < .01$ with a considerable proportion of students preferring SCT as the approach for assessing their learning and regarding SCT as a better approach for promoting learning (Yu, 2013). Descriptive data analyzed further highlighted the potential of SCT for the promotion of knowledge integration (Yu and Su, 2013a) and knowledge elaboration (Yu and Su, 2013b).

Constructing "tests" is different from constructing questions, and it would direct attention to additional criteria. Since a more holistic view of the study content may be obtained (Yu and Su, 2013a), and cognitive processes of different nature and intensity may be mobilized, the learning effects of SCT was the focus of this study. To provide comprehensive information about the observed phenomena in educational context, the learning effects on both cognitive and affective (specifically, the use of cognitive and metacognitive strategies, attitudes toward science and science learning motivation) are examined in this study.

2. Methods

2.1 Participants and instructional content

Students from three intact fifth-grade classes (N=76) taught by the same science teacher were briefed about the purpose of the introduced online learning activity (i.e., support of their science learning; the promotion of higher-order thinking skills, including self-monitoring comprehension level when attending lectures, grasping the main ideas of the study content; evaluating self- and peers' learning by constructing a set of questions of appropriate scope and difficulty) and invited to participate. All students at the participating schools started taking computer classes when they were at the 3rd grade and so had basic computer skills needed to carry out the activity.

Two science units were covered during the study. The first unit is about "how heat affects matter" with three lessons that cover topics including the changes of matter after heating, heat transfer, and insulation. The second unit is on "air and burning" with three lessons. Topics covered include the characteristics of oxygen and carbon dioxide, their uses in daily life and their relationship to burning matter; three elements of combustion; and fire extinguishing and the fire prevention.

2.2 Implementation procedures

The implementation procedure is delineated in Figure 1. A pilot study involving one fifth-grade class in the participating school (N=28) was conducted to ensure that the planned procedures and time allocation for various activities were appropriate prior to the actual study. Data on participants' cognitive and metacognitive strategy use, attitudes toward science and science learning motivation was collected prior to the commencement of this study.

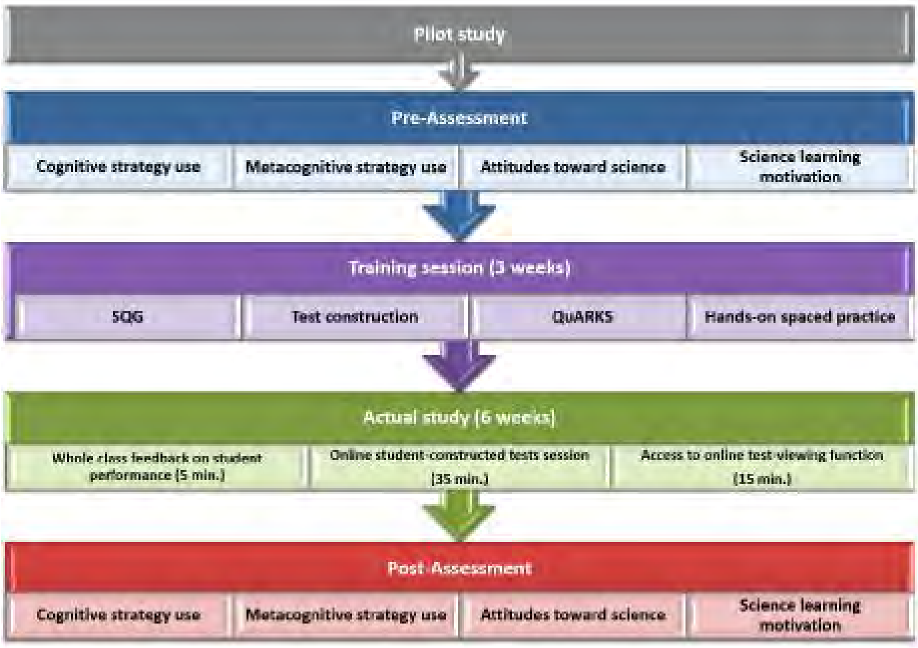


Figure 1. *Experimental procedures of this study*

This study took place right after the school-wide first-term exam. For this study, three of the most frequently encountered question types in primary schools in Taiwan were chosen for the learning activity—true/false, matching and multiple-choice questions. For the duration of the study (i.e., nine consecutive weeks in total), as a routine, participating students would head to the computer lab after attending three 40-minute instructional sessions on science in their respective class. To equip participants with essential knowledge and skills associated with the engaged tasks, three sessions were reserved for training prior to the study. During the training session, quality criteria frequently associated with SQG and SCT and basic principles for item writing for each of the chosen question types were introduced and explained. In addition, the operational procedures for the adopted system were demonstrated, followed up by students' hands-on spaced practice activities.

For each of the following six weeks, at the beginning, whole-class feedback on student performance at the previous SCT activity was arranged with reference to SCT criteria (e.g., covering all main topics, appropriateness of test difficulty level in general, appropriateness of coverage and representation of all main topics). Afterwards, students were directed to construct a test around the science content covered in the prior three instructional sessions by composing a minimum of five question items consisting of at least two out of the three chosen question types. A post-session questionnaire were disseminated to participants for individual completion after the study.

2.3 The Online SCT Learning System

An online learning system supporting associated activities of the study was adopted. Students in this study had access to test-construction and test-view functions of the adopted system.

To construct a test, students first design the overall structure of a test in terms of the number and scoring scheme of each question type. Second, students generate questions out of any of the three question types of their choice. After satisfying a number of questions have been generated, students then view and select individual questions to be included in the test at work. Finally, students can determine and re-arrange the relative sequence of questions both within and among question types before submission.

To promote learning by permitting students to learn from observing peer's work, an observational learning space—test-viewing was created and made accessible at the last 15 minutes of each online learning activity.

2.4 Measurement instrument

Three instruments were used in the study to test the learning effects of online student-constructed tests. First of all, Hung's (2002) "Learning Strategy Use Scale" was adopted. The scale consisted of two parts: "Cognitive Strategies Use Scale" (18 items) and "Metacognitive Strategies Use Scale" (24 items). The former appraises students' use of rehearsal, elaboration and organization learning strategies, and the latter reveals students' activation of metacognitive strategies for cognition regulation, such as planning, monitoring, revising and evaluating one's actions and reasoning while learning. All items were rated on a 6-point Likert scale, with corresponding verbal descriptions ranging from "no consistency" through "very inconsistent," "somewhat inconsistent," "somewhat consistent," "very consistent," to "complete consistency." The internal consistency reliability calculated after this study was .92 and .94 for the "Cognitive Strategies Use Scale" and "Metacognitive Strategies Use Scale," respectively.

Second, "Attitude toward Science in School Assessment" developed by Germann (1988) was adopted to measure students' attitudes toward science. To ensure that the instrument was translated appropriately and adequately, back translation technique was used. Results from the exploratory factor analysis and Cronbach's α with a group of fifth-grade students (N=30) by Tsai (2010) substantiated its validity and reliability. The instrument consisted of fourteen Likert-scale items. Each statement was rated on a five-part discrete scale, with corresponding verbal descriptions ranging from "strongly disagree" through "disagree," "no-opinion," "agree," to "strongly agree." The Cronbach's alpha values calculated after the study (N=149) was 0.88.

Finally, Hung's (2002) "Science Learning Motivation" was adopted for this study. The scale consisted of 14 items and used a six-point Likert scale (ranging from 1=no consistency to 6=complete consistency). The scales validated by a group of 303 sixth-graders evidenced good validity. The Cronbach's alpha values calculated after the study (N=149) was .94.

3. Results

As shown in Table 1, after exposed to the SCT activity, students not only activated more of cognitive and metacognitive strategy while learning science, but also formed better attitudes and exhibited heightened motivation toward science. Nevertheless, the results from one-group paired *t*-tests found significant differences only in attitudes toward science and science learning motivation, but not in the cognitive domains.

Table 1: Descriptive and inferential statistics on four observed variables

Variables	*n*	*M*(*SD*)	*t*	η^2
Cognitive strategy use				
Pretest	76	3.86(0.95)	-1.12	.02
Posttest		4.01(1.20)		
Metacognitive strategy use				
Pretest	76	4.00(0.93)	-1.56	.03
Posttest		4.17(1.19)		
Attitude toward science				
Pretest	76	4.41(1.27)	2.89^*	.10
Posttest		4.76(1.09)		
Science learning motivation				
Pretest	75	4.10(1.14)	3.13^*	.17
Posttest		4.47(1.15)		

$^*p < .05$

4. Discussion and Conclusion

Preliminary studies on the potential of SCT supported its effects on knowledge integration and elaboration (Yu and Su, 2013a; Yu and Su, 2013b). The current study extended prior studies by substantiating its affective effects. By allowing students to construct questions around the study materials they regard as important and relevant, to allocate different weighting among different study topics and to decide the relative sequence of question items within and among question types, SCT in essence is more in alignment with what constructivism, self-regulation and self-determination theories accentuate. As a result, as found in this study, exposing students to SCT helped to increase students' attitudes toward science and science learning motivation.

However, the current study failed to find SCT helped to promote the use of cognitive or metacognitive strategies. Through in-depth analysis of the current study and prior studies, some possible explanation for the unconfirmed results are rendered. First of all, the current study involved fifth-grade students (average age=11), who just reached Piaget's formal operational cognitive development whereas prior studies involved university students, who should be mentally more prepared and ready for the whole range of tasks involved in SCT. Second, participants in this study were directed to construct tests around the study materials on a weekly basis, which may not be in its entirety. Unlike prior studies, SCT activity was arranged around the end of the semester where opportunities for inter-connecting and integrating of all topics are provided.

Based on the findings of this study, it is suggested that instructors can engage students in SCT activity for the promotion of students' affective development, in specific, attitudes and learning motivation toward the learned subject.

Acknowledgements

This paper was funded by research grant from the National Science Council, Taiwan (Project: Online student-generated tests learning system: Development, applicability and learning effects, NSC 102-2511-S-006 -003 -MY3).

References

Brown, S. I., & Walter, M. I. (2005). *The art of problem posing* (3rd Edition). New Jersey: Lawrence Erlbaum Associates.

Chin, C., Brown, D. E., & Bruce, B. C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education, 24*, 521-549.
Dori, Y. J., & Herscovitz, O. (1999). Question-posing capability as an alternative evaluation method: Analysis of an environmental case study. *Journal of Research in Science Teaching, 36*(4), 411-430.
English, L. D. (1997). Promoting a problem-posing classroom. *Teaching Children Mathematics, 4*(3), 172-179.
Germann, P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching, 25*(8), 689-703.
Hung, C-C. (2002). *Effects of question-posing and cooperative learning on students' learning outcomes within a web-based learning environment*. Unpublished Master's Thesis. National Cheng Kung University, Tainan, Taiwan.
Perez, J. A. (1985). *Effects of student-generated problems on problem solving performance*. Unpublished doctoral dissertation. Teachers College, Columbia University, New York, NY.
Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: a review of the intervention studies. *Review of Educational Research, 66*(2), 181-221.
Tsai, H. T. (2010). *The effects of guided short-answer online student question-generation on elementary student science learning*. Unpublished Master's Thesis. National Cheng Kung University, Tainan, Taiwan.
Yu, F. Y. (2013). Preliminary assessment of online student-generated tests for learning. In Tan, S. C., Wu, Y. T., Apko, T. W., Wong, L.-H., Liu, C.-C., Hirashima, T., Sumedi, P., & Lukman, M. (Eds.), *Workshop Proceedings of the 21st International Conference on Computers in Education* 2013 (pp. 183-186). Bali, Indonesia, November 18-22, 2013.
Yu, F. Y., & Liu, Y. H. (2008). The comparative effects of student question-posing and question-answering strategies on promoting college students' academic achievement, cognitive and metacognitive strategies use. *Journal of Education & Psychology, 31*(3), 25-52.
Yu, F. Y. & Su, C-L (2013a). The design, development and preliminary evaluation of an online student-generated tests learning system. In Wong, L.-H., Liu, C.-C., Hirashima, T., Sumedi, P., & Lukman, M. (Eds.), *The Main Proceedings of the 21st International Conference on Computers in Education* (pp. 375-378). Bali, Indonesia, November 18-22, 2013.
Yu, F. Y. & Su, C-L (2013b). The construction and use analysis of student-constructed tests online learning system. *Taiwan Academic Network Conference 2013*, National Chung Hsing University, Taichung, Taiwan, October 23-25.

Exploring the Effects of Student Question-Generation Strategy

Chun-Ping WU[a*], Shu-Ling WU[b] & Ching-Chiu YEH[c]
[a]*Department of Educational Technology, TamKang University, Taiwan*
[b]*Graduate Institute of Human resource and eLearning Technology, National Hsinchu University of Education, Taiwan*
[c]*Department of Educational Technology, TamKang University, Taiwan*
*cpwu303@gmail.com

Abstract: The purpose of this study is to explore the effect of student question generation strategy on students' reported use of cognitive strategies and metacognitive strategies. Furthermore, the relationship between the SQG performance and academic achievement was also examined. A single group experiment was implemented for 7 weeks. Seventy-two junior high school students from two intact history classes were recruited. Participants were engaged in the SQG task followed by the peer-assessment activity. The finding supported the positive effect of the SQG on enhancing students' use of cognitive strategies and metacognitive strategies. Additionally, students' SQG performance was significantly correlated with their academic achievement, which was supported.

Keywords: Student Question Generation, cognitive strategies, metacognitive strategies, academic achievement

1. Introduction and Literature Review

Students' ability to raise a good question relies on their use of existing knowledge to observe and interpret the newly learned content or phenomena. Therefore, it brings the needs to explore how to facilitate students in bridging the new content with their knowledge bases. The student question-generation (SQG) strategy, which is grounded on the constructivism and information processing theory, has gained more attention from the researchers and educators (i.e. Abramovich & Cho, 2006; Berry & Chew, 2008; Brown & Walter, 2005; van Blerkom, van Blerkom, & Bertscho, 2006; Yu & Wu, 2013). The question -composing and revising task could engage students in recalling, organizing or elaborating the newly learned content.

Specifically, the SQG process requires students to recall the content they just learned and identified important concepts and the concepts which their peers might be confused about. Those identified concepts could be used as the focus of the question. For example, while designing the multiple-choice question, students have to examine the interconnection among concepts and translate their understanding into the question stem in their own words or using appropriate examples. Additionally, while designing the correct answer and the three distractors, the question authors experience a micro problem-solving process (Yu, Liu, & Chan, 2005). They have to propose different solutions to the questions and examine and compare the solutions to ensure one best correct answer. Therefore, the question-generation process engages students in organizing, analyzing the learned contents, examining their understanding and misconception (Lee & Hutchison, 1998), and elaborating the contents in a meaningful way which helps to construct their schema (Bangert-Drowns, Hurley &Wilkinson, 2004; Herbert & Burt, 2004).

The SQG effects on enhancing students learning motivation, confidence, understanding of the learning materials, metacognition and so on have been supported in empirical studies. (Abramovich & Cho, 2006; Barlow & Cates, 2006; Belanich, Wisher,& Orvis, 2004; Berry & Chew, 2008; Brown & Walter, 2005; Choi, Land, & Turgeon, 2005;Dori & Herscovitz, 1999,2005; Fellenz, 2004; Ikuenobe, 2001; van Blerkom, van Blerkom, & Bertscho, 2006; Whiten, 2004; Wilson, 2004; Yu, 2005, 2009; Yu & Liu, 2005)

The purpose of this study is to further validate whether the above-mentioned SQG process would enhance students use of cognitive strategies and strategies in the context of junior high school' history course. Furthermore, as suggested, the SQG might help students' deep understanding of the learned content. The second purpose of this study is to examine whether students with better question-generation performance also performed better in the academic achievement tests.

2. Research Method

2.1 Research Design

Seventy-two junior high school students from two intact history classes taught by the same instructor were invited to participate in this study. A single group experimental design was implemented for the seven weeks. At the beginning, the purposes of the question-composing activity were explained to the participants followed by the training.

As suggested by theories and literature, the question composing task is difficult especially for those students without question composing experience (Yu, 2009), thus training on question posing is essential. Students who do not have knowledge of the quality criteria of a good question or are not familiar with the reasoning process of composing a question, might devote efforts to composing questions measuring the facts rather than higher level questions. Furthermore, without developing the schema of question-composing process, they might encounter difficulty in either translating the concepts into question stem or offering the groups of options that are highly related to the question stem. On the basis of the needs for the training, this study incorporated several components into training: the quality criteria of a good multiple-choice question stem and four options, which include one answer and three distractors, the reasoning process of question posing and revision, the explanation of the value of the SQG and hands-on practice followed by feedback.

During the intervention, the participants were required to compose two to three multiple-choice question items in accordance to each of the five instructional topics. A peer-assessment activity was conducted at the instructional session followed by the SQG activity. A whole-class feedback on SQG performance was provided and the peer's comments collected during the peer-assessment activity were sent to the question-author.

2.2 Variables and Instruments

The examined cognitive strategies were defined as students' reported use of rehearsal, elaboration and organization strategies while the metacognitive strategies were defined as their reported use of planning, monitoring and self-judgment during the learning process. These two variables were measured by the translated version of Motivated Strategies for Learning Questionnaire (MSLQ) (Garcia & Pintrich, 1995). The Cronbach's α for cognitive (10 items) and metacognitive strategies (11 items) were 0.90 and 0.89 respectively. Students rated themselves on a seven point Likert scale from "not at all true of me" to "very true of me". Scales were constructed by taking the sum of the scores of items that make up that composite construct of the scale.

Additionally, to examine SQG performance, all the questions were evaluated by two independent raters. The evaluation criteria were adopted from the index, proposed by Yu & Wu (2013) and were revised in accordance with the course instructor's suggestions. The criteria included four dimensions: Importance, fluency, elaboration and cognitive Level.

To establish the inter-rater reliabilities, one third of students composed questions were randomly selected from 822 questions and evaluated by another independent rater (N = 274). The results of the inter-rater reliability were $r = 0.87$, $p < 0.01$, which proved to be satisfactory.

To examine students' academic achievement in the five instructional units, students' performance in the school tests were collected.

3. Results and Conclusions

3.1 Findings of the SQG Effects on Students' perceptions

The descriptive statistics of the variables are listed in Table 1. It can be seen that the post-test scores of students' reported use of cognitive strategies and metacognitive strategies (Mean=4.94, 4.74, respectively) are higher than the pre-test scores (Mean=4.52, 4.33, respectively).

Table 1: Descriptive statistics (N=72)

Variable	Use of Cognitive Strategy	Use of Metacognitive Strategy
Pretest Mean (SD)	4.52 (1.09)	4.33 (0.84)
Posttest Mean (SD)	4.94 (1.11)	4.74 (0.99)

The paired t-test result showed that the participants' post-test score of students' reported use of cognitive strategies is significantly higher than the pretest score. (t =2.91, p = .005). Similarly, the participants' post-test score of reported use of metacognitive strategies is significantly higher than the pretest score. (t =4.12, p < .01). In other words, the participants' reported use of cognitive strategies and metacognitive strategies were significantly enhanced after being engaged in the question-generation activity.

3.2 Findings of the relationship between SQG Performance and Academic Achievement

During the seven-week intervention, 822 questions were generated by 72 participants. As specified in the data analysis section, the questions were evaluated by the two raters using the pre-defined criteria and the ratings were adopted as the indicators for students' question-generation performance.

The mean scores of students' overall achievement, achievement in each unit and question-generation performance as well as the correlations among variables were presented in Table 2. As shown, the participants' overall question-generation performance is significantly correlated with their achievement. Furthermore, in order to explore the relationship between students' question-generation performance and their gained knowledge on each unit, the correlation analyses were conducted. The results show that participants' question-generation performance in unit 1 and 2 is significantly correlated with their achievement scores gained in the test of unit1 and 2.Similiar result was found in the unit3 and 4. The hypotheses that students who generated questions of better quality tended to perform better in the achievement tests were supported in this study. In other words, the questions students posed reflected their understanding and learning of the contents.

Table 2: Correlation among Variables (N=72)

Variable	Achievement (Unit1 &2)	Achievement (Unit3 &4)	Achievement (Unit5)	Overall Achievement	Mean (SD)
QGP at w1and 2	.33** (*p*=.005)	.27* (*p*=.02)	.21 (*p*=.08)	.28* (*p*=.02)	22.89 (5.12)
QGP at w3and 4	.21 (*p*=.08)	.29* (*p*=.01)	.17 (*p*=.15)	.23* (*p*=.05)	27.58 (6.90)
QGP at w5	.12 (*p*=.34)	.15 (*p*=.22)	.13 (*p*=.29)	.14 (*p*=.26)	11.26 (2.74)
Overall QGP	.27* (*p*=.02)	.31* (*p*=.01)	.20 (*p*=.09)	.27* (*p*=.02)	52.72 (11.61)
Mean (SD)	67.50 (21.25)	71.97 (16.58)	72.39 (20.95)	70.62 (18.73)	

Note: a. QGP refers to students' question-generation performance

b. * denotes $p< 0.05$, ** denotes $p<0.01$

3.3 Conclusion

This study contributed to the literature on student question-generation. First, this study validated the effects of SQG on students reported use of cognitive strategies and metacogntive strategies. Second, this study also validated the significant correlations between students' question –generation performance and academic achievement. The instructors who are interested in SQG strategies are suggested to teach students question-generation skills by providing them with guidance, deliberated practice opportunities and in-time feedback on their question-generation performance. As this study adopted the single group experimental design, future research is suggested to take a qualitative approach to explore the difficulty students might encounter during the question –generation process. It might help to understand any potential variables that might moderate the strength of relationship between the SQG and the three examined variables. Furthermore, this study focused on exploring the potential effects of one soft technology, the instructional strategy of SQG. As a result of the limited classroom facilities and teaching time, the hard technology, such as the online question-generation system, was not allowed to integrate in the SQG process. Future research may further explore whether adoption of the online question-posing system will further enhance students' engagement in the SQG process.

Acknowledgements

This paper was funded by the research grant from the National Science Council, Taiwan, ROC (NSC 101-2511-S-032-006). The author would like to thank research assistants, Hao jie Yong for his assistance on data collection.

References

Abramovich, S. & Cho, E. K. (2006). Technology as a medium for elementary preteachers' problem-posing experience in Mathematics. *Journal of Computers in Mathematics and Science Teaching, 25,* 309–323.

Bangert-Drowns, R. L., Hurley, M. M., &Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: a meta-analysis. *Review of Educational Research, 74,* 29–58.

Barlow, A., & Cates, J. M. (2006). The impact of problem posing on elementary teachers' beliefs about mathematics and mathematics teaching. *School Science and Mathematics, 106,* 64–73.

Belanich, J., Wisher, R. A., & Orvis, K. L. (2004). A question-collaboration approach to web-based learning. *American Journal of Distance Education,18*(3), 169-185.

Brown, S. I. & Walter, M. I. (2005). *The art of problem posing* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Choi, I., Land, S. M., & Turgeon, A. (2005). Scaffolding peer-questioning strategies to facilitate metacognition during online small group discussion. *Instructional Science, 33*(5/6), 483–511.

Craik, F. I. M. (2002). Levels of processing: Past, present. and future? *Memory, 10,* 305–318.

Dori, Y. J., & Herscovitz, O. (1999). Question posing capability as an alternative evaluation method: Analysis of an environment case study. *Journal of Research in Science Teaching, 36,* 411– 430.

Dori, Y. J., & Herscovitz, O. (2005). Case-based long-term professional development of science teachers. *International Journal of Science Education, 27*(12), 1413–1446.

Fellenz, M. R. (2004). Using assessment to support higher level learning: The multiple choice item development assignment. *Assessment and Education in Higher Education, 29*(6), 703-719.

Herbert, D. M. B., & Burt, J. S. (2004). What do students remember? Episodic memory and the development of schematization. *Applied Cognitive Psychology, 18,* 77–88.

Ikuenobe, P. (2001). Questioning as an epistemic process of critical thinking. *Educational Philosophy and Theory, 33*(4), 325–41.

Lee, A. Y., & Hutchison, L. (1998). Improving learning from examples through reflection. *Journal of Experimental Psychology: Applied, 4,* 187–210.

van Blerkom, M. L., & van Blerkom, D. L. (2004). Self-monitoring strategies used by developmental and non-developmental college students. *Journal of College Reading and Learning, 34* (2), 45-60.

Whiten, D. J. (2004). Exploring the strategy of problem posing. In G.W. Bright & R. N. Rubenstein (Eds),

Professional development guidebook for perspectives on the teaching of mathematics: 38 *companion to the sixty-sixth yearbook* (pp. 1–4). Reston, VA: National Council of Teachers of Mathematics.
Wilson, E. V. (2004). ExamNet asynchronous learning network: Augmenting face-to-face courses with student-developed exam questions. *Computers & Education*, *42*(1), 87-103.
Yu, F. Y. (2005). Promoting metacognitive strategy development through online question-generation instructional approach. *Proceeding of International Conference on Computers in Education 2005* (pp. 564-571). Singapore: Nanyang Technological University
Yu, F. Y. & Liu, Y. H. (2005). Potential values of incorporating multiple-choice question construction for physics experimentation instruction. *International Journal of Science Education*, *27*, 1319–1335.
Yu, F.-Y., Liu, Y.-H., & Chan, T.-W. (2005). A networked question-posing and peer assessment learning system: A cognitive enhancing tool. *Journal of Educational Technology Systems*, *32*(2 & 3), 211-226.
Yu, F.-Y., & Liu, Y.-H. (2009). Creating a psychologically safe online space for a student-generated questions learning activity via different identity revelation modes. *British Journal of Educational Technology, 40*(6), 1109-1123.
Yu, F.-Y., & Wu, C.-P. (2013). Predictive Effects of Online Peer Feedback Types on Performance Quality. *Educational Technology & Society, 16* (1), 332–341.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Structured Explanation Generation for Conceptual Understanding in Physics

Tomoya HORIGUCHI[a*], Takahito TOUMOTO[b] & Tsukasa HIRASHIMA[c]
[a] *Graduate School of Maritime Sciences, Kobe University, Japan*
[b] *Faculty of Engineering Division II, Tokyo University of Science, Japan*
[c] *Department of Information Engineering, Hiroshima University, Japan*
* horiguti@maritime.kobe-u.ac.jp

Abstract: In science education, usual problem practice hardly helps students reach ‘conceptual understanding’ with which they can solve various problems by making appropriate models of target systems. Students often superficially read the solution of a problem and apply it wrongly to others without understanding the model. It is difficult to teach how to make appropriate models because model-making expertise includes a lot of implicit knowledge. In this paper, we propose a general framework for systematically describe such knowledge, which makes it possible not only to explain various model and difference between them but also to design/sequence a set of problems appropriate for promoting conceptual understanding. Our framework was proved useful through a preliminary experiment in which the explanations generated based on our framework promoted subjects’ conceptual understanding in mechanics.

Keywords: science education, problem practice, conceptual understanding, explanation generation, semantics of constraints

1. Introduction

In science education, one serious drawback of current 'problem practice' is that most students fails to acquire the ability to make an appropriate model for a given task. A domain expert (such as physicist) can model not only the behavior of a system in question, but also she/he can do so in various tasks. Her/his model is always necessary and sufficient for answering the question. Such expertise consists of identifying the structure/state of the system in question and deciding the applicable principles/laws for modeling the behavior of the system. We call such ability 'conceptual understanding' of the domain.

Needless to say, it is very difficult for students to reach such an understanding through problem practice. Instead of considering the model, they often rely on the 'solution' they previously learned (e.g., the procedure of calculation). As a result, they wrongly apply the solution of one problem to another in which the solution is inapplicable. They also can't apply the solution they previously learned to another in which the solution is applicable. Even experience in many problems doesn't help them (Bransford, Brown and Cocking, 2000; VanLehn, 1998; VanLehn and van de Sande, 2009). Without models, the students occasionally succeed (by accident), and fail in many cases.

The major problem is that novice students tend to generate naive representation of a problem focusing on its superficial features (called 'surface structure'). They can't generate the representation based on the structural features (called 'physical structure') (Chi, Feltovich and Glaser, 1981; Larkin, 1983; Larkin, 1985). Therefore, instead of applying principles/laws to make the model, they often apply inappropriate solution based on the superficial similarity between problems (VanLehn, 1998), or use general strategy for operating mathematical equations without considering their physical meanings (Larkin, 1981).

In order to reach conceptual understanding, therefore, students need to learn (1) to infer the structural features of problems from the superficial features, and (2) to apply appropriate principles/laws to structural features to make models necessary for solving problems. For assisting them in problem practice, it is necessary to explain not only how each problem is solved but also why the solution is possible and what physical meaning it has. That is, it must be explicit why the principles/laws are applicable to the given situation (i.e., surface structure) and what physical meaning (physical structure) they imply. Additionally, it is important to explain not only the solution of a problem but also the relation (difference) between problems, that is, how the solution (applicable principles/laws) changes when the situation (problem) is changed. Furthermore, it would promote such learning to provide students with an appropriately designed and sequenced set of problems (Scheiter and Gerjets, 2002; Scheiter and Gerjets, 2003; VanLehn and van de Sande, 2009).

In current problem practice, such instruction has been rarely focused on, at most given by a few (experienced) teachers individually and implicitly. Especially, there have been few intelligent tutoring systems which can explain the relation between arbitrary two problems, and adaptively sequence problems considering the learning effect of order. We think this is because most of the knowledge necessary for such instruction is implicit and difficult to systematize, therefore there have been no general framework for indexing various types of problems.

In this paper, we propose a general framework for indexing problems, based on which explanation generation and problem sequencing mentioned above can be automatized. In our framework, making a model in physics is regarded as a process in which various constraints (applied principles/laws and modeling assumptions) are imposed on the target system and its behavior. A model is regarded as the set of constraints. We first formulate the model-making process in physics, then analyze the constraints which compose a model to systematize them based on their physical meanings and roles (functions). After that, we describe the applicable conditions of principles/laws in physics as a set of constraints. The constraints classified/defined in this manner are easily assigned to the situation of a problem. There are also some groups of constraints which are 'exclusive' each other (i.e., can't be valid simultaneously). Therefore, based on such classification and exclusiveness of constraints, it becomes possible to explain what physical meaning (structural features) superficial features of a problem have, what principles/laws are applicable to them and how applicable principles/laws change when the situation is changed. By indexing problems with this framework (we call it 'Semantics of Constraints: SOC'), it becomes possible to automatically extract the 'differences between problems' which is necessary for the comparison and sequencing of problems.

We first discuss the required knowledge and assistance necessary for conceptual understanding based on current research, then introduce the SOC framework. After that, we show the method for generating SOC-based explanations. The results of preliminary experiment are described which proved the usefulness of our framework. Finally, we conclude this paper and mention our future work.

2. Conceptual Understanding and Assistance

Research on problem-solving has revealed the knowledge structure domain experts in science have (Chi, Feltovich and Glaser, 1981; Larkin, 1981; Larkin, 1983; Larkin, 1985; VanLehn, 1998; VanLehn and van de Sande, 2009). Experts can (1) infer the structural features of problems with scientific concepts from the superficial features and generate the representation to which formal operations are applicable. They can also (2) generate an appropriate plan for solving the problem by operating the representation with the knowledge about qualitatively interpreted principles/laws. It is supposed that experts have acquired such knowledge by inducing the essential features through comparison of many problems and by transforming them into (some kinds of) 'schemata' or 'production rules' (VanLehn and van de Sande, 2009). It is, however, difficult for students reach such an understanding through usual problem practice. Even instructional innovations based on recent learning science research have limitedly improved students' understanding (Bransford, Brown and Cocking, 2000).

In order to promote such knowledge acquisition, it is effective to appropriately design a set of problems which includes positive/negative examples and *'near misses'* of various problem categories and to provide them in appropriate order to students (VanLehn and van de Sande, 2009) (in fact, it is reported problem order greatly influences learning (Scheiter and Gerjets, 2002; Scheiter and Gerjets, 2003)). In order to do that, it is necessary to explicitly describe (1) the superficial/structural features of problems and their relations, and (2) qualitative interpretations of principles/laws and their means of application. However, since most of such knowledge is implicit, there have been no general framework for systematically describe such knowledge. We think this is the reason though knowledge structure necessary for expertise was revealed and an effective instructional method was proposed, it haven't been widely practiced. The framework we propose makes it possible to systematically describe such knowledge, based on which the design of a set/sequence of problems and explanation generation for promoting conceptual understanding become possible.

3. Semantics of Constraints

Given a physics problem (which consists of a physical system and query), one makes a model necessary and sufficient for answering the query by embodying an appropriate part of the domain theory. Domain theory consists of a set of propositions each of which describes a principle/law, its applicable condition and resulting constraint(s) on the attribute(s) of the system. Constraints by embodied principles/laws are called the 'physical phenomenon constraints (PPCs).'

In making a model, various modeling assumptions are set for selecting appropriate principles/laws. Modeling assumptions define the structure/behavioral range of a system and physical phenomena to be considered. Since embodied physical phenomenon constraints are valid under some modeling assumptions, applicable conditions of principles/laws can be described with a set of modeling assumptions. That is, a physical phenomenon constraint always has its corresponding modeling assumptions. Constraints by modeling assumptions are called the 'modeling assumption constraints (MACs).'

Boundary condition of a system is given by the 'boundary condition constraints (BCCs).' They define the influence from the outside of the system. Making the influence which cannot be or need not be calculated with a model means defining the boundary of the model (i.e., what physical processes are considered/ignored). That is, a BCC always has its corresponding modeling assumptions.

In our framework, a model is the union of physical phenomenon constraints, boundary condition constraints and modeling assumption constraints. Usually, only the first two constraints are written as a model while the last constraints are remained implicit. However, MACs gives the validity to PPCs and BCCs. When modeling assumptions are changed, physical phenomena and boundary conditions also qualitatively change. In order to make a model correctly, therefore, it is necessary to under-stand the physical meaning of the constraints based on modeling assumptions (i.e., why an assumption is set and what role it plays). In most cases, such knowledge is acquired by a few students. In this research, we develop a framework for describing such knowledge explicitly, based on which the function for promoting conceptual understanding is designed. In the following two subsections, we elaborate on each class of constraints (BCC is omitted owing to limited space) to systematize their physical meanings and relations.

3.1 Modeling Assumption Constraints (MACs)

Modeling assumption constraints define the physical processes considered/ignored in a model. They are classified in two ways from different viewpoints: structural and functional.

The structural viewpoint focuses on defining the structure and its state of a model. The 'physical structure constraint' specifies what kind of objects, relations and their attributes in a system are considered. It corresponds to selecting a viewpoint, granularity or coordinate system of a system. An example is the specification about whether their mechanical relations/attributes (e.g., mass, applied forces) or their electrical ones (e.g., current, resistance) are considered. On the other hand, the 'operating range constraint' specifies the range within which a model is valid since physical phenomena occur assuming a system is in a specific state. For example, a model of a resistance assuming its value is constant needs the specification that its current and voltage are within the proportional range.

The functional viewpoint focuses on defining the boundary of a model to specify what kind of physical processes are considered/ignored. The 'process consideration constraint' makes such selection about physical processes of the same granularity (where, the 'out-sourcing/black-boxing constraint' ignores a physical process by put-ting it out of the system or into a black box regarding its effect as a boundary condition, and the 'process selection constraint' simply ignores a physical process and its effect). For example, assuming constant voltage supplied from outside is an out-sourcing constraint. Considering two parallel-connected resistors as a compound re-sister is a black-boxing constraint. Considering/ignoring the friction between two objects is a process selection constraint. The 'physical world constraint' maintains the fundamental laws of the physical world, such as 'rigid objects never overlap.' More microscopic physics is necessary to explain why this constraint is valid, that is, it specifies the physical processes of smaller granularity are ignored. The 'process simplifying constraint' substitutes the simplified process for an original complicated process in order to make the (mainly mathematical) calculation with a model easier. An example is to consider the behavior of a pendulum with small amplitude as simple harmonic oscillation (not as circular motion).

Constraint classes from the structural viewpoint are useful for enumerating modeling assumptions because they rather suggest the components and their relations of a system. For example, when a variable in an equation stands for a physical quantity, it is easy to infer an object and its attribute corresponding to the quantity is considered (which are physical structure constraints). Constraint classes from the functional viewpoint are useful for considering the meaning of modeling assumptions because they rather suggest the process structure (processes considered and their relations). For example, considering/ignoring a physical attribute (which is a physical structure constraint) suggests a physical process concerning the attribute is considered/ignored (which is a process selection constraint). That is, the classes from the structural view-point rather concern the surface structure of a problem, while the classes from the functional viewpoint rather concern its physical structure. Furthermore, as shown above, the classes from both viewpoints are related to each other based on their physical meanings. Therefore, with these classifications, it becomes possible to systematically describe the knowledge about the relation between superficial and structural features of problems.

Additionally, there are often sets of modeling assumption constraints which are mutually exclusive (can't be assumed simultaneously). For example, in the same time interval, 'transient state' and 'steady state' (which are operating range constraints) can't be assumed simultaneously. In the same (part of a) system, 'consider friction' and 'not consider friction' (which are process consideration constraints) can't be assumed simultaneously. Such exclusiveness between modeling assumptions gives important clues to identify the differences between models/problems (see section 4.2).

3.2 Physical Phenomenon Constraints (PPCs)

Relatively simpler physical phenomenon constraint is the 'physical device constraint' which arises within a component of a system. That is, it is a 'local constraint.' Since it indicates the physical property of the component, each domain has its specific physical device constraints (for example, Ohm's law constrains the values of current and voltage in an electric device). In contrast, there are 'global constraints' which indicates the behavior of multiple components or the whole system. Global constraints are classified as follows.

In general, a physical system evolves through time, starting from an initial state. It is either (1) changing dynamically, (2) in a steady state or (3) changes discontinuously. Therefore, we call the constraints in these states, the 'dynamic change constraint,' the 'steady state constraint' and the 'discontinuous change constraint,' respectively. Additionally, when a quantity is conserved through time, the constraint which indicates its amount is the same at arbitrary two time points is called 'conservation law constraint.'

Dynamic change constraint constrains the behavior of a system in a time interval during which it is changing dynamically. It often indicates the relation between the driving power of dynamic change and the influences on it. For example, Newton's second law (equation of motion) relates an object's acceleration with the forces applied to it. Steady state constraint constrains the behavior of a system in a time interval during which it is in a steady state. It indicates the balance/cancellation between influences on the driving power of dynamic change. An example is the equation of balance of forces about an object at rest. Discontinuous change constraint constrains the behavior of a system at a time point on which it changes discontinuously. It indicates the relation between the amounts of a quantity before and after the change. An example is the formula of coefficient of restitution. A quantity is called a 'conserved quantity' when its amount is constant during the temporal evolution of a system. Conservation law constraint indicates the amounts of a conserved quantity at arbitrary two time points are the same. An equation of heat exchange between two objects and an equation of conservation of energy/momentum are the examples.

A global physical phenomenon constraint aggregates a set of local physical phenomenon constraints. For example, Newton's second law (equation of motion), which is a dynamic change constraint in mechanics, includes a set of local constraints each of which indicates a force applied to the target object (physical device constraints such as elastic force, friction). Such inclusion relation between PPCs gives important clues to identify the dominant principle(s)/law(s) in solving a problem.

Additionally, there are often sets of physical phenomenon constraints of which modeling assumptions (preconditions) are mutually exclusive. These PPCs are never simultaneously valid in the same state of the same system. For example, since 'static friction' and 'kinetic friction' have exclusive preconditions (operating range constraints) about a contact surface of two objects, they are never valid simultaneously at the same surface. The first three global PPCs (i.e., dynamic change, steady state and discontinuous change constraints) are exclusive for the same reason. They often entirely change each other when preconditions are changed. For example, suppose a mechanical system is in a steady state by assuming 'friction' which cancels other forces. When the assumption is changed to 'frictionless,' the system can dynamically change. Such exclusiveness between PPCs gives important clues to identify the differences between models/problems (see section 4.2).

4. Explanation Generation

4.1 Framework of model-making process description

In our framework, each principle/law is described as a 'model fragment' (Falkenhainer and Forbus, 1991) which consists of its applicable condition and its consequence(s). Applicable condition is described as a set of modeling assumption constraints, while a consequence is described as a physical phenomenon constraint. A model consists of the union of PPCs given by instantiated model fragments, MACs giving applicable conditions for them, and boundary condition constraints given in a problem. (Note that an instantiated 'model fragment' is distinguished from 'model fragment class' which describes a principle/law itself.) A model-making process (i.e., solution) is described as the procedure in which model fragments are applied (instantiated) in turn to the situations (represented with MACs and BCCs) to yield new consequences (represented with PPCs). (Note that a consequence of a model fragment can be the condition for others.)

Figure 1a and 1b show examples, in which it is explicitly described why/how each principle/law is applied to the given situation. In contrast, usual description of solution focuses on the calculation of the required physical amount from the given ones, while the principles/laws and conditions which justify the calculation are attached in the ad hoc way. SOC enables implicit assumptions and physical meanings of calculation to be systematically described.

Additionally, a pair/set of model fragment classes which have similar conditions (situations) but have exclusive MAC(s)/PPC(s) as applicable condition(s)/consequence(s) is called 'exclusive model fragment classes.' Grouping such model fragment classes helps the comparison of models.

4.2 Procedure

4.2.1 Explanation of the model-making process (solution)

The description of model-making process mentioned above makes it possible to explain why/how each principle/law is applied explicitly referring to its modeling assumptions. In figure 1b, for example, the formula '$v^2 - v_0^2 = 2ax$' (dynamic change constraint) is used which relates an object's displacement, velocity and acceleration in a time interval. Note that the constraint *'acceleration is constant in the interval'* (operating range constraint) is explicitly described which is an important precondition for this model fragment to be applied. Many students wrongly use this formula when an object's acceleration temporally varies. The explanation explicitly referring to modeling assumptions would be helpful in avoiding such mistakes.

Additionally, in solving problems, it is important to recognize not only each local principle/law and its consequence, but also the global principle/law which dominates the behavior of the whole system. The solution of domain experts is often 'dominant-principle/law-driven,' that is, they first recognize the dominant principle/law of a problem, then apply local principles/laws to 'fill in the slots' of the global principle/law (Chi, Feltovich and Glaser, 1981; Larkin, 1983; VanLehn and van de Sande, 2009). In our framework, a model fragment of global physical phenomenon constraint (PPC) are defined as the aggregation of the model fragments of local PPC which compose the global one (the applicable condition of a global model fragment is the union of its component model fragments). Global model fragments make it possible to explain the model-making process (solution) focusing on the dominant principle/law. For example, in figure 1a, the model fragment 'balance-of-forces' gives a steady state constraint (global PPC) and its applicable condition includes some physical device constraints (local constraints) given by other model fragments. Based on such inclusion relation, the sequence of explanation can be controlled as follows: first, to indicate the given condition 'a block is at rest' (which means its velocity doesn't temporally vary) suggests 'balance of forces' should be used, then to refer to the laws 'gravity' and 'static friction' which influence the *driving power* of the block's velocity. The generated explanation is shown in figure 2a.

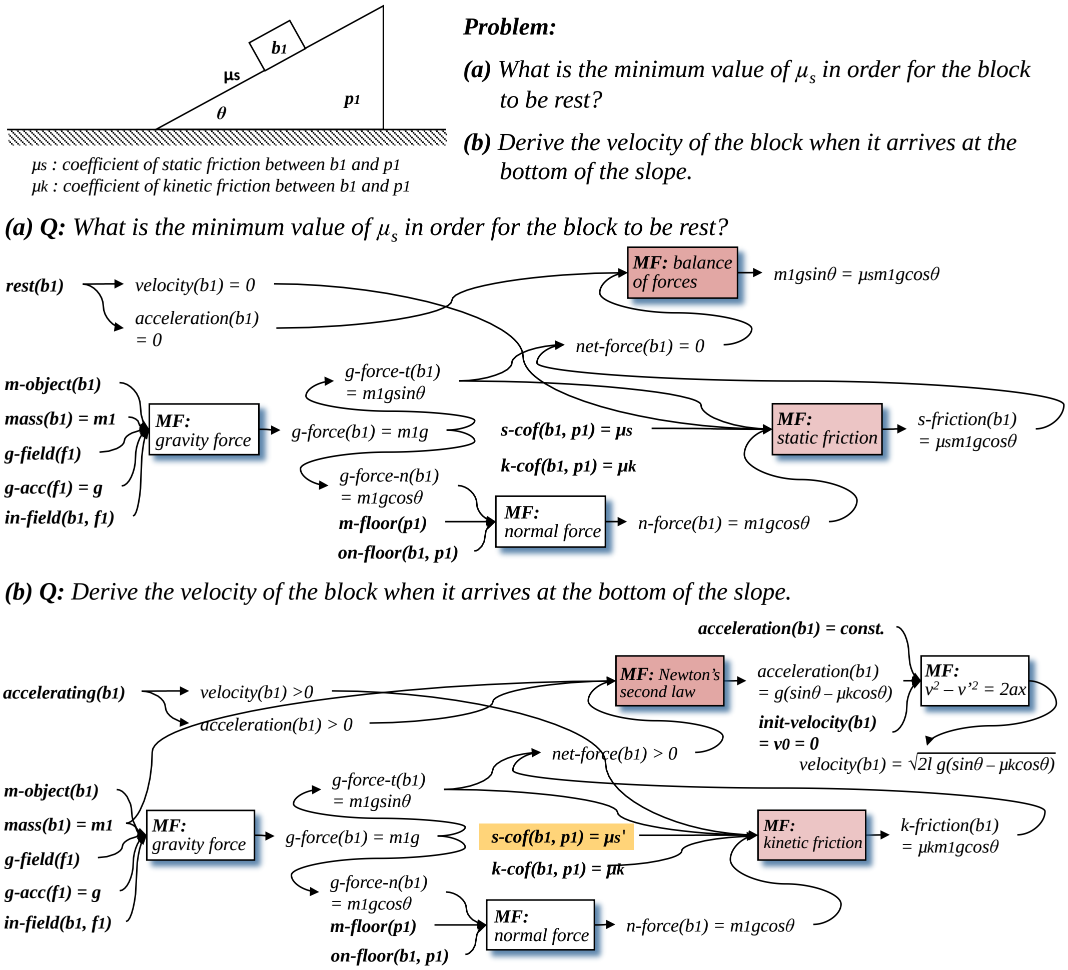


Figure 1. Examples of model-making process.

(a) Since an object keeps at rest, the principle 'balance of forces' is used. So, all the forces applied to the object are required (in the parallel direction to the include plane). Gravity mg is applied downwards and its component parallel to the plane is *mgsinθ*. Since the plane's surface is rough, static friction *μsmgcosθ* is applied. These forces balances with each other, that is, *mgsinθ* = *μsmgcosθ*. From this equation, *μs* must be greater than or equal to *tanθ*. (partly omitted)

(b) Since an object moves at uniform acceleration, the principle 'Newton's second law' is used. So, all the forces applied to the object are required (in the parallel direction to the include plane). Gravity mg is applied downwards and its component parallel to the plane is *mgsinθ*. Since the plane's surface is rough, kinetic friction *μkmgcosθ* is applied. These forces influences the acceleration of the object, that is, *ma* = *mgsinθ* - *μkmgcosθ*. From this equation, *a* = *g(sinθ - μkcosθ)*. Since the acceleration is constant, the formula 'v^2 - v'^2 = *2ax*' is applicable. Consequently, the velocity of the object at the bottom of the slope is $\sqrt{2lg(\sin\theta - \mu k\cos\theta)}$. (partly omitted)

Figure 2. Examples of generated explanation.

4.2.2 Explanation of the difference between models (problems)

The difference between models (problems) can be inferred by comparing their model fragments. There are two types of relations between problems: (1) the problems which have the same/similar surface structures (situations) but have different physical structures (instantiated model fragments belong to different classes) and (2) those which have different surface structures (situations) but have the same/similar physical structures (instantiated model fragments belong to the same classes). Both relations play an important role for promoting conceptual understanding (Scheiter and Gerjets, 2002; Scheiter and Gerjets, 2003). As for the latter, the difference is easily inferred by identifying the corresponding pair of model fragments (each of which belongs to each model) both of which give the (global) PPCs of the same class. The difference can be explained by showing their preconditions (situations) are different.

As for the former, the difference is inferred by identifying the corresponding pair of model fragments (each of which belongs to each model) which belong to exclusive model fragment classes. Since their situations are similar but their modeling assumption constraint(s) and physical phenomenon constraint(s) are exclusively different, they indicate the difference of two models before/after the change of the situation. The type of the difference can be explained by referring to their modeling assumption classes. For example, when two corresponding model fragments have the same physical structure constraints and exclusively different operating range constraints, it is inferred that the difference of two models is change of the operating range about the partial system they match. The model fragments 'static-friction' in figure 1a and 'kinetic-friction' in 1b are in such relation. It can be inferred that the local constraint between a block and slope is changed from 'static-friction' to 'kinetic-friction' by changing the operating range, by which the global constraint 'balance of forces' (steady state constraint) is changed to 'Newton's second law' (dynamic change constraint) (the generated explanation is shown in figure 2b).

Additionally, when comparing models (problems), it is important to recognize not only the change of each local principle/law and its consequence, but also the change of the global principle/law which dominates the behavior of the whole system. Global model fragments which aggregate the model fragments of local PPCs, make it possible to explain the behavioral change of the whole system focusing on the dominant principle/law.

5. Preliminary Experiment

5.1 Design

We conducted an experiment to evaluate the usefulness of our framework. A SOC-based explanation generator was implemented. The purpose was to examine whether the SOC-based explanation promotes students' conceptual understanding, that is, whether their representation of problems was improved and they became able to solve various types of problems by using correct models.

Subjects: Fifteen graduates and under graduates whose majors are engineering participated in.

Instruments: (1) *Two sets of problems in elementary mechanics*: They were called 'problem set 1 (PS-1)' and 'problem set 2 (PS-2).' Each set included fifteen problems of various surface/physical structures. Problems might have similar situations but different solutions, or have different situations but similar solutions. The sets had no common problem. (2) *Usual explanation about the solutions of eleven problems in PS-1*: The calculation of the required physical amount from the given ones was mainly explained. (3) *SOC-based explanation about the solutions of the same problems as usual explanation*: In addition to the solution of each problem, the differences between problems were explained about eight pairs of problems which had similar surface/physical structures. (4) *Explanation generator used for generating SOC-based explanation*: Model-making processes described by the experimenter (first author) were input and their explanations were output, which were rewritten into readable natural language by the experimenter (without changing the point).

Procedure: First, subjects were given PS-1 and asked to group the problems into some categories based on some kind of 'similarity' they suppose (any number/size of categories were allowed), then asked to label each category they made (called 'categorization task 1'). After that, they were asked to solve eight problems in PS-1 (called 'pre-test'). After a week, the subjects were divided into two groups: one was the 'control group' (seven subjects) and another was the 'experimental group' (eight subjects). The average scores of both groups in pre-test were made equivalent. The subjects in control group were given the usual explanation and asked to learn it. The subjects in experimental group were given the SOC-based explanation and asked to learn it. After that, by using PS-2, 'categorization task 2' was conducted in the same way as above. Finally, subjects were asked to solve eight problems in PS-2 (called 'post-test').

Measure: The quality of the representation of problems was measured with the categories, their 'frequencies' (number of problems accounted for) and the time required in each categorization task. The ability to solve various types of problems was measured with the scores in each test. The effect of learning with usual/SOC-based explanation on the quality of representation and the ability of problem-solving was measured with the comparison of the results of two categorization tasks and pre-/post-tests. The superiority of SOC-based explanation to usual explanation was measured with the differences of improvement of categorization and problem-solving between experimental and control groups.

5.2 Results

The categories made by subjects and their frequencies in categorization task 1 are shown in table 1. Most of the subjects categorized the problems based on the similarity of their superficial features, such as the components of the system (e.g., inclined plane, springs), the figures of motion (e.g., circular motion, free fall). Additionally, all subjects finished the task within ten minutes. The results of categorization task 2 are shown in table 2 (for control group) and table 3 (for experimental group). Many subjects of control group still categorized the problems based on the similarity of their superficial features, while many subjects of experimental group became to categorize the problems based on the similarity of their structural features, that is, the dominant principles/laws of problems (e.g., Newton's second law, balance of forces, conservation of energy). Additionally, all subjects of control group finished the task within ten minutes again, while the subjects of experimental group required from twenty-five to thirty-five minutes. These results suggest that the learning with SOC-based explanation promoted representing problems based on their structural features rather than their superficial features (the increase of the time required suggests the subjects of experimental group inferred the physical structure from surface structure).

The average scores in pre- and post-tests are shown in figure 3 (in both tests, full marks were 52). In pre-test, there was no significant difference of average scores between groups (control group: 36.0 and experimental group: 33.6, t-test $p > .10$). In post-test, though there was also no significant difference of average scores between groups (control group: 42.7 and experimental group: 47.6, t-test: $p > .10$), the increase of average score of experimental group was larger than that of control group. This result suggests that the learning with SOC-based explanation promoted the ability to solve various types of problems, that is, to make appropriate models regardless of their superficial features.

These results suggest that SOC-based explanation about the solution of problems and their differences can assist students in reaching conceptual understanding.

Table 1: Categories in task-1

	Number of subjects using category labels (N_1=15)	Average size of category (N_2=15)	Number of problems accounted for ($N=N_1 \times N_2$ =225)	Number of problems wrongly accounted for (N^*=225)	Number of problems correctly accounted for (N^C=N-N^*)
Springs	**12**	**3.1**	**37**	**2**	**35**
Free fall etc.	**9**	**4.1**	**37**	**2**	**35**
Collision	**12**	**2.0**	**24**	**0**	**24**
Circular motion	**12**	**1.9**	**23**	**1**	**22**
Acceleration	**3**	**5.7**	**17**	**1**	**16**
Strings	**7**	**2.0**	**14**	**0**	**14**
Inclined planes	**5**	**2.2**	**11**	**0**	**11**
Balance	5	2.4	12	4	8
Object only	1	5.0	5	0	5
Friction	3	1.7	5	0	5
Second law	2	2.5	5	2	3
Pulleys	1	2.0	2	0	2
Balance of energies	1	4.0	4	2	2
Motion of weight	1	2.0	2	0	2

Table 2: Categories in task-2 (usual)

	Number of subjects using category labels (N_1=7)	Average size of category (N_2=15)	Number of problems accounted for ($N=N_1 \times N_2$ =105)	Number of problems wrongly accounted for (N^*=105)	Number of problems correctly accounted for (N^C=N-N^*)
Springs	**4**	**4.5**	**18**	**0**	**18**
Inclined planes	**4**	**3.3**	**13**	**0**	**13**
Balance of forces	**3**	**3.7**	**11**	**0**	**11**
Conservation of energy	3	6.0	18	9	9
Second law	3	3.7	11	2	9
Pulley and string	2	3.5	7	0	7
Circular motion	4	1.5	6	0	6
Pendulum	3	1.7	5	0	5
Simple harmonic motion	2	2.0	4	1	3
Collision	2	1.0	2	0	2

Table 3: Categories in task-2 (SOC)

	Number of subjects using category labels (N_1=8)	Average size of category (N_2=15)	Number of problems accounted for ($N=N_1 \times N_2$ =120)	Number of problems wrongly accounted for (N^*=120)	Number of problems correctly accounted for (N^C=N-N^*)
Balance of forces	**7**	**4.4**	**31**	**5**	**26**
Second law	**7**	**3.6**	**25**	**1**	**24**
Conservation of energy	**8**	**4.1**	**33**	**12**	**21**
Linear accelerated motion	3	3.3	10	2	8
Conservation of momentum	3	1.3	4	1	3
Acceleration	1	3	3	0	3
Springs	1	3	3	0	3
Pulleys	1	3	3	0	3
Simple harmonic motion and period	2	1	2	0	2
String and tension	1	2	2	0	2
Time	1	2	2	0	2

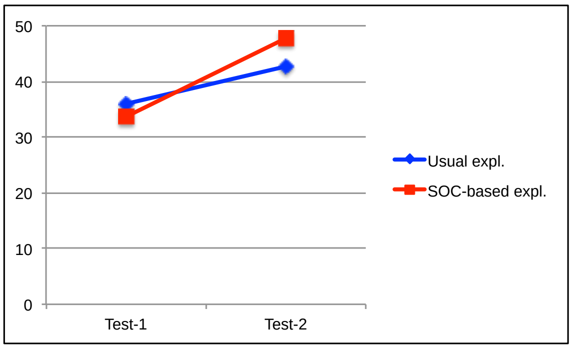


Figure 3. Average scores of tests.

6. Conclusion

Aiming at promoting conceptual understanding through problem practice, we proposed the SOC framework based on which the knowledge necessary for designing a set of problems, sequencing them and generating explanations can be described. We showed the explanations generated with our framework could promote conceptual understanding through a preliminary experiment. SOC-based explanation generator can provide a basic function for designing various instructional methods (e.g., a detailed explanation is gradually simplified (scaffolding-fading), a sequence of problems is given which promotes spontaneous induction). Design of such instructional methods and verification of their effectiveness are our future work.

References

Bransford, J.D., Brown, A.L. & Cocking, R.R. (Eds.) (2000). How People Learn: Brain, Mind, Experience, and School, National Academy Press.

Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices, *Cognitive Science*, 5, 121-152.

Falkenhainer, B. & Forbus, K.D. (1991). Compositional Modeling: Finding the Right Model for the Job, *Artificial Intelligence*, 51, 95-143.

Larkin, J.H. (1981). Enriching formal knowledge: A model for learning to solve textbook physics problems, In Anderson, J.R. (Ed.) Cognitive skills and their acquisition, Hillsdale, N.J.: Lawrence Erlbaum Assocoates.

Larkin, J.H. (1983). The role of problem representation in physics, In Gentner, D. and Stevens, A.L. (Eds.) Mental models, Hillsdale, N.J.: Lawrence Erlbaum Assocoates.

Larkin, J.H. (1985). Understanding, problem representations, and skills in physics, In Chipman, S.F., Segel, J.W. and Glaser, R. (Eds.) Thinking and learning skills, Vol.2: Research and open questions, Hillsdale, N.J.: Lawrence Erlbaum Assocoates.

Scheiter, K. & Gerjets, P. (2002). The Impact of Problem Order: Sequencing Problems as a Strategy for Improving One's Performance, *Proc. of CogSci2002*.

Scheiter, K. & Gerjets, P. (2003). Sequence Effects in Solving Knowledge-Rich Problems:The Ambiguous Role of Surface Similarities, *Proc. of CogSci2003*, 1035-1040.

VanLehn, K. (1998). Analogy Events: How Examples are Used During Problem Solving, *Cognitive Science*, 22(3), 347-388.

VanLehn, K. & van de Sande, B. (2009). Acquiring conceptual expertise from modeling: The case of elementary physics, In Ericsson, K. A. (Ed.) The Development of Professional Performance: Toward Measurement of Expert Performance and Design of Optimal Learning Environments (Cambridge University Press), 356-378.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Practical Use of Interactive Environment for Learning by Problem-posing for One-step Multiplication and Division Word Problems

Sho YAMAMOTO[a*], Yuki AKAO[a], Mitsutaka MUROTSU[a], Takehiro KANBE[a], Yuta YOSHIDA[a], Kazushige MAEDA[b], Yusuke HAYASHI[a] & Tsukasa HIRASHIMA[a]

[a]*Graduate School of Engineering, Hiroshima University, Japan*

[b]*Elementary School Attached to Hiroshima University, Japan*

*sho@lel.hiroshima-u.ac.jp

Abstract: Problem-posing is known that effective method for promoting to master the use of solution methods. However, these methods have been rarely used because of the activity for posing problem and assessment the posed problem. So, we design and develop an interactive environment for learning by problem-posing continually. In previous research, we have already designed and developed the learning environment targeting one-step addition or subtraction word problems and one-step multiplication word problems. In this paper, we have designed and developed the learning environment targeting one-step multiplication or division word problem and its assignment newly. For realizing this system, the property of quantity and its relation is suggested as problem structure. And the diagnosis and feedback of posed problem are defined based on the property and relation. The level of assignment is defined too. Developed learning environment and its practical use in an elementary school are reported.

Keywords: Problem-posing, sentence-integration, multiplication word problems, division word problems, problem structure

1. Introduction

Learning by problem-posing is well known as effective method for promoting learners to master the use of solution methods (Polya, 1945; NF Ellerton, 1986; Silver, CAI, 1996). Moreover, it has been proposed that poor problem solvers often fail to elicit problem structures from problem (Brown, VanLehn, 1980; Mayer, 1982; Kintsch, Greeno, 1985). So, it is postulate that learning by problem-posing is effective for learner to promote to acquire the problem structure. However, this exercise also known that it is difficult to perform because of the cost of activity for posing problem and assessment the posed problem. Therefore, we design and develop the learning environment which learners acquiring the structure of arithmetic word problem by exercising the problem-posing continually (Nakano, et al, 1999; Hirashima, et al, 2007; Hirashima, et al, 2011). This learning environment is required a learner to pose a problem by selecting three cards from a set of given sentence cards and arranging them in proper order (Hirashima, et al, 2014). Also, the learning environment can generate feedback about posed problem. We call this learning environment as "MONSAKUN".

Until now, one-step addition or subtraction word problems and multiplication word problems are analyzed, and the structures of these problems are implemented on tablet PC for realizing the problem-posing exercise and an assessment of posed problem (Yamamoto, et al, 2012; Yamamoto, et al, 2013). MONSAKUN consists of MONSAKUN Touch and MONSKAUN Analyzer. By using this environment, a learner can exercise the problem-posing on MONSAKUN Touch, and a teacher can confirm the result of learner's problem-posing on MONSAKUN Analyzer via network. For this implementation, teacher can use our learning environment in their arithmetic class and lecture the arithmetic word problem by the problem-posing. Actually, in addition to the development of the learning environment, we have performed two experimental uses with elementary school teacher.

First experimental use is, which first grade students were used by MONSAKUN Touch for posing problem that can be solved by one-step addition or subtraction word problems because first grade students have just learned one-step addition or subtraction in arithmetic class. In second experimental use, MONSAKUN Touch 2 is used by the second grade students, which for posing problem that can be solved by one-step multiplication word problems because second grade students have just learned one-step multiplication in arithmetic class. The results of these experimental uses have proposed that not only the learner improve the problem solving performance, but also this learning environment was effective for the learner who can't judge the problem structure to acquire it.

Based on these researches, we have designed and developed a learning environment for posing problems that can be solved by one-step multiplication or division. Since third grade students learn not only one-step multiplication word problems but also one-step division word problems in contrast to second grade students, these scopes are targeted in our research as next step. In order to realize this, the design of a model of problem-posing and an assignment of problem-posing for the learning environment based on analysis of targeted problem should be performed. In this paper, a problem structure is explained in the following chapter. A design of developed learning environment based on this structure is described in section 3. A sequence of assignment is also expressed. Subsequently, a procedure of its practical use and an analysis of the results are reported.

2. Problem Structure of One-step Multiplication or Division Word Problem

In this section, the model of one-step multiplication or division arithmetic word problem is explained. One-step word problems can be expressed by three sentences in our research. Example is shown in Figure 1. Because there are three values in one-step arithmetic word problem, this problem can be expressed by three sentences. These sentences consist of two sentences mean existence and one sentence means relation between other two values. We call each sentence as existence sentence and relation sentence. In this example, "There are three boxes" and "There are several apples" are existence sentence. "There are four apples in each box" is relation sentence because this sentence shows the relation between the apple and box. These sentences consist of value, object and predicate.

In addition to the kind of sentence, multiplication and division word problem have a property of quantity (Yamamoto, et al, 2013). Generally, multiplication is expressed by “multiplicand multiplied by multiplier is product” (Greer, 1992; Vergnaud, 1983; Davies, 1841). In other words, each quantity has different property. In Figure 1, multiplicand is expressed as "There are five apples in each box", multiplier is "There are three boxes" and product is "There are several apples". That problem contains the story that the value of apples is expressed as the amount of apple when there are three boxes and the value of apples in each box is basis. Since, in Japanese Education, multiplicand is also called "base quantity", multiplier is "proportion" and product is "compared quantity". Then, the arithmetic word problems that can be solved by one-step multiplication or division has three types of story. This is, (1) Compared quantity divided by base quantity is proportion, (2) Base quantity multiplied by proportion is compared quantity, (3) Compered quantity divided by proportion is base quantity. The story of the problem in Figure 1 is (2).

All of these stories contain the relation that is “Base quantity multiplied by proportion is compared quantity”. One-step multiplication or division word problems are expressed by changing the one quantity to required value in each story. Therefore, it is important to extract the base quantity, proportion and compared quantity from problem and to make the relation between these quantities with “Base quantity multiplied by proportion is compared quantity”. Next section, the implementation of the problem structure mentioned above to tablet PC is described.

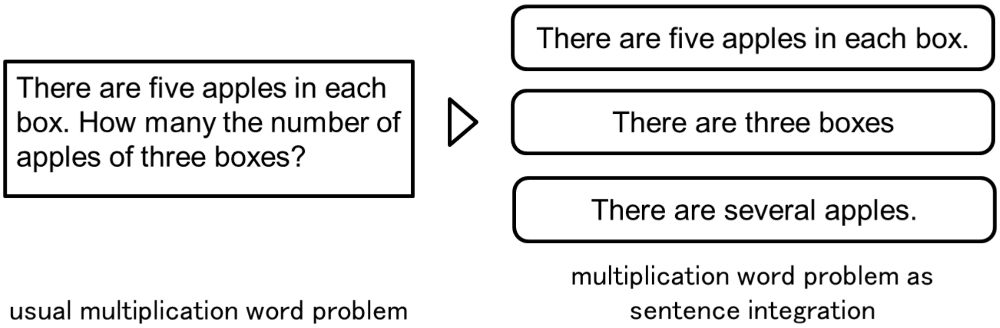


Figure 1. Example of Problem Expression as Sentence Integration.

3. Learning Environment for Problem-posing "MONSAKUN Touch3"

3.1 Framework of Learning Environment

This learning environment consists of MONSAKUN Touch 3 for learners and MONSAKUN Analyzer 3 for teachers. A result of the learner's learning by problem-posing on MONSAKUN Touch 3 is sent to database server via network. This framework is shown in Figure 2. MONSAKUN Touch 3 developed by using Android, MONSAKUN Analyzer 3 by using PHP and JavaScript. The each software can be run on Android Tablet. RDBMS is used MySQL. The teacher can confirm the graph of learner’s learning by using MONSAKUN Analyzer 3 that receives a learning data from database server. The learning data are saved as three data: the number of correct problem, the number of incorrect problem, the number of the each incorrectness and the learner’s log. Category of incorrectness is explained in next section. MONSAKUN Analyzer 3 generates some graphs by using these data and displays teacher it. Teacher can limit to an assignment that learners can exercise on MONSAKUN Touch 3 by using MONSAKUN Analyzer 3.

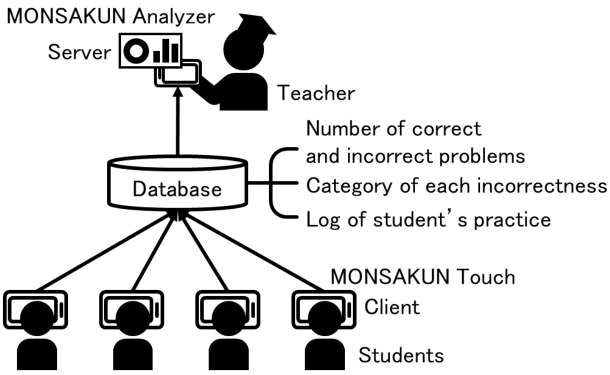


Figure 2. The Framework of Learning Environment.

3.2 MONSAKUN Touch 3

3.2.1 Flow by Using MONSAKUN Touch 3

In this section, the flow of problem-posing by using MONSAKUN Touch 3 is described. This is same as the MONSAKUN Touch 1 and 2 but an interface of problem-posing is a little changed. First, a learner logs in this system by selecting his/her grade, class and number. After that, the learner selects a level of assignment on an interface for selecting level. The learner is also able to select particular assignment and switch feedback on or off. The level of assignment is elaborated in 3.2.3. After selecting the level, he/she sees an interface of problem-posing is shown in Figure 3. This interface presents the assignment for posing problem, the set of given sentence card and three blank for arranging given sentence cards. Sentence cards are written sentence like "There are three apples".

These sentence cards were mentioned in previous section. Hence the learner can pose the problem by selecting three sentence cards from given cards and arranging them in proper order. Given sentence cards are consists of correct card set and dummy card set for leading to errors. In the MONSAKUN Touch 3, the text means the property of quantity is shown in the left side of each blank because the environment lets the learner consider the property of each sentence card when they learn by using this environment. If three blank is filled with three sentence cards, diagnosis button will be active. Then, the learner can tap this button and the system diagnoses and feed back his/her posed problem. When the learner finishes answering all assignment in selected level correctly, the interface for posing problem backs to the interface for selecting level.

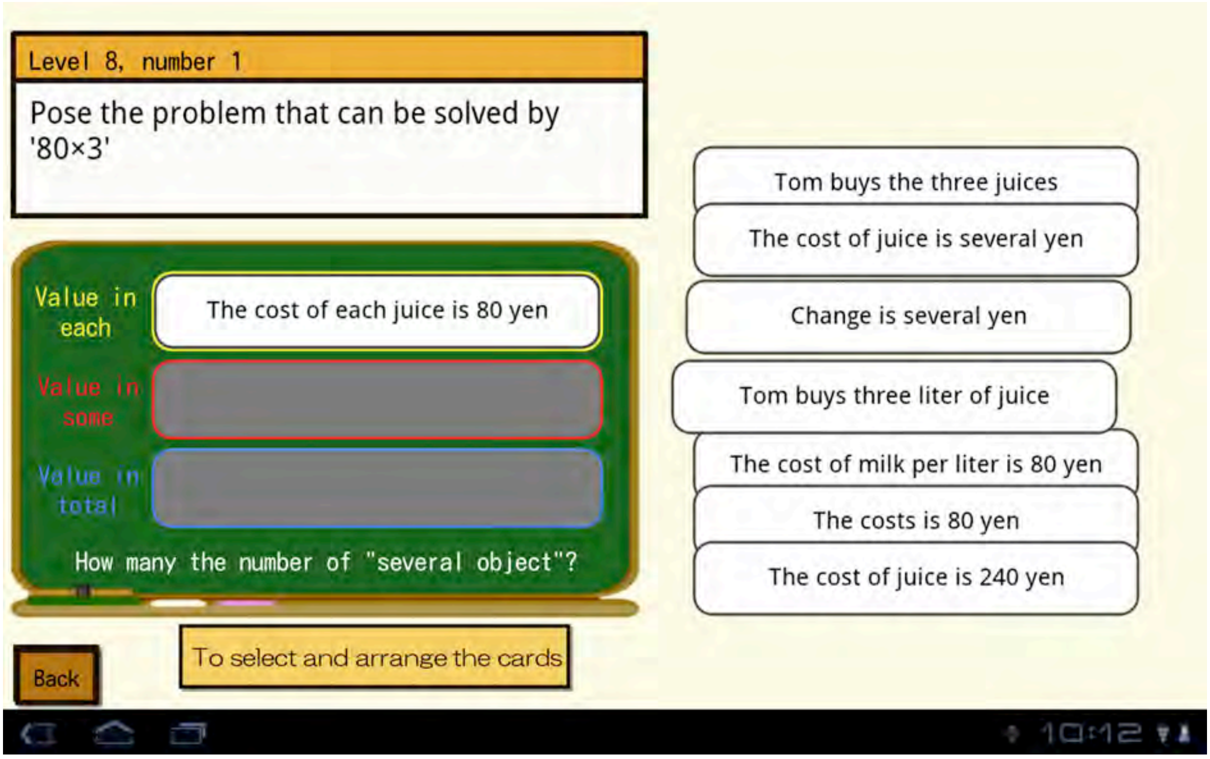


Figure 3. Main Interface of MONSAKUN Touch 3.

3.2.2 Diagnosis and Dummy Card

Figure 4 is the procedure for diagnosing of posed problem. This procedure is processed based on the problem structure is described in chapter 2. The posed problem is required to satisfy these several constraint. First, MONSAKUN Touch 3 assesses the composition of sentence. This is, the combination of existence sentences and relation sentence, and the setting cards and its property of quantities. If these are not correct, MONSAKUN Touch 3 gives a learner feedback it. If these are correct, next, MONSAKUN Touch 3 assesses whether a composition of story is correct or not. This means that the system checks the relation of each object, value, unit in answered three sentence cards. In addition to this, the learning environment assesses not only the relation between two objects of base quantity and object of other sentence cards but also the base value of base quantity like “apple per 2 boxes”. After that, MONSAKUN Touch 3 assesses whether the calculation expression of posed problem and given calculation by assignment is same or not. If learner causes the error, the system feedbacks its reason based on diagnosis in Figure 4. When posed problem satisfies the above constraint, this problem is diagnosed as correct.

Dummy card is included in given sentence card in order to let the learner cause the error that is show in Figure 4. These cards make by changing object, number or predicate of correct card. Because of dummy cards, the learner needs to be vividly aware of problem structure when they posed problem on MONSAKUN Touch 3.

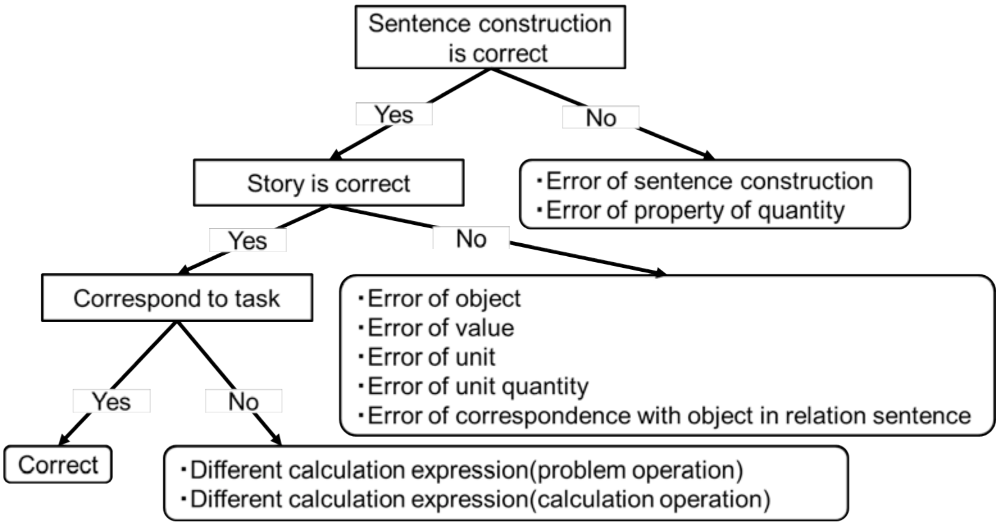


Figure 4. The Procedure for Diagnosing the Posed Problem.

3.2.3 Designing the Level of Assignment

In this research, we have designed the level of assignment gradually so that the learner acquires the problem structure seamlessly. Table 1 shows the all level of assignment by dividing into the number of level, assignment, required activity, contents of assignment and number of assignment. Then, each level designed on the basis of "Base quantity multiplied by proportion is compared quantity" for third grade students on elementary school. The learner is required to pose the story from level 2 to 7, to pose the problem from level 8 to 9.

In level 1, the learner is given the story of one-step multiplication and four calculation expression which are expressed by "Base quantity multiplied by proportion is compared quantity", "Proportion multiplied by base quantity is compared quantity" and the cumulation of same number like “4+4+4=12” and “3+3+3+3=12”. This assignment is the confirmation of the relation between multiplication and addition. Then, the learner is required to select the correct calculation expression. The purpose of this level is which let the learner comprehend the relation of multiplication story and addition calculation. The learning environment gives the story and several sentence cards to the learner in level 2. The given story as sentence integration consists of two fixed sentence cards and one blank. The learner is required to fill this blank by considering the property of quantity. In this assignment, they learn the property of quantity that is contained each given sentence card. Given sentence cards in level 3 are included two sentence cards that have different text representation and same property. For example, "There are two boxes." and "The number of box is two.". In this level, let the learner learn that the sentence cards include the same property of quantity have various text representation. MONSAKUN Touch 3 presents the three blank for putting the sentence cards and several sentence cards in level 4. Then, the learner is required to pose the story by selecting three sentence cards and by arranging them in proper order based on the relation of "Base quantity multiplied by proportion is compared quantity". The assignment of level 5 requires the learner to pose the two stories by using one common sentence card. For example, "There are three apples in each box. There are six boxes. There are eighteen apples." and "There are two boxes in each shelf. There are three shelves. There are six boxes" are used same card that is "There are six boxes.". Through this exercise, the learner comprehends that existence sentence card is able to have two property of quantity. In other words, both proportion and compared quantity are expressed by existence sentence. After that, in level 6, the learner learns that the story has three kinds of calculations expression that are mentioned in section 2. This purpose is that the learner notices the multiplication story contains the calculation (a) and (c). Thus, the learner is given the multiplication and division calculation expression as assignment for posing story. In order to let the learner confirm three properties of quantity and its relation again, assignments of level 7 includes improper assignment which cannot solve because of lack of one proper sentence card. Then, the learner is given a specific sentence card for posing the story in this level, which is labeled "proper sentence card is not given" instead of lacking sentence card. Because the assignments in level 7 are composed of usual assignment and assignment which mentioned above, the learner is required to consider each property of quantity and its relation again. As the next step, the learner is required to pose problem in level 8 because the learner learn to pose the story through level 2 to 7. Finally, in level 9, the learner is required to pose the two problems by using one common sentence card. This assignment is same as assignment of level 5. Through the exercise from level 1 to 9, the learner can acquire the problem structure gradually.

Table 1: The Assignment Level on MONSAKUN Touch 3.

Level	Required activity	Contents of assignment	Number of assignment
1	Select calculation expression	Select calculation express given story	12
2	Pose story	Pose story that is expressed by given calculation (one-step multiplication) Required story has already given two sentence cards	12
3	Pose story	Same as assignment of level 2 Include same property and different text representation	12
4	Pose story	Pose story that is expressed by given calculation (one-step multiplication) Select three sentence cards and arrange them	10
5	Pose story	Pose two stories by using same sentence card	10
6	Pose story	Pose story that is expressed by given calculation But given calculation expression is one-step multiplication or division	12
7	Pose story	Same as assignment of level 6 But one proper sentence card is not given	12
8	Pose problem	Pose problem that is expressed by given calculation Select three sentence cards and arrange them	12
9	Pose problem	Pose two problems by using same sentence card	12

3.3 MONSAKUN Analyzer 3

Here, MONSAKUN Analyzer 3 for visualizing the learner's learning data on MONSAKUN Touch 3 is explained in line with the function of MONSAKUN Analyzer 3. After the teacher loged in the learning environment by inputting id and password, MONSAKUN Analyzer 3 changes the interface that is shown in Figure 5. In this interface, the learning environment displays the average of student's learning data in each lesson that are received some learning data from database server. This interface generates and shows a three bar charts and a doughnut chart. Three bar charts consist of the average number of posed problem in total, the average number of correct problem and the average number of incorrect problem. A doughnut chart shows the rate of each error that category is classified in 3.2.2. In addition to this information, MONSAKUN Analyzer 3 indicates the average progress of the level and assignment number in class. These displayed learning data can be filtered out based on the each level and assignment in each class. MONSAKUN Analyzer 3 displays these graphs not only in each lesson but also in each student by clicking the link "see the each student's data". The interface element is same as Figure 4 but each graph are visualized in each student's learning data in total. This interface also only displays the learning data in each level or each assignment. Then, the teacher can see the posed problem of each student by clicking "see the each student's log". These data are updated in real-time. The teacher is able to arrange his/her lesson on the basis of these visualized data. For example, by seeing the learning data in each student, the teacher can know the students who failure in the exercise of MONSAKUN Touch 3 and support their exercise.

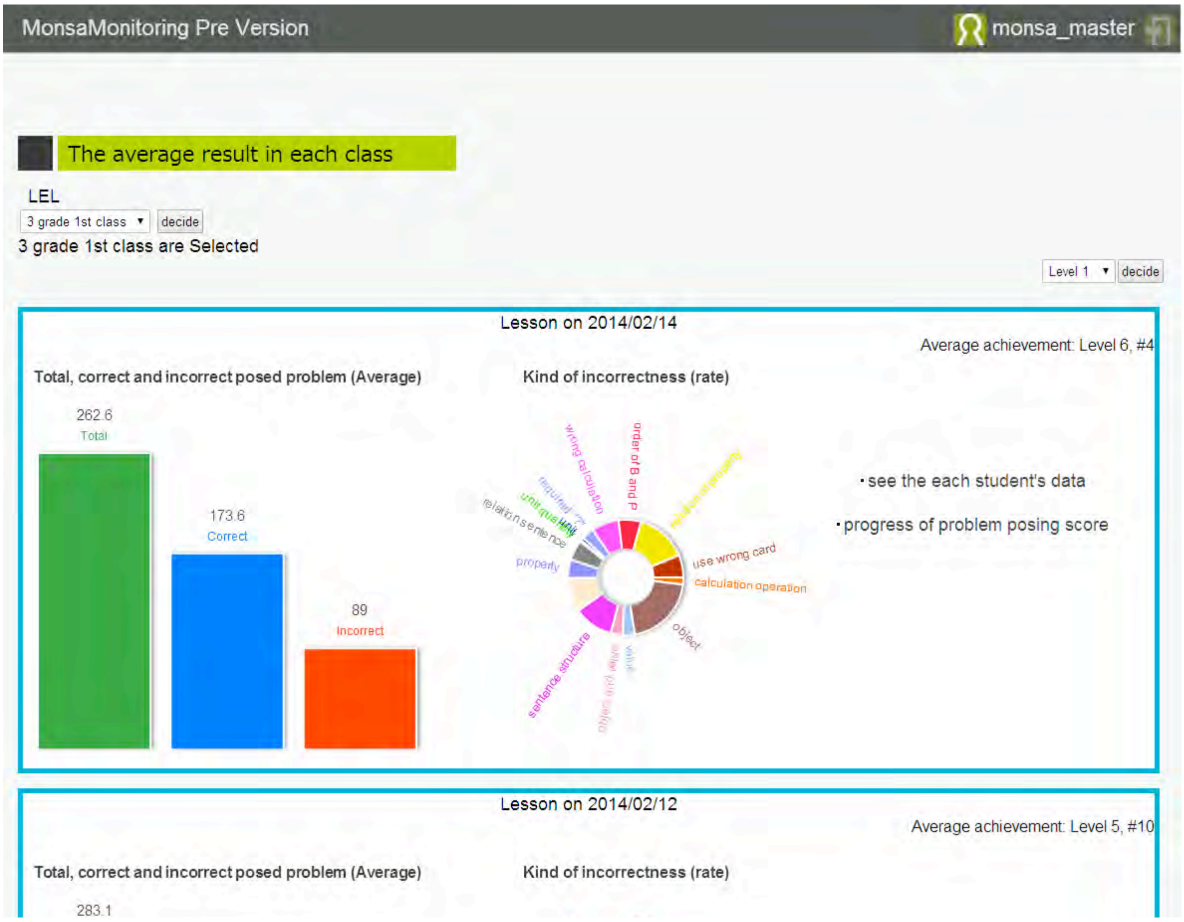


Figure 5. The Main Interface of MONSAKUN Analyzer 3.

4. Experimental Use of MONSAKUN Touch 3

4.1 Procedure of Experimental Use

Subjects were thirty-nine students in the third grade of an elementary school. They were divided into subjects who experienced MONSAKUN and who did not experience it in previous experimental use of our research (Yamamoto, et al, 2012; Yamamoto, et al, 2013). Inexperienced group of MONSAKUN has learned one-step addition, subtraction or multiplication word problem by usual lesson only. Moreover, they had just learned to solve arithmetic word problems that can be solved by one-step multiplication or division. This experimental use has been performed during thirteen lessons that consist of pretest in one lesson, eleven lessons by using MONSAKUN and posttest in one lesson (45 minutes per lesson, in 5 weeks). A lesson by using MONSAKUN has composed of teaching about problem-posing by a teacher and problem-posing exercise by using MONSAKUN Touch 3. The teacher decided the time of using MONSAKUN Touch 3 based on the progress of each lesson. If the subjects have finished twice the current level when they exercise the problem-posing after teaching, they were allowed to work on the previous level. The purpose of this experimental use is to examine the effects of MONSAKUN Touch 3 by using a usual problem solving test, an extraneous problem solving test and a problem-posing test, and the effects of experience MONSAKUN is also examined.

We used these three tests: the problem solving test, the extraneous problem solving test and the problem-posing test. The problem solving test is the usual problem solving test can be solved by one-step multiplication or division that is expressed by three sentences. This test is included five stories that are made as the permutation of base quantity, proportion and compared quantity without "Proportion multiplied by base quantity is compared quantity" because of commutative law. Therefore, the usual problem-posing test has fifteen questions because each quantity can be the required value in these five stories. Extraneous problem solving test includes extraneous information that is not necessary to solve the problem. It is more difficult for learner to solve the extraneous problem than to solve the usual problem solving (Muth, 1992). The subjects are required to judge the relevance of each sentence and find the sentence including as the extraneous information for solving the problem. Therefore, the extraneous problem solving test is useful to assess leaner's comprehension of the problem structure. These problems consists of twelve problems that including the two kinds of extraneous information that change sentence cards except sentence contains required value in each six stories. The problem-posing test examines the problem-posing performance to let the subject pose the problem as he/she can within the time limit. The subject pose problem from scratch. The time limit is ten minutes in each test. The difference between pretest and posttest is order of each problem.

4.2 Analysis of Pretest and Posttest

An analysis of pretest and posttest are reported in this section. And the level by using lecture is described. The teacher has performed the lecture based on the level on MONSAKUN Touch 3 and treated one level in one lecture. However, the subjects can not relate between multiplication calculation expression and text representation contain "cut" because "cut" is associated with division calculation expression. Thus, the teacher has to spend three lessons for resolving this difficulty. The lessons have been performed from sixth lesson for level 4 continually and the subjects have worked on level 9 in eleventh lesson.

The result of average score and SD in usual problem solving, extraneous problem solving and problem-posing test are shown in Table 2. These scores are divided into experienced and inexperienced group of MONSAKUN that the subjects learn by problem-posing in the scope of one-step addition, subtraction or multiplication word problem. In addition to this result, the results of ANOVA in each test are shown in Table 3. There was an interaction in the score of usual problem-posing test between experience of MONSAKUN and pre-posttest (p=.03). So, we analyzed simple effect. There was a significant difference in the score of posttest between experienced group and inexperienced group ($F(1, 36)$=3.193, p=.008). This result suggested that it is effective for the subjects to experience the learning by using MONSAKUN for improving their usual problem solving performance. Next, there was a significant difference in the score of extraneous problem solving test between experienced group and inexperienced group (p=.04), and effect size is medium ($|\eta^2|$=.10). Also, there was a significant difference in the score between pretest and posttest (p=.02), and effect size is small ($|\eta^2|$=.02). In addition to this analysis in the score of extraneous problem solving test, we analyzed the correlation between the pretest score and the difference posttest and posttest score. In this result, there are a negative correlation between them (Spearman's rank-correlation coefficient, $|rs|$=.59, p=2.5E-06). These results suggested that the lesson by using our learning environment promote the subjects to improve their problem structure, in particular, more effective to the subjects who the score of extraneous problem solving test is lower. MONSAKUN is more effective for the subjects who have experienced MONSAKUN to comprehend the problem structure particularly. This result same as the result of experimental use by MONSAKUN Touch 2 (Yamamoto, et al, 2013) so this effect is that the main effect of learning by problem-posing on MONSAKUN Touch. Last, there was no significant difference in the number of posed problem between experienced and inexperienced group. But, there was a significant difference between pretest and posttest (p=.005), and effect size is medium ($|\eta^2|$=.07). These results suggested that MONSAKUN is useful for the subjects to improve their problem-posing performance regardless of whether the subjects have experienced MONSAKUN.

Table 2: Result of Each Pretest and Posttest in experienced group (N=18) and inexperienced (N=20).

Test	Experience of MONSAKUN	Pretest		Posttest	
		M	SD	M	SD
Problem-posing	experienced	2.50	1.57	3.56	1.42
	inexperienced	2.10	1.45	2.7	1.52
Usual Problem solving	experienced	13.61	1.34	14.28	0.80
	inexperienced	13.30	1.52	12.75	2.05
Extraneous Problem Solving	experienced	10.83	2.14	11.38	0.76
	inexperienced	8.80	3.54	9.75	3.40

Table 3: Two factor ANOVA of Each Pretest and Posttest.

(a) Result of the score of usual problem solving test

factor	SS	df	MS	F	
experienced × inexperienced group	16.02	1	16.01	4.56	*
pre × post-test	0.06	1	0.06	0.05	n.s.
interaction	7.01	1	7.01	5.32	*
total variation	196.88	75			

(b) Result of the score of extraneous problem solving test

factor	SS	df	MS	F	
experienced × inexperienced group	63.88	1	63.88	4.54	*
pre × post-test	10.74	1	10.74	5.63	*
interaction	0.74	1	0.74	0.39	n.s.
total variation	651.41	75			

(c) Result of the score of problem-posing test

factor	SS	df	MS	F	
experienced × inexperienced group	7.47	1	7.47	2.28	n.s.
pre × post-test	12.98	1	12.98	9.19	**
interaction	0.98	1	0.98	0.70	n.s.
total variation	190.04	75			

* $p<.05$, ** $p<.01$

5. Conclusion

In this paper, we have described the model of problem and problem-posing in one-step multiplication or division arithmetic word problem, the development of the interactive environment for problem-posing based on the model, and the results of its practical use. We analyze the problem structure as processing information and realize the development of interactive environment for learning by problem-posing based on its problem structure continually. This learning environment called MONSAKUN Touch. Until now, MONSAKUN Touch is introduced by elementary school teacher into first and second grade arithmetic class on an elementary school for practical use. MONSAKUN Touch is developed for learning by problem-posing in the scope of one-step addition, subtraction or multiplication arithmetic word problem. Therefore, as the next step, we designed and developed the learning environment by posing problem that can be solved by one-step multiplication or division word problem. In order to realize this system, firstly, we have mentioned that three quantities and its relation define one-step multiplication or division word problem. These three quantities are called base quantity, proportion and compared quantity. Its relations are "Base quantity multiplied by proportion is compared quantity". At the second step, the problem-posing based on this problem structure and the diagnosis and feedback of posed problem are defined. The levels of assignment are designed by this problem structure as the learner can acquire the problem structure gradually. After that, we have developed learning environment for problem-posing as sentence integration. This environment consists of MONSAKUN Touch 3 and MONSAKUN Analyzer 3. The learner can exercise the problem-posing on MONSAKUN Touch 3 and MONSAKUN Analyzer 3 provides the visualized student's learning data on MONSAKUN Touch 3 to the teacher. Lastly, an eleven lesson experimental use is reported. The results of brief analysis suggested that the third grade students who have learned by using MONSAKUN in the past time are improved their problem solving performance and sophisticated their acquired problem structure. In addition to this result, the third grade students who didn't acquire the problem structure well are improved their problem-posing and sophisticated their acquired problem structure. This result same as the result of experimental use by MONSAKUN Touch 2 so it is suggested that this result is main effect of learning by problem-posing on MONSAKUN Touch.

As our future works, we need to verify the quantity of the effect to high group by using MONSAKUN Touch 3. Furthermore, we should perform the practical use for problem-posing in one-step addition, subtraction, multiplication or division to fourth grade students of an elementary school continuously.

References

Brown, J.S., VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. Cognitive Science, 4, 379-426.

Davies, C. (1841). Arithmetic: Designed for Academies and Schools. A.S. Barners.

E.A. Silver, J. CAI. (1996). An analysis of arithmetic problem posing by middle school students. Journal for Research in Mathematics Education, 27(5), 521-539.

Greer, B. (1992). Multiplication and Division as Models of Situations. Handbook of Research on Mathematics Teaching and Learning, 276-295.

Hirashima, T., Kurayama, M. (2011). Learning by Problem-Posing for Reverse-Thinking Problems, Proc. of AIED 2011, 123-130.

Hirashima, T., Yamamoto, S., & Hayashi, Y. (2014). Triplet Structure Model of Arithmetical Word Problems for Learning by Problem-posing. In Human Interface and the Management of Information. Information and Knowledge in Applications and Services (pp. 42-50). Springer International Publishing.

Hirashima, T., Yokoyama, T. Okamoto, M. ,Takeuchi, A. (2007). Learning by Problem-posing as Sentence-Integration and Experimental Use. Proc. of AIED2007, 254-261.

Kintsch, W., Greeno, J.G. (1985). Understanding and Solving Word Arithmetic Problem. Psychological Review, 92(1), 109-129.

Mayer, R. E. (1982). Memory for algebra story problem. Journal of Educational Psychology, 74, 199-216.

Muth K.D. (1992). Extraneous information and extra steps in arithmetic word problems, Contemporary educational psychology. Vol. 17, pp.278-285.

Nakano, A., Hirashima, T., Takeuchi, A. (1999). Problem-Making Practice to Master Solution-Methods in Intelligent Learning Environment. Proc. of ICCE'99, 891-898.

NF Ellerton. (1986). Children's made-up mathematics problems: a new perspective on talented mathematicians. Educational Studies in Mathematics, Vol.17, 261-271.

Polya, G. (1945). How to Solve It. Princeton University Press.

Vergnaud, G. (1983). Multiplicative Structures. Acquisition of mathematics concepts and processes, 127-174.

Yamamoto, S., Kanbe, T., Yoshida, Y., Maeda, K., Hirashima, T. (2012). A Case Study of Learning by Problem-Posing in Introductory Phase of Arithmetic Word Problems. Proc. of ICCE2012, Main Conference E-Book, 25-32.

Yamamoto, S., Takuya H., Kanbe, T., Yoshida, Y., Maeda, K., Hirashima, T. (2013). Interactive Environment for Learning by Problem-Posing of Arithmetic Word Problems Solved by One-step Multiplication, Proc. of ICCE2013, Main Conference E-Book, 51-60.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Revealing Students' Thinking Process in Problem-Posing Exercises: Analysis of First Sentence Selection

Nur HASANAH[a*], Yusuke HAYASHI[a] & Tsukasa HIRASHIMA[a]
[a]*Dept. of Information Engineering, Hiroshima University, Japan*
*nur@lel.hiroshima-u.ac.jp

Abstract: We have developed a computer-based learning environment called MONSAKUN to realize learning by problem-posing, where students select and arrange several sentence cards to pose arithmetical word problems. We call this type of problem-posing “sentence integration”. As the next step of MONSAKUN development, we have been analyzing the sentence selection process which is considered to reflect students' thinking process. In the first step of analysis, we focused on the first sentence card selected in the process of posing a problem. We found that the selection changed based on different type of approach, type of story and students’ exercise experience. This result is an important step towards building elaborate process model of problem-posing and adaptive support of the process.

Keywords: Arithmetical word problems, learning analytics, problem posing, reverse thinking problem, sentence integration

1. Introduction

1.1 Background

Two activities that have been identified to be central themes in mathematics education are problem posing and problem solving. Problem solving practice, as the most popular way of teaching the solution method, has been long integrated into school mathematics (Stanic & Kilpatrick, 1988). Problem posing practice involves the generation of new problems in addition to solving pre-formulated problems (English, 1997; Silver & Cai, 1996). Although learning by problem posing has been suggested as an important way to promote learner's understanding (Ellerton, 1986; Polya, 1957), it was not until recently that the recommendations for the reform in mathematics education suggested the problem posing inclusion in students' activities (NCTM, 2000). Several investigations of various aspects of problem posing activities have been conducted as more educators and researchers realized its importance in mathematics education (English, 1998; English, 2003).

One of the most important issues in learning by problem posing is the way to assess and give feedback to posed problems. In traditional problem posing method, teachers and students were faced with difficulties to conduct the learning activities effectively. It is not easy for students to pose mathematically correct problems in a given time, especially students in lower grade of elementary school. Teachers were having problems to assess and give feedback to the wide variation of problems that students pose in a given time of class activity. The inefficiency of time and available method made problem posing activity less attractive for most mathematics educators.

In order to realize learning by problem-posing in a practical way, we have been investigating a computer-based learning environment to assess and give feedback to problems posed by students (Nakano, Hirashima & Takeuchi, 1999; Hirashima, Nakano & Takeuchi, 2000; Nakano, Hirashima & Takeuchi, 2002; Hirashima et al, 2007; Kurayama & Hirashima, 2010). The software, named MONSAKUN (means “Problem-posing Boy” in Japanese), provides an interactive support for learning arithmetical word problems solved by one operation of addition/subtraction.

The interface of an assignment in MONSAKUN is explained in Figure 1. A learner is provided with a set of sentence cards and a numerical expression, and then he/she is required to pose an arithmetical word problem using the numerical expression by selecting and arranging appropriate cards. Although learners do not create their own problem statements, they are required to interpret the provided sentences and integrate them into one problem, which is essentially the same as ordinary problem-posing activity. Hirashima & Kurayama (2011) call this style as “problem-posing as sentence-integration” and assert that this integration process is an essential activity in learning.

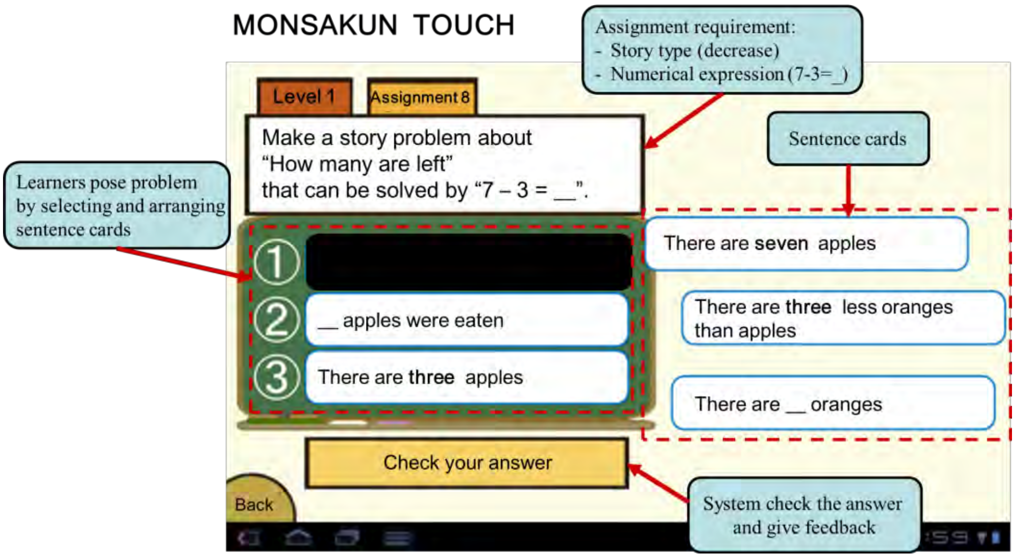


Figure 1. Interface of MONSAKUN

The practical use of MONSAKUN at several elementary schools has been reported in previous studies (Hirashima et al., 2008; Kurayama & Hirashima, 2010; Hirashima & Kurayama, 2011; Yamamoto et al., 2013). Through the analysis of pre-test and post-test of high-score group and low-score group of the students, effect of learning by problem-posing with MONSAKUN was investigated. It has been confirmed that the problem-posing exercise is effective to improve both problem-posing and problem-categorization abilities. Furthermore, after long term use of MONSAKUN in an elementary school, the result showed that both the students and teachers enjoyed using this system continuously and considered it useful for learning.

1.2 Purpose

One important direction in investigation of problem posing activities is to examine thinking processes related to problem posing (Brown & Walter, 1990). As the next step of MONSAKUN development, the purpose of this study is to examine learners’ problem-posing process and to develop technologies for identifying learners’ thinking process. By identifying learners’ thinking process, we will be able to provide a better individualized feedback based on understanding of each learner.

Through previous practical use, we observed different ways of sentence selection in problem-posing process by the students. We assume that it is caused by the different way of thinking depending on the nature of problems and learner's understanding. Therefore, by examining the selection process of sentences, we aim to infer about a learner’s thinking process in problem-posing.

While it is difficult to trace thinking process in a free problem posing activity, we can trace learners’ card selection in MONSAKUN which can be considered to reflect their thinking process. Problem posing in MONSAKUN is defined as integration of provided sentences into one problem. Learner’s assignment is to choose appropriate cards from several sentence cards provided by the system in order to fill the requirement of numerical expression and story type. This can be considered as search problem. Figure 2 illustrate a search space of an assignment in MONSAKUN which provides six sentence cards. The search space is a tree structure of combination of cards. Here, the root is the starting point and the numbers represent ID of cards, for example, the starting point is empty and the combination of cards 1, 2 and 3 indicates the correct answer. The nodes and arrows with bold line represents the paths of the learner’s card selections during his problem posing activity. This learner committed mistakes twice and then got the correct answer. As shown in this figure, problem-posing in MONSAKUN is defined as a search problem in the structure of transition of card combinations. The rules for valid combination of the sentence cards are explained in the next section.

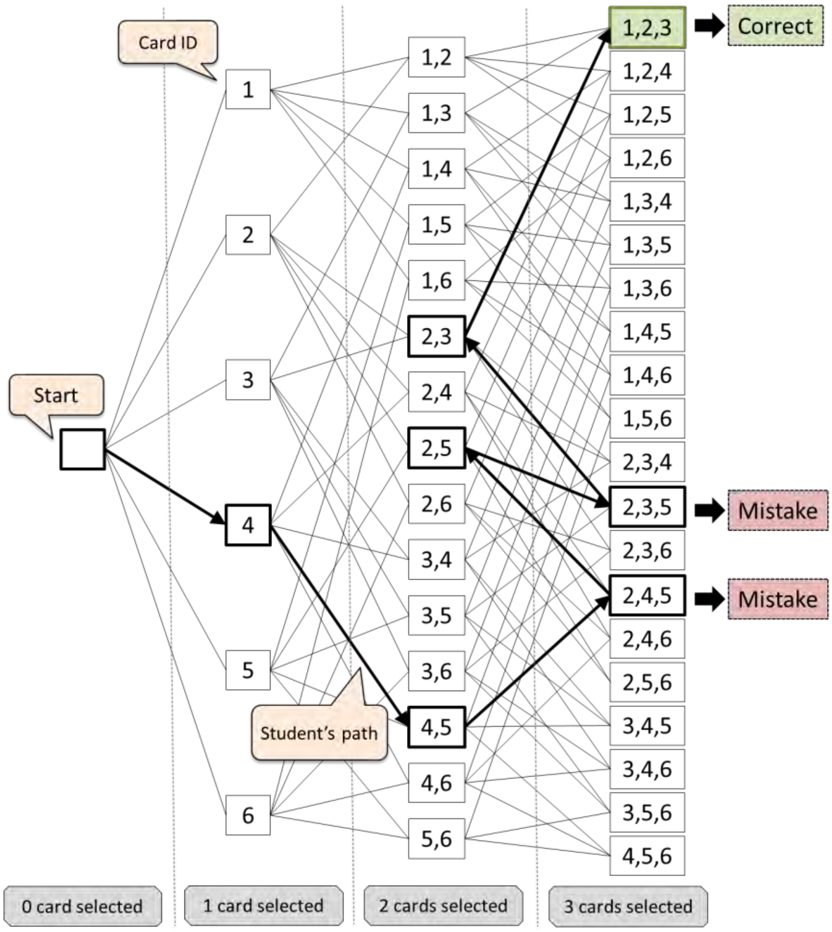


Figure 2. Example of learner's card selections shown in a graph

In this study, we examine how learners pose arithmetical word problems as sentence integration on MONSAKUN. Our assumption is learners do not choose sentence cards randomly - they arrange sentence cards based on some sort of thinking. In the analysis, as the first step toward analyzing problem-posing activity, we especially focus on what kind of sentence card was firstly selected by the learners.

The composition of this paper is as follows. The next section gives an overview of MONSAKUN and the definition of two types of problems: forward-thinking and reverse-thinking problem. Section 3 describes data about initial card selection by learners and discusses what happens in problem posing on MONSAKUN. Finally section 4 concludes this paper and shows some prospects for future study.

2. Problem Posing Activity in MONSAKUN

2.1 Categorization of Problem-posing Exercises

In arithmetical word problems, sentences are divided into two types: existence sentence and relational sentence. An *existence sentence* represents a number of single objects. A *relational sentence* contains keyword that represents a story type. An arithmetic word problem of binary operation is integration of two existence sentences and one relational sentence.

There are four types of story in arithmetic word problems of addition and subtraction: 1) combination, 2) increase, 3) decrease, and 4) comparison (Riley, Greeno and Heller, 1983). In MONSAKUN, the differences among them are defined as differences of integration of sentences. For example, a decrease story type problem is composed as follows:

a) There are seven apples *(existence sentence)*,
b) Several apples were eaten *(relational sentence that contains decrease story type)*, and
c) There are three apples *(existence sentence)*.

2.2 Forward-thinking and Reverse-thinking Problem

An arithmetical word problem includes two kinds of numerical relations: story operation structure and calculation operation structure. Story operation structure is the equation expressing the numerical relation according to the story, while calculation operation structure is the equation used to derive the required number in the assignment.

Based on this relation, there are two groups of problem in arithmetical word problems: forward-thinking problem and reverse-thinking problem. In forward-thinking problem, a story represented in the problem has the same structure with the calculation to derive the answer, while in reverse-thinking problem, the story and the calculation operation structures are different (Hirashima and Kurayama, 2011).

For example, in the following problem:

There are seven apples. Three apples were eaten. There are several apples.
How many apples are there?

Based on the sentence "Three apples were eaten", we understand that the story focuses on "decrease" number of an object. The story operation structure is "7-3=?", and the calculation structure is also "7-3(=?)", which can be found easily by reading the story in order from the first sentence. Since the two structures are the same, this type of problem can usually be solved easily by the learners.

Meanwhile, in the following problem:

There are seven apples. Several apples were eaten. There are three apples.
How many apples were eaten?

We can derive the story operation structure as "7 - _ = 3", and the calculation operation as "7 - 3 = _". Since the two structures are different, a learner is required not only to understand the story but also to derive the calculation operation structure from the story. This type of problem is called "reverse-thinking problem".

2.3 Task Model of Problem-Posing

Based on the consideration of problem types, we have proposed a task model of problem posing as sentence-integration shown in Figure 3 (Kurayama & Hirashima, 2010). There are four main tasks in problem posing activity: (1) deciding calculation operation structure, (2) deciding story operation structure, (3) deciding story structure, and (4) deciding problem sentences. A learner should complete these tasks to pose a correct problem, although the execution procedure of the tasks is not decided in the model.

In the first step of MONSAKUN, subtraction or addition is selected as a *calculation operation structure*. In the second step, a story operation structure is decided. For example, for subtraction, four *story operation structures* can be selected. Among them, only one story operation structure is the same with the calculation operation structure (subtraction), and two of them have completely different story operation, that is, addition. Because this is an abstract transformation, it is often very difficult for learners to decide.

The next task of deciding *story structure* involves selection from four types of story: combination, increase, decrease, or comparison problem. Each type of story has its own structure, as explained in the sections above.

In the last task of deciding *problem sentences*, sentences are put into the story structure following the story operation structure. This task is divided into three more tasks: deciding sentence structure, deciding concept structure and deciding number structure. Deciding *sentence structure* is to select and order sentences following the story structure. For example, if the story structure is decrease problem, learner should make a sentence structure composed of an existence sentence, a decrease type of relational sentence, then followed by another existence sentence.

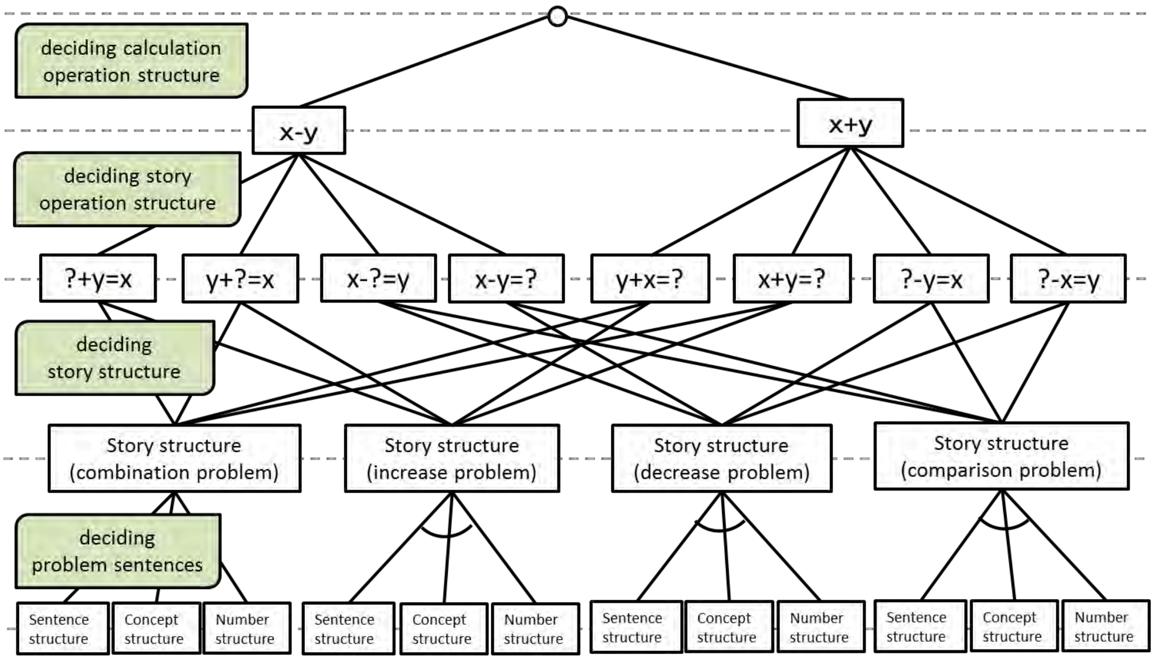


Figure 3. Task model of problem-posing as sentence integration

For the decision of *concept structure*, concepts dealt with the problem are decided. For example, if the problem requested learner to answer about the total number of apples and oranges, then the sentences should be dealt with apples and oranges as the concept. For the decision of *number structure*, the numbers dealt with the problem is decided. In arithmetic word problems, a negative number should not be used.

3. Analysis of MONSAKUN Log Data

In this section, the analysis of MONSAKUN log data from an experiment of MONSAKUN used by eleven undergraduate students from Faculty of Education is reported. Although MONSAKUN is intended for elementary school students, the subjects of this experiment are undergraduate students. The reason is that undergraduate students are supposed to be able to solve both forward-thinking and reverse-thinking problems rather easily, because they have already understood the structure of simple arithmetic word problems. They are only expected to learn how to make problems through the use of MONSAKUN.

On the other hand, elementary school students firstly learn about the problem structure through the use of MONSAKUN before they become able to pose problems, which takes several times of class schedule. Because the undergraduate students do not need to learn but only to recognize the problem structure, they are expected to show clearer changes in thinking process towards different problem types than elementary school students. For this reason, our study analyzed data from the experimental use of MONSAKUN by university students as the subjects. In the experiment, the subjects are firstly given explanation about the software, and then posed problems in a given time.

Our aim in this study is to examine learners' way of thinking from selection of sentence, especially the first selected sentence in each assignment. We analyzed the subjects' log data in assignments at Level 1 and Level 5 which require the subjects to pose forward-thinking problems and reverse-thinking problems, respectively. Both levels consist of 12 assignments that include four types of stories: combination, increase, decrease, and comparison. Each type of story has three assignments. Subjects carried out the assignments in order, and they can only move on to the next assignment when the current assignment has been answered correctly.

3.1 *Difference in First Selected Card between Level 1 and Level 5*

Figure 4 shows log data in Level 5 Assignment 1 mapped on a graph. Black nodes and links represent the ones selected by the subjects, while gray ones represent the ones not selected. In this experiment, not all paths were observed in subjects' selection. Subjects only took some particular paths, which show that the card selections are not random. In addition, focusing on the card firstly selected by subjects, most of them chose Card 4 (8 subjects out of 11). In this assignment, there is a decided tendency based on some sort of thinking. If we can clarify the tendency for learners to choose a specific first sentence card, it will be useful to diagnose learners' understanding. Therefore, as the initial step of analysis of students' thinking process, this study aims at revealing the characteristics of first selected sentence card.

The sentence cards in MONSAKUN contain different number according to the numerical expression in the given assignment. For example, in an assignment "Make a story problem about 'how many are the difference' that can be solved by 7 – 3.", the required calculation expression is "7 – 3 = _". In MONSAKUN, several sentence cards with numbers are provided to the users. The cards are distinguished by the order of numbers in the required calculation expression. If a card contains the first number in the required expression, for example, 7 in the example above, it is called *"first number card"*. Similarly, if it contains the second number or the third number, it is called *"second number card"* or *"third number card"*, respectively. One of the numbers in every assignment is an unknown number, which is represented by the blank mark.

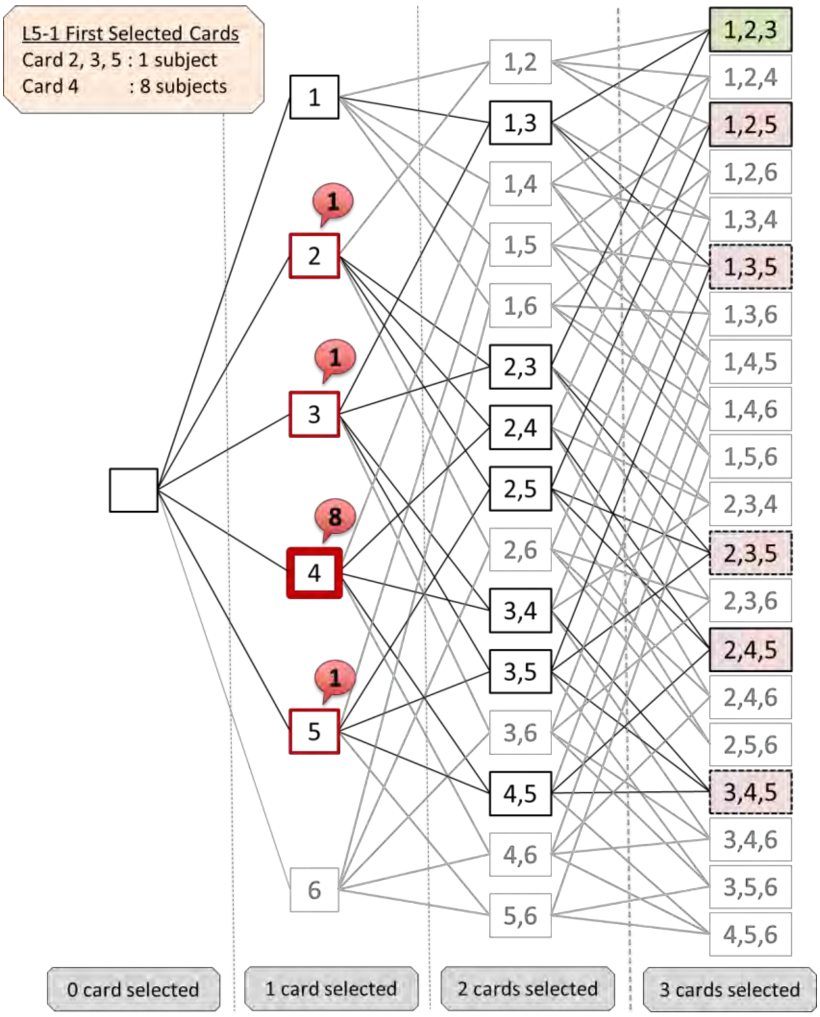


Figure 4. University students' card selection in Level 5 Assignment 1 shown in graph

From the analysis of all subjects' first card selection in Level 1 and Level 5 assignments, we found that the proportion of each sentence card to be selected firstly is entirely not even. Table 1 shows the proportion of first card selected by subjects in Level 1 and 5. In every assignment, only one or two cards are significantly chosen by them. We found that there is a bias against first selected card. This finding proves our assumption that subjects did not choose a card randomly, but with some sort of approach.

Table 1 Percentage of first selected card by the subjects

Type of first selected card	Level 1 (%)	Level 5 (%)
First number card	91.8	58.7
Second number card	3.3	16.5
Third number card (blank mark)	4.9	24.8

Furthermore, we found different trends of first card selection between Level 1 and 5. We presume that this difference appeared because subjects had different approach to pose either forward-thinking or reverse-thinking problems. In forward-thinking problem, the approach to order cards following the order of numbers in the numerical expression can be applied easily. However, in reverse-thinking problem they cannot pose problem with the same approach. This type of problem requires learners to think about the numerical relation in the given problem and reflect it to the choice of cards.

3.2 Change of Approach through the Exercise

In the previous section, we have presumed that subjects had different approach to solve forward-thinking and reverse-thinking problems. In this section, we would like to explain further how the subjects change their way of thinking during problem posing exercise by looking at the type of story, order of assignment, type of first selected card, as well as the type of sentence. We especially analyzed subjects' selection in Level 5 assignments, where they posed challenging reverse-thinking problems.

Table 2 shows the characteristics of first selected sentence card from each assignment at Level 5 that has marginal or significant difference in number of selection from the average. These results were analyzed with binomial test to the amount of each card being firstly chosen or not in each assignment. Binomial test is an exact test of the statistical significance of deviations from a theoretically expected distribution of observations into two categories. Based on our assumption that students posed problems by selecting cards through a thinking process, we expect the distribution of firstly chosen cards to have a significant difference in comparison with other cards.

Table 2. Result of binomial test of first selected card in Level 5 assignments

No	Type of story	Order of assignment	Type of first selected card	Type of sentence	p-value	
1	**Combination**	1st	*First number card*	*Existence*	$7.05*10^{-5}$	**
2		2nd	First number card	Relational	$1.88*10^{-7}$	**
3		3rd	First number card	Relational	$1.97*10^{-3}$	**
4	**Increase**	1st	*First number card*	*Existence*	$1.89*10^{-5}$	**
5		2nd	Second number card	Existence	0.0504	+
6		3rd	*First number card*	*Existence*	0.0504	+
7	**Decrease**	1st	*First number card*	*Existence*	$2.35*10^{-4}$	**
8		2nd	Second number card	Existence	$2.35*10^{-4}$	**
9		3rd	Second number card	Existence	$2.35*10^{-4}$	**
10	**Comparison**	1st	-	-	-	
11		2nd	Third number card	Relational	0.0266	*
12		3rd	Third number card	Relational	0.0266	*

**: significant difference (p<.01), *: significant difference (p<.05), +:, marginal difference (p<.1)

When firstly used MONSAKUN, subjects are given simple forward-thinking problems to pose at Level 1. From the analysis mentioned in Section 3.1, we found that they first simply chose a card with the first number in the required numerical expression (*"first number card"*), and then proceeded to choose other appropriate cards. This approach worked well for assignments in Level 1, where all of the assignments are forward-thinking problems.

When subjects arrived at the first assignment of Level 5, they initially approached the assignment with the same way of thinking in choosing the first sentence card. However, this did not work well, and they tend to make more mistakes than in the previous levels. We presumed that the subjects were aware that the previous approach did not work for reverse-thinking problems, because in the second assignment of Level 5 they tend to choose another type of card.

In a similar way, subjects changed their approach from the first assignment in a story type to the second and third assignment in the same story type. As shown in Table 2, in the first assignment in each type of problem, they generally took the simple approach to firstly select a *"first number card"* containing an *existence sentence*. Only in the case of comparison story there was no significant difference in cards selected by subjects in the first assignment. On the other hand, in the second and third assignments of the same type of story, they did not choose it as the first card. For example, in combination stories, most of them firstly did not select *existence sentence*, but "*first number card"* containing *relational sentence*. This is also the same as in the decrease story type.

Meanwhile, in the case of increase stories, we did not found any evident change between the assignments. At the second assignment they tend to select *"second number card"* containing *existence sentence* (shown by a marginal p-value), while at the third assignment they took the simple approach just like in previous levels of forward-thinking problems.

Furthermore, in comparison stories, there is no trend in first card selection at the first assignment. However, at the second and third assignment, there is a trend to select *"third number card"*, that is a blank mark, containing a *relational sentence*. Consequently, we observed that there is a change of approach in comparison story compared to the previous story types.

This leads to two findings about changes in subjects' way of thinking through the exercises. The first one is that subjects change their approach to pose problems after they had experienced posing the same type of story. As shown in Table 2, trends of first card selection are different between the first assignment and the rest in the same story type. The next finding is that the change of approach depends on the type of story, as we can see that subjects made different first card selection in different story type.

Table 3. Average of steps and mistakes in Level 5 assignments

Story Type	Average No. of Steps		Average No. of Mistakes	
	Assignment		Assignment	
	1st	2nd & 3rd	1st	2nd & 3rd
Combination	11.60	4.20	1.5	0.3
Increase	45.50	16.50	8.4	1.4
Decrease	24.90	16.30	3.3	1.9
Comparison	10.00	9.80	1	0.5

These changes of approach seem to bring a good effect to subjects' thinking process in posing reverse-thinking problems. Our analysis of the average of steps and mistakes in Level 5 problems showed that in comparison to the first assignment of each story type, the average of steps and mistakes in the second and third assignments of the same story type are mostly decreased, as shown in Table 3.

4. Concluding Remarks

In this research, we have conducted analysis of MONSAKUN log data of university students' problem posing activity to investigate their way of thinking in posing different types of arithmetical word problems. From the analysis, we found that the first sentence selected in each assignment were different in several ways. In forward-thinking problems, subjects generally used a simple approach to select *“first number card”*. However, in reverse-thinking problems, they changed the approach to select *“second number card”* or *“third number card”*. Depending on the type of story and subjects’ exercise experience, they applied different approach of first card selection. These finding proves our assumption that learners who used MONSAKUN did not chose cards randomly, but with some sort of thinking process. Furthermore, we infer that learners who used MONSAKUN were able to recognize the differences in structure of problems depending on types of story, as they changed their approach to pose problems for different story types. The recognition of the difference is important for learners to understand the nature of arithmetic word problems.

For the next step of this research, we plan to perform the analysis to a larger data of MONSAKUN used by elementary school students to infer their thinking process of the same problem posing activity. The results will be used to make an elaborate process model of the problem-posing and adaptive support of the process.

References

Brown, S. I., & Walter, M. (1990). The art of problem posing. NJ: Lawrence Erlbaum Associates.

Ellerton, N.F. (1986). Children’s Made Up Mathematics Problems: A New Perspective on Talented Mathematicians. *Educational Studies in Mathematics*, Vol.17, 261-271.

English, L. D. (1997). The development of fifth-grade children’s problem-posing abilities. *Educational Studies in Mathematics*, 34(3), 183-217.

English, L. D. (1998). Children’s problem posing within formal and informal contexts. *Journal for Research in Mathematics Education*, 29(1), 83-106.

English, L. D. (2003). Problem posing in elementary curriculum. In F. Lester & R. Charles (Eds.), *Teaching Mathematics through Problem Solving*. Reston, Virginia: National Council of Teachers of Mathematics.

Hirashima, T., Nakano, A., & Takeuchi, A. (2000). A Diagnosis Function of Arithmetical Word Problems for Learning by Problem Posing. *Proceedings of PRICAI2000*, 745-755.

Hirashima, T., Yokoyama, T., Okamoto, M., & Takeuchi, A. (2007). Learning by Problem-Posing as Sentence-Integration and Experimental Use. *Proceedings of AIED2007*, 254-261.

Hirashima, T. & Kurayama, M. (2011). Learning by Problem-Posing for Reverse-Thinking Problems, *Proceedings of AIED2011*, 123-130.

Hirashima, T., Yokoyama, T., Okamoto, M., & Takeuchi, A. (2008). Long-term use of learning environment for problem-posing in arithmetical word problems. *Proceedings of ICCE2008*, 817-824.

Kurayama, M., & Hirashima, T. (2010). Interactive Learning Environment Designed Based on Task Model of Problem-Posing. *Proceedings of ICCE2010*.

Nakano, A., Hirashima, T., Takeuchi, A. (1999) Problem-Making Practice to Master Solution-Methods in Intelligent Learning Environment, *Proceedings of ICCE'99*, pp.891-898.

Nakano, A., Hirashima, T., Takeuchi, A. (2002) An Evaluation of Intelligent Learning Environment for Problem Posing, *Proceedings of ITS2002*, pp.861-872.

National Council of Teachers of Mathematics (2000). Principles and standards for school mathematics. Reston: Va, NCTM.

Polya, G. (1957). How to solve it: A new aspect of mathematical method. Princeton University Press.

Riley, M.S., Greeno, J.G., & Heller J.I. (1983). Development of Children’s Problem-Solving Ability in Arithmetic. The Development of Mathematical Thinking, Ginsburg H. (ed.), Academic Press, 153-196.

Silver, E.A. & Cai, J. (1996). An Analysis of Arithmetic Problem Posing by Middle School Students. *Journal of Research in Mathematics Education*, vol.27, No.5, 521-539.

Stanic, G., & Kilpatrick, J. (1988). Historical perspectives on problem solving in the mathematics curriculum. In R. Charles & E. Silver (Eds.), *The teaching and assessing of mathematical problem solving (pp. 1–22).* Reston, VA: National Council of Teachers of Mathematics.

Yamamoto, S., Kanbe, T., Yoshida, Y., Maeda, K., & Hirashima, T. (2013). Learning by Problem-Posing with Online Connected Media Tablets. *Proceedings of HIMI/HCII 2013*, Part III, 165-174.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Balance Control of Question-Posing Focusing on Learning Target Words on the Self-Study Material Contribution and Sharing System

Toshihiro HAYASHI[a*], Yuji HIRAI[b], Kazuhiro URA[a], Akihiro IWAKI[c], Rihito YAEGASHI[b], Hiroshi MURAI[c] & Hiroyuki TARUMI[b]
[a]*Information Technology Center, Kagawa University, Japan*
[b]*Faculty of Engineering, Kagawa University, Japan*
[c]*University Consortium for E-Learning, Shikoku Center, Kagawa University, Japan*
*hayashi@eng.kagawa-u.ac.jp

Abstract: Recently, education and learning by using e-Learning style have become popular. We have developed a self-Study Material Contribution and Sharing System called "S-Quiz" wherein students can pose questions and share them. Through the educational practice with S-Quiz, we found that balance of posed questions was not suitable to use for learning after question-posing. This paper describes balance control of question posing in order to solve this problem of unbalanced questions. This balance control aims to increase questions of which posed number is a few by system-centered navigation to students on S-Quiz. In addition, we propose a control method which is not biased to a part of students.

Keywords: e-Learning, Self-Study Material Contribution and Sharing System, S-Quiz, Balance Control of Question-Posing

1. Introduction

Recently, education and learning by e-Learning have become popular in especially higher education, job training, and so on. We have developed a self-Study Material Contribution and Sharing System called "S-Quiz" wherein students can pose questions and share them(Mizuno *et al*, 2007, Hayashi *et al*, 2008). S-Quiz is categorized into educational system by using question-posing. There have been developed various kinds of such question-posing systems(Nakano *et al* 2002, Yu *et al* 2005, Takagi and Teshigawara 2006, Hirai and Hazeyama 2007, Denny *et al* 2008).

We used S-Quiz in home work for a special subject of university. In the home work, each student could freely pose questions about several important keywords which s/he understands well. We call the important keywords which can be selected in a lecture LTWs: Learning Target Words in this paper. In this educational practice, we found students posed many questions for several LTWs as a positive result but other several LTWs have no posed question as a negative result. We can think that the unbalance of question-posing can happen by allowing such style of question-posing mentioned-above.

Posed questions are expected to be shared and be used their learning after question-posing. Therefore, unbalanced question sets are not suitable for student to use for their learning. In order to avoid the unbalance of questions, we try to add a balance control function into S-Quiz. In this balance control function, "good balance of question posing" is defined that the number of questions about each LTW is bigger than the threshold value (for example, minimum number of posed-questions). Based on this idea, we set following conditions:

(1) Balance Control function navigates student to pose question about LTWs of which the number of posed questions is smaller than the threshold value, and
(2) The navigation is not concentrated to a part of students.

In this paper, we propose two balance control methods: basic method and improved method. Especially, explain about the simulation result of the improved method and show this method satisfy the conditions (1) and (2) on simulation level.

2. S-Quiz

S-Quiz provides a learning environment wherein students can pose questions freely and share the question with other students. Fig.1 shows a snapshot of question posing interface of S-Quiz. After choosing question category, students can make a question by inputting a question text, one correct answer, and three incorrect answers on the question posing interface and then post it. In addition, students can input hint information of questions by text and image data. In the case of Fig.1, the student poses a question about ENIAC. Generally, ENIAC becomes a LTW which can be selected in the genre about computer system. As for the genre, student can select a proper genre from the pull down menu. The set of genre is chosen by the teacher in advance.

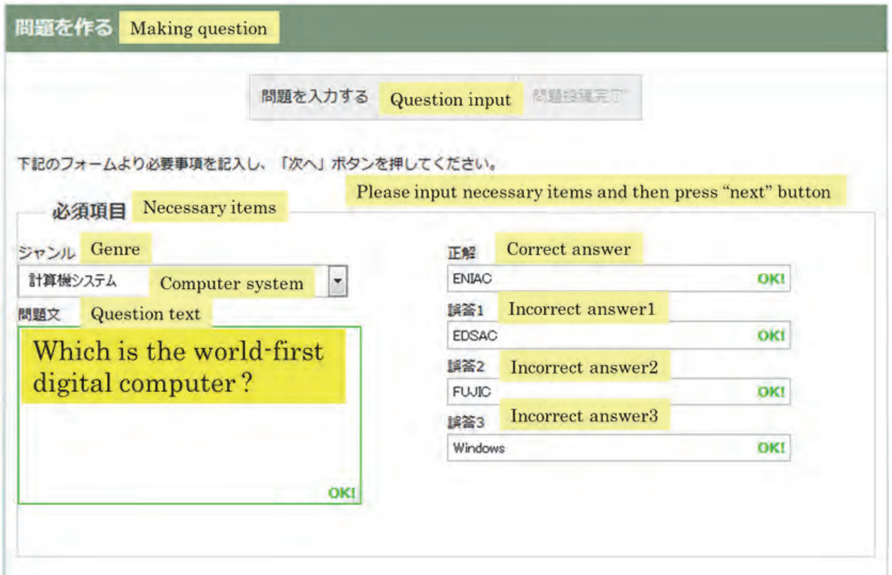


Figure.1 Question posing interface of S-Quiz

Fig.2 shows a snapshot of question answering interface of S-Quiz. Students can also share all questions posed by other students and use them for students’ free learning. Beside the multiple choice question, the genre, the maker (student) and the evaluations of the questions quality are displayed. As for the evaluation, students can vote three kinds of evaluations: “Good”, “Average” and “Not good” to questions which the student has answered. Hint message and image can be also displayed if they are prepared. Basically, the student can answer the question by selecting a correct answer candidate from four choices, but s/he can select “Pass” to obviously inform that s/he does not know.

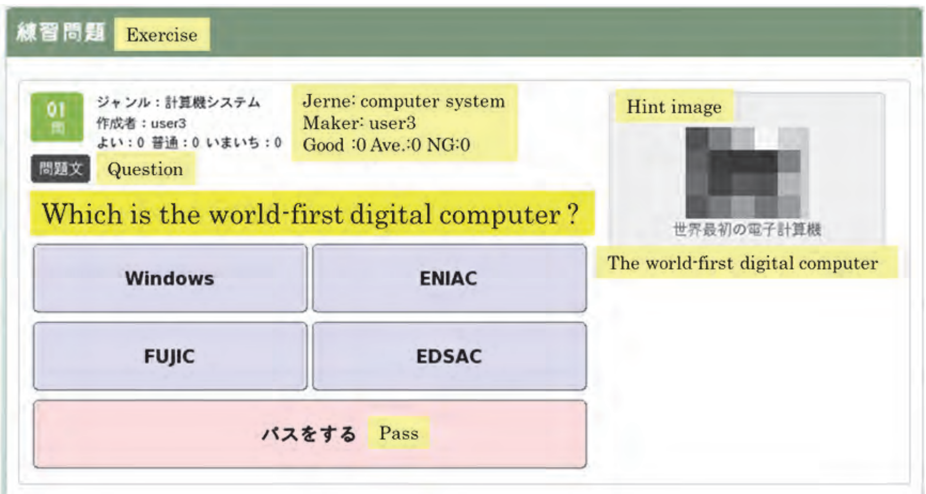


Figure.2 Question answering interface of S-Quiz

3. Balance control of question-posing

3.1 The reason of unbalance of question posing

In S-Quiz, student's actions are mainly classified into question-posing and question-answering. LTWs are important words in a lecture. Therefore, it is ideal style of question-posing in S-Quiz that all students make questions about all LTWs. However, we should consider the other factors: the number of students, understanding levels for each LTW, and so on. As for the number of students, if it is large number, the number of posed questions also becomes large. Sometimes it becomes exceed to use. In addition, the possibility of low quality of questions becomes high, because it is not guaranteed that all students can good quality of question for every LTW from the viewpoint of understanding levels for each LTW. Therefore, it is a realistic and loose solution for answering to the mentioned-above that each student poses questions about several LTWs which s/he understands well. In contrast, the unbalance of question-posing can happen by allowing such style of question-posing.

3.2 Conditions of Balance control

Balance control of question-posing is to increase question-posing about LTWs of which the posed number is smaller than the threshold value in order to avoid unbalance of question-posing. S-Quiz focuses on such LTW and navigate student to pose question about the LTW. This control makes question-posing balance good with satisfying the following conditions: (1) Balance Control function navigates student to pose question about LTW of which the number of posed questions is smaller than the threshold value, and (2) The navigation is not concentrated to a part of students.

3.3 Basic method

Considering only balance of question-posing, we can just focus on the condition (1). As for LTWs of which the number of posed questions is smaller than the threshold value, S-Quiz navigates students to pose questions about the LTW. We call this navigation as "basic method." The followings are concreate process of basic method:
(BM1) Estimation of the number of questions which students will pose from now,
(BM2) Calculation of probability that each LTW is included in questions, and
(BM3) Decision of LTWs that navigation of question-posing is needed by using results of the above mentioned (BM1) and (BM2).
The basic method is simple to be implemented. However, navigation based on the basic method tends to become frequent at the late stage of question-posing. In other words, the navigation is concentrated to a part of students who pose questions at such stage. Unless students accept the navigation, the balance of question-posing cannot be improved.

3.4 Improved method

Focusing on both conditions (1) and (2), we propose "improved method" for balance control of question-posing. This improved method positively navigates students even at early stage of question-posing. In order to realize this feature, the necessary number of posed question of each LTW is calculated automatically according to the situation. If the current number of posed question about a LTW is smaller than the necessary number, navigation works for question posing about the LTW. To fulfil the condition (2), the improved method activates navigation at early stage of question-posing even when navigation is not needed.

3.5 Discussion

Generally, it can be useful for keeping the balance of numbers to use both upper and lower limits. However, both basic method and improved method use only a lower limit. If upper limit is set, it can occur that a student is not allowed to pose the question about LTWs on upper limit. This is a bad situation for motivation of the student and also reduces the possibility that the student can pose questions. The root purpose of balance control is to guarantee the minimum number of posed questions about each LTW. If the root purpose is satisfied, there is no need to set the upper limit. In addition, lower limits control by both methods loosely restricts exceed question-posing for each LTW. Through these considerations, the basic and improved methods put focus on only lower limits.

4. Simulation experiment

We did simulation experiments for evaluating performance of the improved method. In this section, we describe the outline of experiment, the simulation results, and discussion about the results.

4.1 Outline of simulation

Based on various kinds of data about educational practice with S-quiz, we assume student behavior and tendency of question-posing as followings:
(a) Students pose the required number (for home work) questions at once,
(b) Students do not use LTW which they have used before, and
(c) Students pose a question by using a LTW.
In addition, the following feature is included for simplification of group level behavior of students:
(d) Question-posing tends to be distributed and to be not overlapped.
Fig.3 shows illustrated figure about the features (a) and (d).

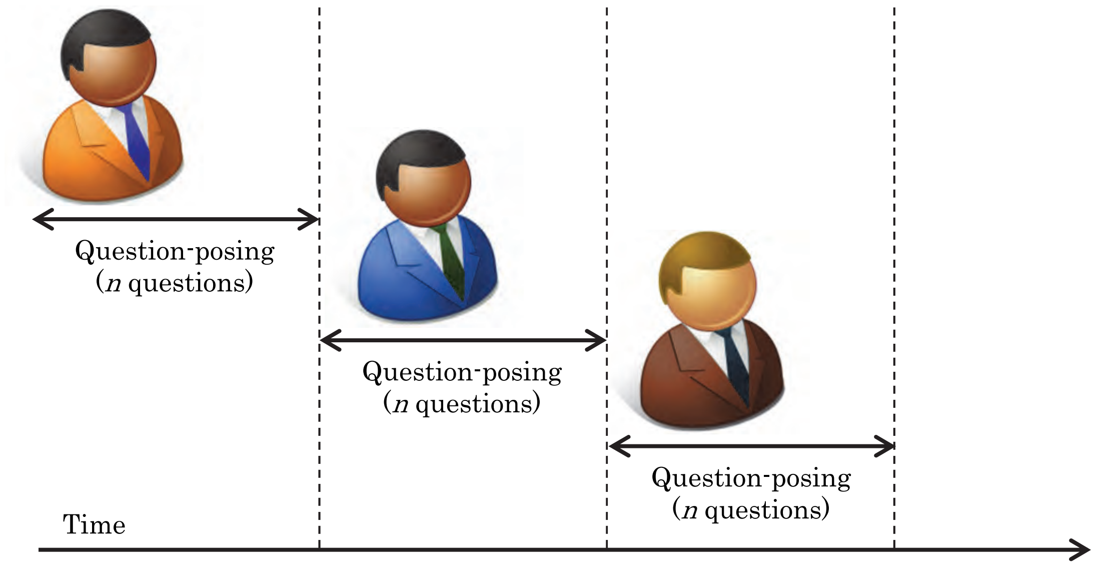


Figure 3 Question-posing tendencies of all students

4.2 Simulation results

In this simulation, we set parameters as Table 1. The threshold value means the minimum number for each LTW. We obtained the results from the large number of simulation with the parameters and calculated the average.

Table 1 Simulation parameters

Parameters	Simulation value
the number of students	80
the number of question posing per one student	10
the number of LTW	20
the threshold value	20
the probability of acceptance ratio to navigation	50%

Fig. 4 shows the number of posed questions about each LTW. The number of each LTW shows tendency of use. Small number means question of the LTW is easy to be posed. In contrast, big number means question of the LTW is hard to be posed. Therefore, it becomes high possibility that the number of posed question about LTW with big number is not over the threshold value without the navigation.

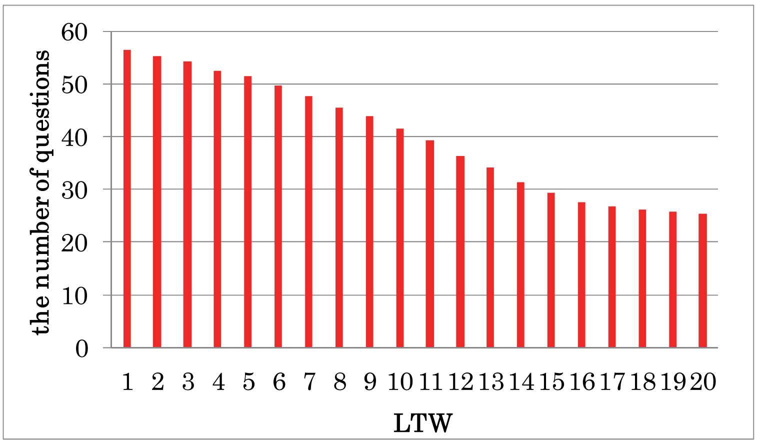


Figure 4 The number of posed questions about each LTW (by Improved method)

In addition, Fig.5 shows the number of navigation to each student. The number of student ID shows order of question-posing. For example, student 2 poses questions after student 1 poses questions (refer to Fig.3). This means that students with big number pose questions at late stage. Such students have to pose questions under frequent navigation.

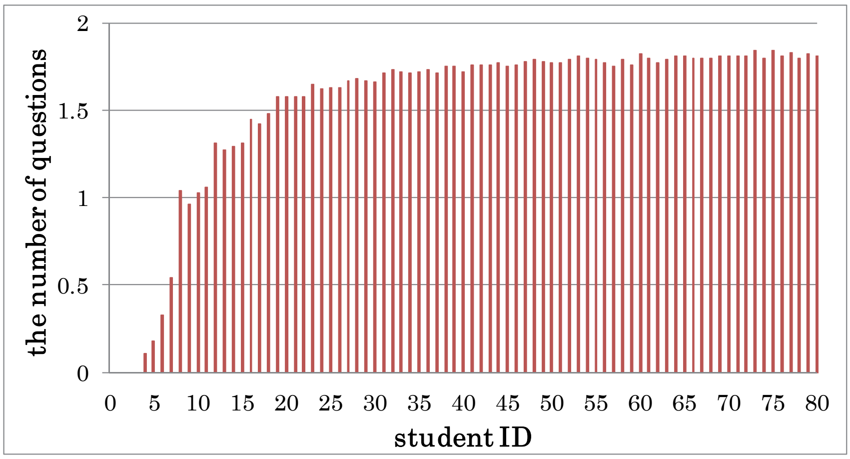


Figure 5 The number of navigation to each student (by Improved method)

4.3 Discussion

As for simulation results, we can find the every number of posed-questions for each LTW is over the threshold value (the minimum number) from Fig.3. In addition, we can find the number of navigation is under two times and there is no big difference of navigation among all students from Fig. 4. We can say the improved method satisfies the condition (1) and (2). We also did simulation experiments with various parameter settings and obtained good results for the improved method. However, we omit those results because of paper length limitation.

5. Summary

In this paper, we explained about development and the educational practice issue of S-Quiz: a self-Study Material Contribution and Sharing System wherein students can pose questions and share them. In S-Quiz, each student is allowed to pose questions about several LTWs which s/he understands well. S/he does not have to pose questions about every LTW. In contrast, the unbalance of question-posing can happen by allowing such style of question-posing. Then, we proposed two methods for balance control of question-posing focusing on LTWs and simulation experiments. From the experiment we found our proposed method fulfil the following conditions: (1) Balance Control function navigates student to pose question about LTW of which the number of posed questions is smaller than the threshold value, and (2) The navigation is not concentrated to a part of students. The performance evaluation at real educational practice is remained.

Acknowledgements

This work was supported by MEXT (The Ministry of Education, Culture, Sports, Science and Technology) KAKENHI 26330401.

References

Mizuno, T., Hayashi, T., Tominaga, H. and Yamasaki, T.(2007) Design and Prototyping of a Self Study Material Contribution and Sharing System, *Proc. of ICCE2007*, 37-38.

Hayashi, T., Mizuno, T., Tominaga, H., Tarumi, H. and Yamasaki, T.(2008) A Learner-centered e-Learning Environment by Contributing and Sharing Problems Created by Learners, *Proc. of CollabTech'08*, 94-98.

Nakano, A., Hirashima, T. and Takeuchi, A.(2002) An Evaluation of Intelligent Learning Environment for Problem Posing, *Proc. of ITS'2002*, 861-872.

Yu, F.-Y., Liu, Y.-H. and Chan, T.-W.(2005) A web-based learning system for question posing and peer assessment, *Innovations in Education and Teaching International*, 42(4), 337-348.

Takagi, M. and Teshigawara, Y.(2006) A WBT System Enabling to Create New or Similar Quizzes Collaboratively by Students, *Proceedings of ICET2006*, 263-268

Hirai, Y., and Hazeyama, A.(2007) A Learning Support System based on Question-posing and Its Evaluation, *Proc. of C5*, 180-185.

Denny, P., Hamer, J., L.-Reily, A. and Purchase, H.(2008) PeerWise: Students Sharing their Multiple Choice Question, *Proc. of ICER'08*, 51-58.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Development of a Customized English Learning System based on Augmented Reality Technology

Gwo-Haur Hwang[1*], Beyin Chen[2] & Hen-Lin Huang[2]
[1]*Department of Information Networking and System Administration, Ling Tung University, Taiwan*
[2]*Department of Information Management, Ling Tung University, Taiwan*
*ghhwang@teamail.ltu.edu.tw

Abstract: The augmented reality (AR) is a kind of technology that can combine the virtual information and the real image. Many scholars indicate that AR can effectively enhance learners' learning motivation and effectiveness. This study proposes an English learning system based on AR technology. For customized design, the system adds the functions of related words and 3 learning scopes: phrases, and sentences, related words. In addition, an easy mode and an advanced mode are provided. These functions are less considered in the past and will also be regarded as the basis of future personalized preference analysis. Through the system, it is hoped that the learning motivation and learning effectiveness of learners can be enhanced. In the future, personalized preference will be analyzed according the collected historical data of learners' behavior.

Keywords: customized design, English learning, augmented reality, personalized preference

1. Introduction

With the advances of sciences and technologies and the rise of the concept of the global village, people around the world can cross national boundaries through networks and communicate in the aspects of information and culture. Among these developments, the language is essential as a communication bridge. In the global trend of internationalization, learning English is no longer just the needs of individuals only, but also a trend of into a new era (Hwang, Lee, Hwang, Huang, Lin & Cai, 2013). Taiwan is an EFL (English as Foreign Language) English learning environment. In Taiwan, a formal exposure to English is until entering elementary schools (Lu, 2012). In the traditional teaching environments, through explaining the rigid text of the textbooks by teachers, the English learning can be conducted, which is passive learning (Savigan, 1988). In other words, the teachers just unilaterally give students the knowledge. In addition, due to the restrictions of the teaching time, students usually recite the knowledge, so that the knowledge cannot be applied to everyday life (Brown, Collins & Duguid, 1989).

Looi, Seow, Zhang, So, Chen and Wong (2010) have used a context-aware ubiquitous learning system to record the situation of learners. It allowed teachers to rapidly understand students' learning process and individual differences and to develop better teaching methods. By this, students can get the most necessary information in a timely and appropriate manner (Chen, Lien & Lu, 2009). The development process of context-aware technology first is RFID (Radio Frequency Identification) technology, which can complete a non-contact identification process by using the RFID reader to induce the RFID tags (Landt, 2005). Although the reliability and the identification speed of RFID technology are very high, but most smart phones and tablets do not support the reading of the RFID tags. In addition, the cost is also much more expensive than QR (Quick Response) codes. Thus, in the field of context-aware ubiquitous learning, QR Code technology has gradually replaced RFID technology. However, the virtual information and reality information provided by these two technologies are completely separated. This will result in that the users do not have coherent information when they receive information. According the "Spatial Contiguity Principle" and "Temporal Contiguity Principle" of 12 multimedia design principles proposed by Mayer (2009), the learning effectiveness of learners can be increased if at the time of scanning real objects, the corresponding and associated information can be immediately generated beside the objects.

The augmented reality (AR) is a kind of technology that can combine the virtual information and the real image (Azuma, 1997). The technology allows that the learning process can more meet the above principles proposed by Mayer. Billinghurst (2003) has pointed out that AR can provide unique educational benefits. First, the use of AR as teaching aids allows learners smoothly to interact with virtual objects (interactive learning concept) in virtual and real environments. Second, the use of AR as teaching aids will extend as a new teaching and learning strategies. The learning mode can be conducted even if the students do not have any computer experience. Finally, AR has the characteristics that let learners be immersed in learning contents. It allows that learning is no longer just to face boring textbooks. Lai, Hwang, and Chen (2012) have pointed out that AR can indeed effectively enhance students' learning motivation. Therefore, if a language learning system is built through AR technology, learners will be able to scan the learning objects directly in the real scene. The real-time multimedia information or teaching materials can be obtained, without destroying the original scene, but also saving the additional cost of teaching objects.

In addition, past researcheres have indicated that learners like to have options to set the functions according to their own preference and actual situation especially when the learning environment has many variables (Mitchell, Chen and Macredie, 2005). Therefore, this study proposes a customized AR English learning system. It is different from the general paper textbooks or traditional media textbooks which are just flat texts or pictures and lack context and interactivity. We expect that the customized AR English learning system can increase the learning wishes of learners through the good interactivity of AR technology. In addition, it can record the learning process which provides the future teaching reference for teachers.

2. Literature Review

2.1 The problems of English learning

In the past, second/foreign language learning relied upon teachers' lecturing to explain the learning materials of textbooks (Savignon, 1988). In other words, because the teaching time is limited, traditional teaching is considered unable to enhance learning motivation and interest. It may result in the students' to learn English by a rote style rather than by a style of increasing knowledge and problem-solving ability (Brown, Collins, & Duguid, 1989). Looi, Seow, Zhang, So, Chen and Wong (2010) have indicated that if the technologies of context-aware and ubiquitous learning are imported in teaching, learning will be able to be conducted at any time and any place. Thus, the role of teachers will transfer from imparter to guider. Teachers will guide students to learn actively and attract their attention so that the learners' ability of observation of the real world and the ability to actually solve problems can be enhanced. (Chen, Lien & Lu, 2009). Among the technologies of context-awareness, RFID and QR code technologies have the disadvantages of information discontinuous problem. Therefore, this paper uses AR technology to implement our learning system.

2.2 The applications of AR technology on education

AR is to import the images, objects, and scenes generated by computers to the real environment. Its purpose is to enhance the effect of perception. That is, the virtual objects are added to the real environment. This technology must have 3 characteristics: "combine the virtual and the real world", "be able to interact immediately", "be necessary in 3D space" (Azuma, 1997). Milgram, Takemura, Utsumi and Kishino (1994) regard real and virtual environments as a closed set as shown in Figure 1. The left is a purely real environment and the right is a purely virtual environment. The virtual reality attempts to replace the real world, while the augmented reality is to augment the virtual picture generated by computers into the real environment.

Figure 1. The definition of AR (Milgram et al., 1994).

AR is currently widely used, for example, education, medical science, military training, engineering, industrial design, art, entertainment, etc. (Azuma, 1997). Among these applications, the application of AR on education has obtained the attention of many scholars. The reason is that the traditional way of multimedia learning lacks immediacy and interactivity. During the learning, the virtual and real information is separated. This disadvantage has been greatly improved after the rise of AR technology. Therefore, there are many scholars applying AR technologies to mobile navigation and context-aware ubiquitous learning.

There are many applications of AR technologies on mobile navigation. For example, Kuo (2008) applied AR to "epidemic battle Camp" exhibition at the National Science and Technology Museum. It allowed the audience to interact with objects in a 3D environment. The objects may be not able to actually take to the exhibition; may be not able to let the audience touch; or may be necessary to be amplified. Thus, the better communication effectiveness can be achieved. In addition, the stay time of visitors in front of display units can be extended. Sejin and Woontack (2009) proposed the guide learning of context-aware applied in the Museum of Art. It allowed users to follow the guiding indicator to deepen the art articles. In addition, according to different users, it provided personalized guidance interface. However, this study did not carry out effective analysis. Lin, Tang and Peng (2011) also used AR technologies to build digital teaching materials of arts and humanities learning of an elementary school. The results showed that the learners felt comfortable and easy when learning and operating AR aids. In addition, they have also maintained a certain degree of concentration and a sense of curiosity to AR aids. Wernhuar and Ou (2012) proposed an AR system of butterfly virtual ecological learning environment used on smart phones. The system was easy to maintain and solved the problem of insufficient butterfly species. The experimental results showed that the use of AR technology can improve the learning effectiveness. Chen and Tsai (2013) also applied AR to a library of an elementary school. They explored the effect of gender, prior knowledge and cognitive style on learning. The results found that AR learning way can enhance the overall learning satisfaction of learners. In summary, AR applied to mobile navigation can increase the learning willing and motivation of learners, but also enhance the learning effectiveness.

Many studies have also applied the AR technology to classroom teaching and reading. For example, Dünser and Hornecker (2007) studied and looked at the children' learning status of reading textbooks. They explored how for children of 6 to 7 years old to operate interactive teaching media. The results showed that a rich interface may increase the willingness of students to learn. Amir and Vineet (2012) proposed a system to assist construction engineering students with AR technology to simulate the actual construction. Thus, the shortcomings of traditional construction which cannot practice actually can be improved. Under the simulation of AR environments, the students can conduct learning according to the real situation. Wu, Li, Yao and Pai (2012) designed an interactive AR system of chemistry experiment. Students can operate the picture cards of experiment equipment to complete a virtual chemistry experiment via AR technology. The difference between the traditional teaching style and the AR teaching style is that the students can watch the entire chemical reaction process through the 3D model. In addition, the experimental dangers that may occur can also be avoided. Chu and Lin (2013) have built an AR system which combined Kelly grid. It was used in natural science courses of an elementary school. The results showed that students' learning attitude was improved significantly. Chen, Zhao, Liu, Lin and Lu (2013) have built an AR system of gear teaching which imported the game concept. The experimental results indicated that the system can effectively enhance the concentration and the learning interest of students. In conclusion, we found that the use of AR technologies in the classroom can improve the shortcomings of the previous learning environment. In addition, it also enhanced the concentration of learners in classroom.

In addition, there are many applications of AR technology on language learning. For example, Hsieh and Lin (2009) designed an AR system of English vocabulary learning which had immersive learning outcomes. The experimental results showed that the learners will be willing to use the system. Chang, Chen, Huang and Huang (2010) also built an AR game-based English vocabulary learning system. The above two English vocabulary learning systems are both based on AR technology, but the methods are different. The former is to use English vocabulary magic books so that learners can scan the learning objects in the books while the latter is to scan the 3D learning objects directly to conduct learning. The advantage of the latter is able to scan real objects in a real environment. It does not need extra teaching objects. So, this study adopts the latter method. Tsai, Li and Wu (2011) applied AR technology to Chinese learning. The learning system allowed foreigners to learn Chinese by combining learning objects of text boxes and the instant interactivity of AR technology.

However, the above teaching systems are all for vocabulary. The applications of related words or example sentences are lacking. In addition, it is inconvenient because additional objects are required. Therefore, in this study, we will develop a customized English learning system based on AR technology. It is hoped that through the combination of the real situation and the virtual information, the learning motivation and effectiveness of learners can be enhanced.

2.3 Customized design

Personalized design is to design system architecture or user interface according to the unique and special requirement of every user (Fink & Kobsa, 2000). Personalized design can be provided by 2 styles. One is customized and the other is adaptive. Regarding customized, users have the right of selection. That is, they are allowed to modify the content presentation styles, user interfaces and navigation tools themselves. Regarding adaptive, the system actively provides appropriate content presentation styles, user interfaces and navigation tools according to the observed user behavior. In other words, the former is user-oriented while the latter is system-oriented. Both have advantages respectively. The former gives users considerable options, but needs extra effort to choose their suitable manners. The latter although can decrease users' loading, but may erroneously compute the users' preferences. However, the past researches have shown that most users prefer to have options (Mitchell, Chen and Macredie, 2005). Especially when the learning environment has many variables, the users like to set the content presentation styles, user interfaces, and navigation tools according to their own preference and actual situation.

3. System Implementation

3.1 The hardware architecture

Learners only need to install the customized AR English learning system to a smart phone or a tablet and then can scan the learning object using the smart phone or the tablet. When the identification is completed, the interactive learning in the real object and the virtual teaching material will be conducted. The operation of the system is very easy. So, even if the learners do not have any computer experience, they can use the system very easily. The hardware architecture of the system is shown in Figure 2.

Figure 2. The hardware architecture.

3.2 The software architecture

The software architecture of the system is shown in Figure3. There are 4 databases, which store text material, voice material, picture material, and the usage behavior record of learners, respectively. The text material database stores the English words, phrases, and sentences of all objects. The voice material database stores the English and Chinese pronunciation of all teaching material. The picture material database stores the animation and image of all objects. The learners' portfolio database records the behavior operation process, for example, the setting of personalized preference, the time of all objects scanned, the time of the function button clicked.

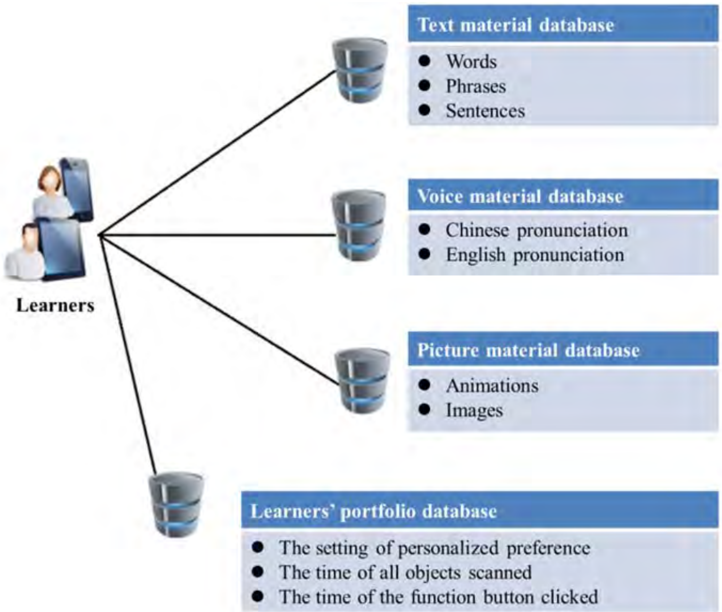


Figure 3. The software architecture.

3.3 The snapshots when the system was executed

The operation of the system is very simple. After starting the system, learners just need to scan the real objects with smart phones or tablets as shown in Figure 4.

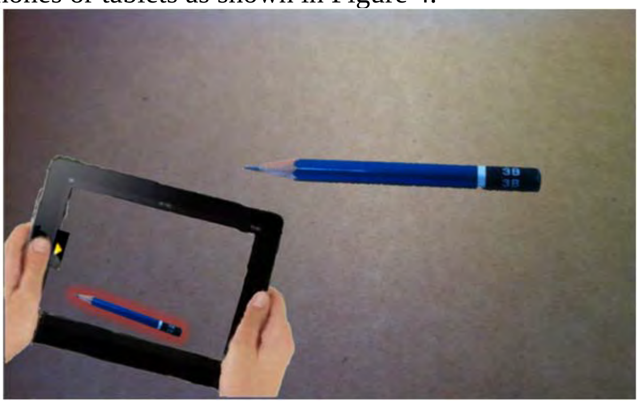


Figure 4. Scan the actual scene objects with a tablet.

The system provides 3 learning scopes: phrases, sentences, related words. In the learning process, learners can turn on or off these functions with their own preferences. In addition, there are 2 learning modes can be selected. One is the easy mode and the other is the advanced mode. These functions are for customized design and as shown in Figure 5.

Figure 5. Customized selection.

When learners successfully scan an object, the screen will show the main teaching material (for example, a pen) and the teaching material of related word (for example, an eraser). A flashing light will appear around the two teaching materials to enhance the display, so that the learners can clearly know where the operable objects are. After the learners click the object that want to learn (For example, the eraser), the corresponding function buttons will appear, as shown in Figure 6. Such design is to meet the "Spatial Contiguity Principle" and "Temporal Contiguity Principle" of 12 multimedia design principles proposed by Mayer (2009). The design allows learners can simultaneously observe the real learning object and corresponding virtual teaching material. Thus, the best learning effectiveness can be obtained.

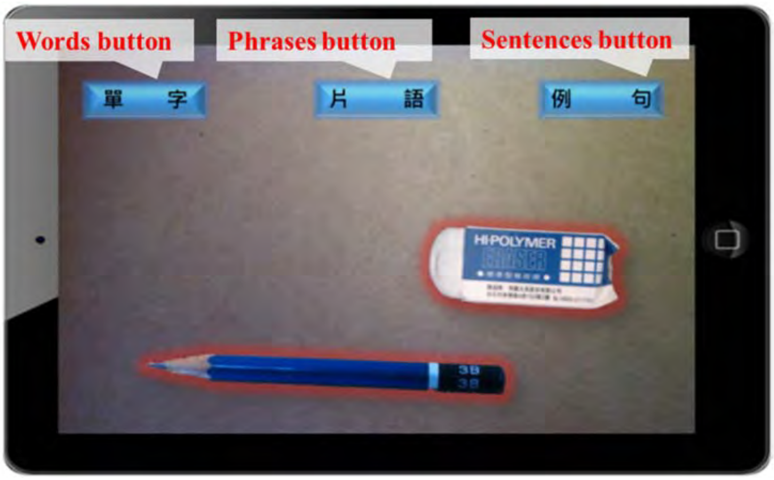


Figure 6. Learners successfully scan the learning object.

When learners click the function button of "words", the system will first split the word into letters and read out each letter. Meanwhile, the screen also appears the corresponding letters in marquee style. Then, the word will be read out once. Then, the Chinese meaning will be explained using Chinese voice, as shown in Figure 7. Such design is to meet the "Modality Principle" of 12 multimedia design principles proposed by Mayer (2009). The design allows the learners can use auditory and visual multisensory to receive a single message. Thus, the best learning outcomes can be achieved.

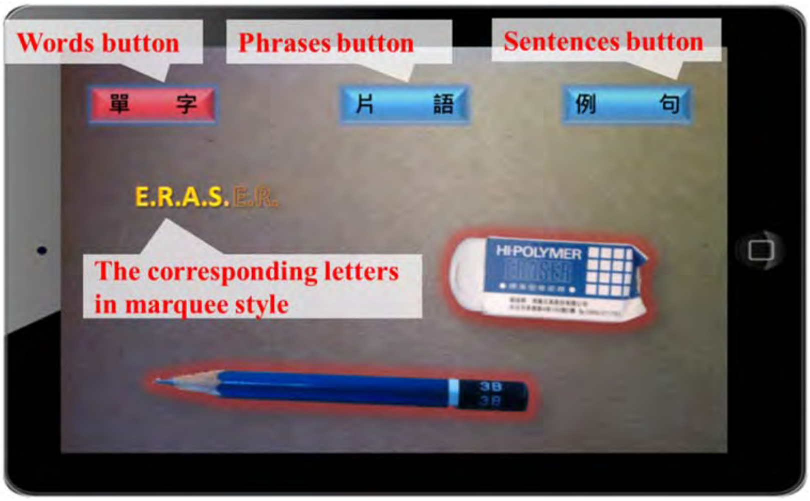


Figure 7. The function of “words” button.

If learners click the function button of "phrases", the teaching material of phrases will appear on the screen. Then, the phrase will be read out using English and the Chinese meaning will be explained using Chinese voice. In addition, the corresponding animation will appear. The system will give simple phrase in the easy mode (as shown in Figure 8) and give more difficult phrase in the advanced mode (as shown in Figure 9).

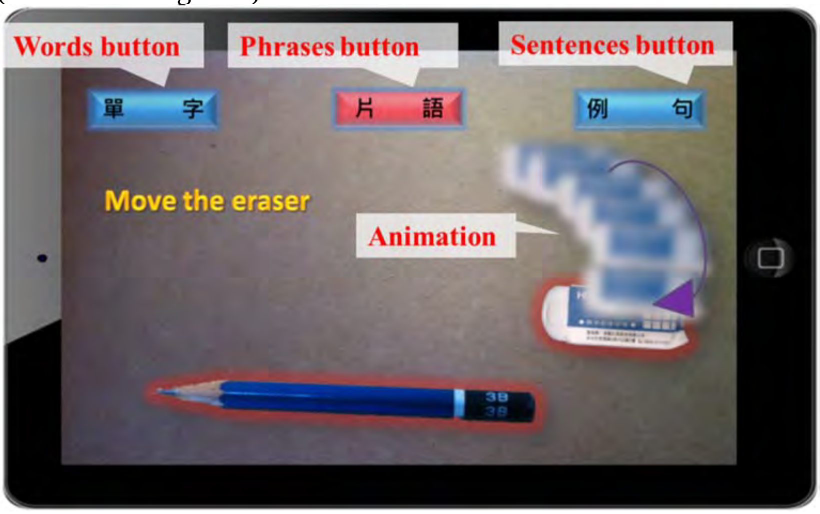


Figure 8. The function of “phrases” button in easy mode.

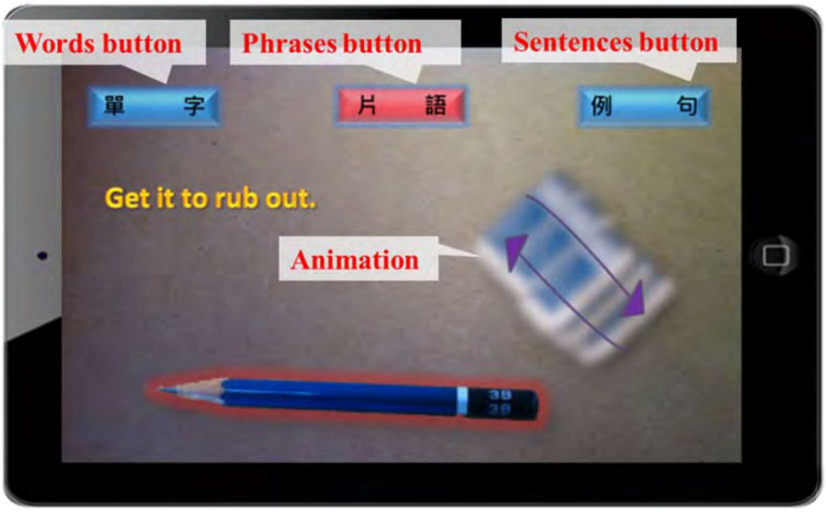


Figure 9. The function of “phrases” button in advanced mode.

If learners click the function button of "sentences", the teaching material of sentences will appear on the screen. Then, the sentence will be read out using English and the Chinese meaning will be explained using Chinese voice. In addition, the corresponding animation will appear. According the easy or advanced mode, the degree of difficulty of the sentence is also different, as shown in Figure 10.

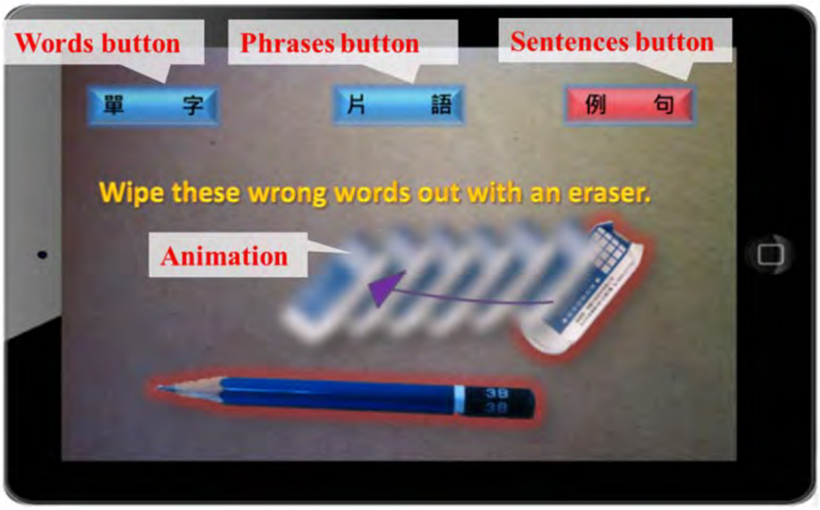


Figure 10. The function of “sentences” button.

4. Conclusions and Future Works

This study proposed a customized AR English learning system. The system provided 3 learning scopes (phrases, sentences, and related words) and 2 modes (the easy mode and the advanced mode). By AR technology and customized design, the learning motivation and effectiveness of learners are expected to be enhanced.

In the future, we will invite 2 teachers in an elementary school in the central region of Taiwan to provide 100 English vocabularies suitable for grade 5 students and to conduct experimental teaching. The experimental subjects are 100 students of 4 classes. All the operation process will be recorded in database, so that further personalized preference analysis can be conducted.

Acknowledgements

This study is supported in part by the National Science Council of the Republic of China under Contract No. MOST 103-2511-S-275-002-MY2.

References

Amir, H. B., & Vineet, R. K. (2012, November). *A framework for utilizing context-aware augmented reality visualization in engineering education*. Paper presented at the 12th International Conference on Construction Application of Virtual Reality, Taipei, Taiwan.

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*, 355-385.

Billinghurst, M. (2003). *Augemented reality in education.* Retrieved on Dec. 27, 2013 from http://www.newhorizons.org/strategies/technology/billinghurst.htm.

Billinghurst, M. (2003). *Augemented reality in education*. Retrieved December 27, 2013 from http://www.newhorizons.org/strategies/technology/billinghurst.htm.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, *18*(1), 32-42.

Chang, Y. J., Chen, C. H., Huang, W. T., & Huang, W. S. (2010, December). *Application of Augmented Reality on Implementation of the English Vocabulary Learning system*. Paper presented at the 16th Conference on Information Management & Practice (IMP 2010).

Chen, C. M., Tsai, Y. N. (2013). Interactive Augmented Reality System for Enhancing Library Instruction in Elementary Schools. *Museum News of Taipei Public Library, 30*(3), 29-49。

Chen, Y. H., Lien, C. J. & Lu, Y. L. (2009, May). *Influences of different learning preferences on learning achievement in E-learning*. Paper presented at the 13th Global Chinese Conference on Computers in Education (GCCCE 2009). Taipei, Taiwan.

Chen, Z. Y., Zhao, Z. Y., Liu, C. X., Lin, K. Z., & Lu, G. F. (2013, May). *Development and Design of Gear Unit in Game-based Interactive Teaching Materials Using Augmented Reality Technology*. Paper presented at the 17th Global Chinese Conference on Computers in Education (GCCCE 2013). Beijing, China.

Chu, H. C., & Lin, C.-W. (2013, September). *The development and application of a repertory grid-oriented ubiquitous augmented reality learning system*. Paper presented at the 2nd IIAI International Conference on Advanced Applied Informatics (IIAI AAI 2013) Conference, Kunibiki Messe, Matsue, Japan.

Dünser, A., & Hornecker, E. (2007). Lessons from an AR book study. *Proceedings of the First International Conference on Tangible and Embedded Interaction*, 179-192.

Fink, J. & Kobsa, A. (2000). A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web. *User Modeling and User-Adapted Interaction, 9*(34), 209-249.

Hsieh, M. C., Lin, H. C. (2009, September). *Employing augmented reality to promote English vocabulary learning*. Paper presented at the 5th Taiwan E-Learning Forum (TWELF 2009), Tainan, Taiwan.

Hwang, G. H., Lee, C. Y., Hwang, H. L., Huang, G. L., Lin, J. Y., & Cai, J. J. (2013). *Using augmented reality to assist an interactive multi-language learning system in an elementary school.* Proceedings of the 21st International Conference on Computers in Education. Bali, Indonesia.

Kuo, S. W. (2008, December). Preliminary study on the application of augmented reality in museum exhibition. *Technology Museum Review*, *12*(4), 25-37.

Lai, Q. L., Hwang G. J., Chen, H. L. (2012, May). *The learning effectiveness analysis of teaching imported augmented reality*. Paper presented at the 16th Global Chinese Conference on Computers in Education (GCCCE 2012). Kenting, Taiwan.

Landt, J. (2005). The history of RFID. *IEEE Potentials, 24*(4), 8-11.

Lin, H. C., Tang, K. W. & Peng, Y. Y. (2011). Effectiveness analysis of integrating augmented reality into learning domain of arts and humanities: A Perspective of user playfulness. *International Journal on Digital Learning Technology, 3*(3), 39-56.

Looi, C. K., Seow, P., Zhang, B., So, H. J., Chen, W., & Wong, L. H. (2010). Leveraging mobile technology for sustainable seamless learning: a research agenda. *British Journal of Educational Technology, 41*(2), 154-169. doi:10.1111/j.1467-8535.2008.00912.x.

Lu, C.M. (2012). *Application of Simple Quasi-markless Augmented Reality: A Case Study of English Conversation* (Unpublished master’s thesis). National Taipei University of Education, Taipei City.

Mayer, R. E. (2009). *Multimedia learning, second edition.* Cambridge University Press.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: a class of displays on the reality-virtuality continuum. *Proceedings of Society of Photo-Optical Instrumentation Engineers (SPIE), 2351*, 282-292.

Mitchell, T. J. F., Chen, S. Y. & Macredie, R. D. (2005). Cognitive Styles and Adaptive Web-based Learning. *Psychology of Education Review, 29*(1), 34-42.

Savignon, S. J. (1988). In second language acquisition/foreign language learning, nothing is more practical than a good theory. *Issue and Developments in English and Applied Linguistics*, *3*, 83-98.

Sejin, O., & Woontack, W. (2009, January). *CAMAR: Context-aware Mobile Augmented Reality in Smart Space*. Paper presented at the 3rd International Workshop on Ubiquitous Virtual Reality (IWUVR 2009), Adelaide.

Tsai, Y. S., Li, L. M., & Wu, A. H. (2011, November). *Augmented Reality Applications in Chinese Digital Teaching Materials: On Chinese Cubes*. Paper presented at the 7th Taiwan E-Learning Forum (TWELF 2011), Taipei, Taiwan.

Wernhuar, T., & Ou, K. L. (2012, March). *A study of campus butterfly ecology learning system based on augmented reality and mobile learning*. Paper presented at the 7th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technologies on Education (WMUTE 2012), Kagawa, Japan.

Wu, M. N., Li, M. H., Yao, Y. J., & Pai, T. F. (2012, October). *Interactive digital teaching for chemistry experiment using augmented reality*. Paper presented at the 8nd Taiwan E-Learning Forum (TWELF 2012). Tainan, Taiwan.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Development of a Multi-Device Data Structures Course Item Bank Practice System with Self-Regulated Learning Strategy on Bloom's Taxonomy of Educational Objectives

Gwo-Haur HWANG[a*], Jing-Fang CHEN[a], Yu-Ting SHIH[a], Yong-Sheng JHANG[a], Yi-Xuan LIN[a] & Yu-Syuan WANG[b]

[a] *Department of Information Networking and System Administration, Ling Tung University, Taiwan*

[b] *Department of Mathematics and Information Education, National Taipei University of Education, Taiwan*

*ghhwang@teamail.ltu.edu.tw

Abstract: With the rapid development of information technology, traditional paper-and-pencil testing is lacked for immediate analysis and feedback. Thus, it has been replaced by computerized item bank practice systems. Because of the popularity of wireless networks and multi-devices, developing an item bank practice system which can support on multi-device has become a new trend gradually. This makes learning become more convenient. In this paper, we classify items by Bloom's taxonomy of educational objectives, and found the weaknesses of learners in the practice in order to improve the weaknesses. At the same time, we use self-regulated learning strategy in this system, so that learners can set their learning goals. In this system, it provides self-monitoring, standard setting, evaluative judgment, self-appraisal, and affective self-reaction, so that learners can learn in the best environment. In the future, we will combine this system with data structures course and hope to improve learners' learning motivation and effectiveness.

Keywords: Multi-device system, data structures, item bank practice system, self-regulated learning, Bloom's taxonomy of educational objectives

1. Introduction

The traditional practice method is a paper-and-pencil testing, and it is replaced by computers. When learners encounter problems, they need teachers help, otherwise students cannot get the correct information immediately. With the rapid development of information technology and internet, computer online practice has become the new trend gradually. Tu (2003) developed an on-line assessment system, which combined a natural sciences course. The system can be repeated practicing and provide the feedback immediately. In addition, Lee, Tseng and Tsai (2003) and Chen, Chang and

Wang (2008) mentioned, because of the popularity of wireless networks and multi-devices, learners can learn online via multi-device to improve their learning effectiveness. Therefore, multi-device item banks practice come into being. For example, Kung, Huang and Chung (2007) developed a multi-device learning assessment system which combined class b of computer software application technicians. The results of their study showed that the system improved learners' learning effectiveness. However, these papers only developed multi-device item bank practice systems, but did not enhance leaners' weaknesses. Therefore, we use the Bloom's taxonomy of educational objectives to analyze the weaknesses of learners.

Bloom, Engelahar, Frust, Hill and Krathwohl (1956) proposed the "Bloom's Taxonomy of Educational Objectives", which was used by many educators and updated to the new version (Anderson & Krathwohl, 2001). The Bloom's Taxonomy of Educational Objectives has good effectiveness on designing items and diagnosing learners' weakness (Crowe, Dirks, & Wenderoth, 2008; Hwang, Chen, Loe, & Huang, 2013). Therefore, we import the taxonomy and wish let learners can understand their weaknesses, so that they can adjust the learning goal by themselves and improve their learning motivation. However, initiative is a critical factor for learners' learning effectiveness (Govaere Jan, de Kruif, & Valcke, 2012). Thus, we also import the self-regulated learning strategy in this system to improve learners' initiative.

Learners can set a goal to carry on self-monitoring, standard setting, evaluative judgment, self-appraisal, and affective self-reaction, which is spirit of self-regulated learning (Bandura, 1991), i.e., learners can learn through setting and adjusting goals. We hope to improve learners' learning motivation and effectiveness via self-regulated learning. However, if learners only depend on the behavior of initiative investment and self-monitoring without the effective adjusting strategies, learners' learning effectiveness cannot be improved (Zimmerman & Kitsantas, 1997). Therefore, the effective adjusting strategy is a critical factor. Wang (2011) indicated that adding the self-adjust learning mechanism into formative assessment and using the Peer-Driven Assessment Module (PDA) strategy can lead to learners' learning motivation. The formative assessment can be learners' learning basis. Learners can adjust their goals repeatedly. The experimental results indicated that learners who use the PDA-WATA (Web-based Assessment and Test Analysis) system are better than learners who use the N (None)-WATA system on initiative and learning effectiveness. Therefore, we also provide all learners complete rate ranking, which can motivate learners via comparing.

In the past, about the self-regulated learning, many scholars found that learners can achieve good learning effectiveness via the self-regulated learning strategy in mathematics (Hackett & Betz, 1989; Malpass, O'Neil, & Hocevar, 1999; Pajares & Miller, 1994; Parker, Marsh, Ciarrochi, Marshall, & Abduljabbar, 2014) and natural sciences (Betz & Hackett, 1983; Britner & Pajares, 2006; Chen & Usher, 2013) domains.

Therefore, according to above mentions, we develop a multi-device data structures course item bank practice system with self-regulated learning strategy on Bloom's taxonomy of educational objectives. The Bloom's taxonomy of educational objectives and the self-regulated learning strategy can help learners understanding their weaknesses to practice and give the appropriate feedback. At the same time, learners can use multi-device learning tools to learn without any limitation of space and time. We expect that learners can improve their learning motivation and effectiveness via this system.

2. Literature

2.1 Item Bank Practice System

Bunderson, Inouye, and Olsen (1989) indicated that computerized testing includes four steps of development: (1) computerized testing: the traditional paper-and-pencil testing is replaced by using computer to practice; (2) computerized adaptive testing: item response theory (IRT) is applied to conduct the computerized adaptive test (CAT); (3) continuous measurement: history records are added; (4) intelligent measurement: expert systems with artificial intelligence provide learning suggestion to learners. The first three steps are lacked for immediate feedback to carry on effective improvement in traditional teaching. However, with the popularity of internet, item bank practice systems are changed from single-computer practice to computer-online practice (Hwang, Chen, Huang, & Loe, 2013; Tu, 2003; Yo, Jan, & Li, 2011). It is an issue how to improve learners' learning effectiveness by item bank practice systems. In this paper, the system combines the history records of the third step and the expert system of the fourth step, and assists learners to proceed to effective learning. On the other hand, multi-device learning tools make learners learn conveniently.

2.2 Multi-Device Learning

Gay, Stefanone, Grace-Martin and Hembrooke (2001) used laptops as learning tools on the communication course and the computer science course. Using of wireless Internet technology, learners can discuss course anywhere. Thus, learning by mobile device is feasible. Guerreroa, Ochoaa and Collazosb (2010) built a learning system. In this system, learners can use PDAs to carry on the grammar practice online. Learners are divided into some groups can discuss courses immediately. Teachers can see the learners' answers and comments and reduce the time of marking learners' homework. The experimental indicated that more than 70% of the learners who used the system can improve the language grammar ability, and more than 86% of the learners thought the system is operated easily. The system can not only improve learners' learning motivation and effectiveness, but also reduce the load for teachers. With the rapid development of multi-device learning tools (smart phones, tablets, notebooks, PDAs and computers), learners have more choices. In the empirical researches, Chen et al. (2008) and Hwang et al. (2013) had established two learning websites which allowing learners to use multi-device learning tools for learning without limitation all the time. At the same time, the learners' portfolios are recorded on the websites, so that teachers can see the learning situation of learners. The experimental results indicated that the webs can improve the performance of learners. As the above literatures mentioned, we can know that using multi-device learning tools will make learners learn conveniently. However, it is important to this paper how to recognize learners for items understanding. Thus, we use the Bloom's taxonomy of educational objectives to find the learners' weaknesses and improve them.

2.3 Bloom's Taxonomy of Educational Objectives

Bloom et al. (1956) propose cognitive domain taxonomy of educational objectives as "Bloom's taxonomy of educational objectives", which includes six classes. From easy to difficult, the six classes included knowledge, comprehension, application, analysis, synthesis and evaluation. With the rapid development of educational psychology, Anderson and Krathwohl (2001) corrected the Bloom's taxonomy of educational objectives. This taxonomy includes knowledge dimension and cognitive process dimension. Knowledge dimension is subdivided into factual knowledge, conceptual knowledge, procedural knowledge and metacognitive knowledge. Cognitive process dimension is divided into six categories, i.e., remember, understand, apply, analyze, evaluate and create. Therefore, many scholars have used this taxonomy to classify items. Lan and Chern (2010) classified the university entrance exams in English reading questions by the Bloom's taxonomy of educational objectives, which can let teachers catch the emphasis to carry on teaching easily and let learners more understand the topic and the problem-solving elements. Therefore, in this paper, we use remember, understand, apply and analyze of cognitive process dimension to classify items of the data structures course. We hope to analyze learners' weaknesses effectively and let learners improve their learning motivation and effectiveness by the self-regulated learning strategy.

2.4 Self-Regulated Learning

Bandura (1977) indicated the self-regulated learning conception as self-efficacy includes the efficacy expectation and the outcome expectation. The efficacy expectation means to set goals and to do self-evaluation for these goals. The outcome expectation means that, if learners think that their ability can't reach their goals, they cannot keep their learning motivation. Therefore, we think that self-efficacy (self-regulated learning) is important to impact learners' learning motivation. Subsequently, Schunk and Zimmerman (1994) indicated that the self-regulated learning includes four steps, i.e., "self-evaluation and self-monitoring", "goal setting and strategy planning", "using and monitoring of strategies" and "monitoring of results of strategies". Learners can learn according to the learning pace by themselves via these four steps. In summary, all scholars think that setting goal and self-evaluation in self-regulated learning are very important. And Schunk and Zimmerman also proposed "using and monitoring of strategies" and "monitoring of results of strategies". In the self-regulated learning environment, learners can not only set goals and carry on self-evaluation, but also they can carry out self-adjustment via the learning results. In addition, Multon, Brown and Lent (1991) aim related 39 papers of self-regulated learning to carry on integrated analysis in the past. The results of their study showed that the self-regulated learning can improve learners learning effectiveness in different subject areas, and different assessment methods. Therefore, we develop a multi-device data structures course item bank practice system with self-regulated learning strategy on Bloom's taxonomy of educational objectives. We hope this system can improve learners' motivation and effectiveness of learning.

3. Development of Our System

3.1 System Architecture

In this paper, we use Windows Server 2008 to set up an Internet Information Services (IIS) web server and a Microsoft SQL database server, and develop all modules by Visual Studio 2010 ASP.NET C# language. This system is divided into two interfaces, i.e., the learner interface and the teacher interface.

The learner interface includes three modules, i.e., “goal management module”, “item bank practicing module” and “personal learning portfolio inquiry module”. Learners can carry on operating of three modules with multi-device learning tools, e.g., cellphones, tablets and computers. In order to implement the self-regulated learning strategy, “goal management module” can provide learners setting goals which include selecting range, Bloom's Taxonomy of Educational Objectives (remember, understand, apply and analyze), and passing score. “Item bank practicing module” can provide learners to carry on practice via selected range by learners, and it is based on the Bloom's taxonomy of educational objectives to display the weaknesses of learners. Learners can understand their weaknesses and adjust the goals in “goal management module”. “Personal learning portfolio inquiry module” can provide learners to inquiry the practice records in the past. On the other hand, the teacher interface includes three modules, i.e., “learners’ basic information management module”, “item bank management module” and “learning portfolio management module”. Teachers can operate these three modules by computers. “Learners’ basic information management module” can provide teachers to modify learners' basic information. “Item bank management module” can provide teachers to modify items and set the Bloom's Taxonomy of Educational Objectives (remember, understand, apply and analyze) of items. “Learning portfolio management module” can provide teachers to inquire the practice records of all learners’ learning situation. All modules can access “learners’ basic information database”, “item bank database” and “learning portfolio database”, as shown in Figure 1.

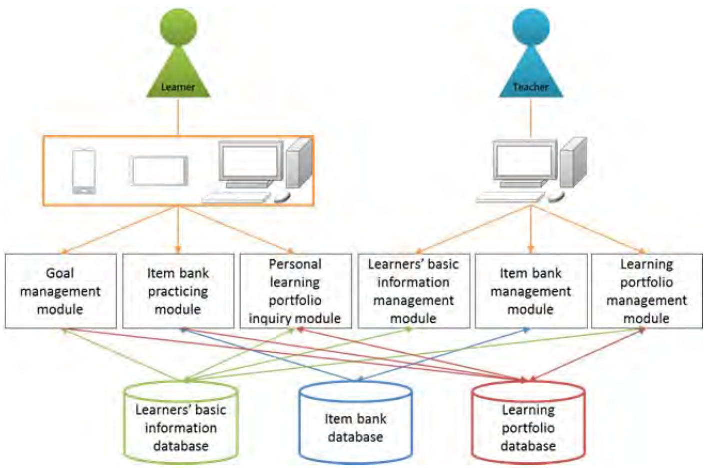


Figure 1. System architecture.

3.2 Operating interface

Learners can watch the currently complete rate ranking on the login snapshot of the system, as shown in Figure 2. After learners sign in the system, they can see the three buttons (“system homepage”, “goal management and item bank practicing” and “personal learning portfolio inquiry”) on left side, as shown in Figure 3.

Figure 2. Login interface.

Figure 3. Multi-device of system homepage.

When learners click the “goal management and item bank practicing” button, and enter “goal management module”, they can select chapters to practice by themselves. The items of every chapter are classified according to Bloom's Taxonomy of Educational Objectives (remember, understand, apply and analyze), and learners can do practice for their selected types. In addition, the system also provides learners a setting goal function. Learners can set the passing score of the current practice, as shown in Figure 4. Subsequently, learners can carry on practice in “item bank practicing module”. The practice interface of learners is shown in Figure 5.

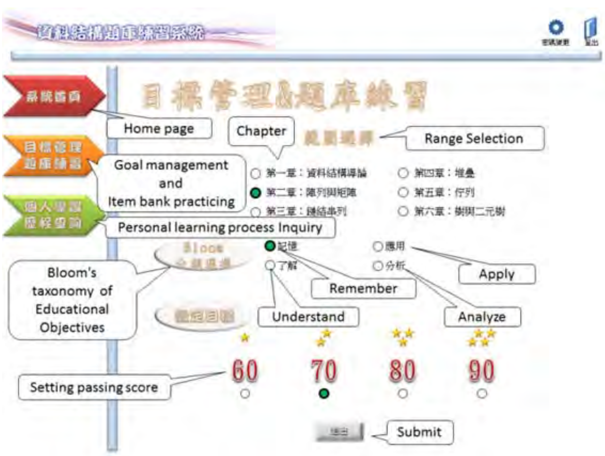


Figure 4. Item bank of setting goal snapshot.

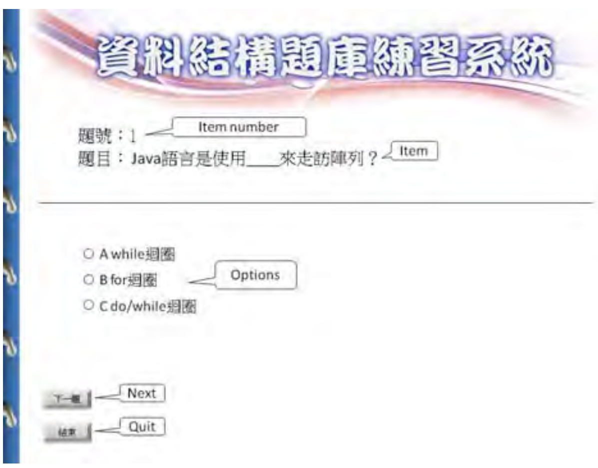


Figure 5. Practice snapshot.

End of the practice, the interface will change to the “goal management module”. Learners can see the correct rate, the courses complete rate and the feedback of practicing. There are two buttons at the bottom of “goal management module”, i.e., the “reset goal” button and the “practice again” button, are provided learners resetting goal to practice or practice again of the same goal, as shown in Figure 6, 7.

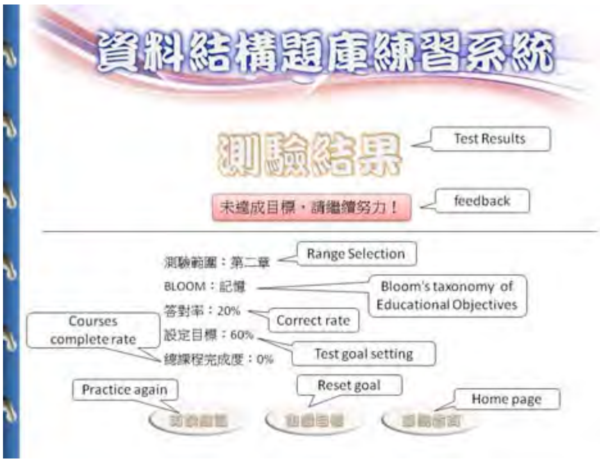


Figure 6. The snapshot of practice result without achievement.

Figure 7. The snapshot of practice result with achievement.

When learners click the “personal learning portfolio inquiry” button and enter the “personal learning portfolio inquiry module”, they can see the practice records quickly via the drop-down list above the interface. In the drop-down list, the red items mean that the practice is no-passed, and the black items mean passed, as shown in Figure 8. Learners can inquiry practice records after practicing, and they can watch all practice items or only wrong items, so that they can review and correct. Leaners can also watch items quickly via the drop-down list on the right hand of the interface. In the drop-down list, the red items mean that learners answer incorrectly, and the black items mean that leaners answer correctly right, as shown in Figure 9.

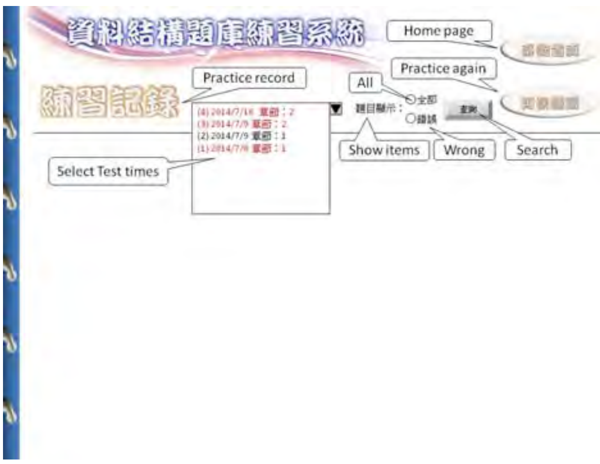


Figure 8. Practice records inquiry interface.

Figure 9. View practice items.

In the teacher interface of “learners’ basic information module”, which is provided teachers to inquiry the basic information of all learners. In “item bank management module”, teachers can click the "item bank management" button and carry on the management of items, as shown in Figure 10. In "learning portfolio management module", teachers can watch all learners' portfolios.

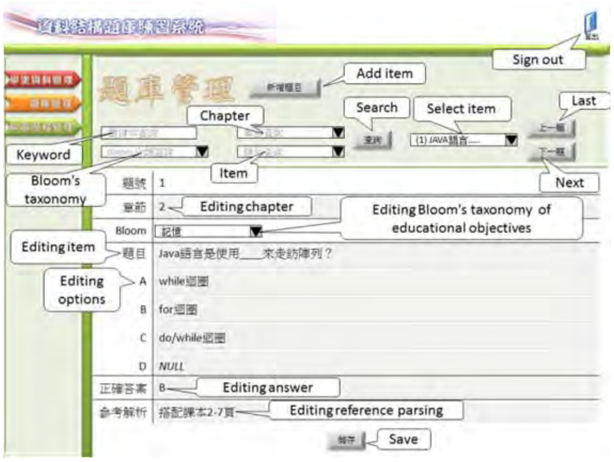


Figure 10. Item bank management.

4. Conclusion and future works

4.1 Conclusion

In this paper, we develop a multi-device data structures course item bank practice system with self-regulated learning strategy on Bloom's taxonomy of educational objectives. Learners can use multi-device to practice item bank. At the same time, they can set a goal to learn and budget time by themselves. In addition, every question in the item bank is set from a Data Structures teacher, and they are classified according to remember, understand, apply, and analyze. Therefore, this system can make learners clearly know what their weaknesses are. We expect learners can improve their motivation and effectiveness of learning via this system.

4.2 Future work

In the future, we will carry on an empirical research with two classes of attending data structure course which combine this system. The participants are about 99 learners in two classes. One class will be the experiment group which will use the self-regulated strategy. Learners of the experiment group can set a goal includes selecting range, Bloom's Taxonomy of Educational Objectives (remember, understand, apply and analyze), and passing score by themselves. The other class will be the control group which will use the non-self-regulated strategy. Learners of the control group will go by what the homework request of teacher to do practice. If they won't complete in time, they can't do practice. Their passing score will be set to 75. Two classes will proceed with pre-test, post-test and questionnaire (learning effectiveness, learning motivation, learning attitude, learning satisfaction and cognitive load). After the experiment, we will analyze to learners' learning motivation and learning effectiveness between the experimental group and the control group, as shown in Figure 11.

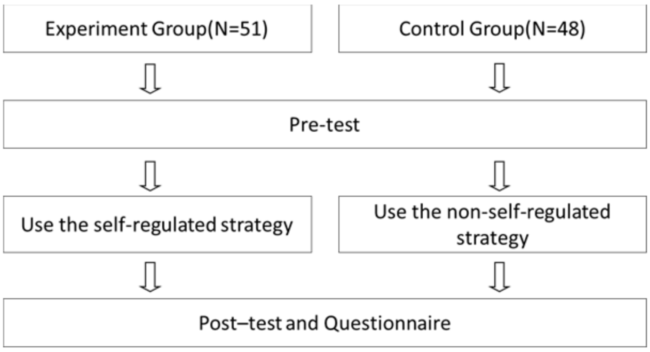


Figure 11. The flowchart of the experiment.

Acknowledgements

This study is sponsored by the Ministry of Science and Technology in Taiwan under the contracts no. MOST 103-2511-S-275-002-MY2.

References

Anderson, W., & Krathwohl, D. R. (Eds.) (2001). *A taxonomy for learning, teaching, and assessing: a revision of bloom's educational objectives*. NY: Longamn.

Bandura, A. (1977). Self-efficacy toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248-287.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational behavior, 23*(3), 329-345.

Bloom, B. S., Engelahar, M. D., Frust, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectiv*e. handbook1: cognitive domain. NY: David McKay.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching, 43*(5), 485-499.

Bunderson C. V., Inouye D. I., & Olsen J. B. (1989). The four generations of computerized educational measurement. In R. L. Linn (Ed.). *Educational measurement* (3rd ed., pp. 367-407). New York: American Council on Education.

Chen, G. D., Chang, C. K., & Wang, C. Y. (2008). Ubiquitous learning website: scaffold learners by mobile devices with information-aware techniques. *Computers & Education, 50*(1), 77-90.

Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences, 24*, 11-21.

Crowe, A., Dirks C., & Wenderoth M. P.(2008), Biology in Bloom: Implementing Bloom's Taxonomy to Enhance Student Learning in Biology. *CBE—Life Sciences Education, 7*(4), 368-381.

Gay, G., Stefanone, M., Grace-Martin, M. & Hembrooke, H. (2001). The effects of wireless computing in collaborative learning environments. *International Journal of Human–Computer Interaction*, *13*(2), 257–276.

Govaere Jan, L. J., de Kruif, A., & Valcke, M. (2012). Differential impact of unguided versus guided use of a multimedia introduction to equine obstetrics in veterinary education. *Computers & Education, 58*(4), 1076-1084.

Guerreroa, L. A., Ochoaa, S., & Collazosb, C. (2010). A mobile learning tool for improving grammar skills. *Procedia - Social and Behavioral Sciences, 2*(2), 1735-1739.

Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for research in Mathematics Education, 20*(3), 261-273.

Hwang, G. H., Chen, B., Huang, C. W., & Loe, D. M. (2013). Development of a personalized ubiquitous multi-device certification tutoring system based on bloom's taxonomy of educational objectives. *The 2013 Global Chinese Conference on Computers in Education (GCCCE 2013)*, Beijing. (In Chinese).

Hwang, G. H., Chen, B., Loe, D. M., & Huang, C. W. (2013). A cooperative learning certification examination tutoring system combining hamming distance with bloom's taxonomy of education objectives. *Association for Educational Communications and Technology - International Conference on the Frontier in e-Learning Research 2013*, Taichung.

Kung, H. Y., Huang, C. Y., & Chung, W. K. (2007). The assessment and effect analyses for mobile learning －the case of class b of computer software application technicians for the senior vocational school. *Journal of Computer Science and Application, 2*(2), 1-20. (In Chinese).

Lan, W. H., & Chern, C. L. (2010). Using revised bloom's taxonomy to analyze reading comprehension questions on the SAET and the DRET. *Contemporary Educational Research Quarterly, 18*(3), 165-206.

Lee, C. J., Tseng, K. H., & Tsai, H. L. (2003). Exploration of the mobile technology integrated into the primary school science and technology curriculum. *International conference for curriculum and instruction in technology education*, Kaohsiung. (In Chinese).

Malpass, J. R., O' Neil H. F. Jr., & Hocevar, D. (1999). Self-Regulation, goal orientation, Self-Efficacy, worry, and high-stakes math achievement for mathematically gifted high school students. *Roeper Review, 21*(4), 281-289.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of educational psychology, 86*(2), 193.

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology, 34*(1), 29-48.

Schunk, D. H., & Zimmerman, B. J. (1994). *Self-regulation of learning and performance: Issues and educational applications.* Hillsdale, NJ: Erlbaum.

Tu, C. C. (2003). *Research on the development and application of an on-line assessment system* (Master's thesis). Available from national digital library of theses and dissertations in Taiwan. (UMI No. 091NCUE0112009) (In Chinese).

Wang, T. H. (2011). Developing web-based assessment strategies for facilitating junior high school students to perform self-regulated learning in an e-learning environment. *Computers & Education, 57*(2), 1801-1812.

Yo, Y. C., Jan, C. W., & Li, B. J. (2011). The effect of online English proficiency practice tests on low-proficiency EFL learners' achievement. *Journal of China University of Science and Technology, 49,* 169-190.

Zimmerman, B. J., & Kitsantas, A. (1997). Development phases in self-regulation: shifting from process goals to outcome goals. *Journal of Educational Psychology, 89*, 29-36.

Students' Self-efficacy and Acceptance toward Context-Aware Ubiquitous Learning in Biology Education: A Case of Photosynthesis in Plant

Chuntanet NASARO[a] & Niwat SRISAWASDI[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
*niwsri@kku.ac.th

Abstract: Recent progression in wireless and sensor technologies has led to a new development of learning environment, called context-aware ubiquitous learning environment, which is able to transform students' learning experience into more authenticity. With the benefit of the context-aware u-learning environment, this study aims to investigate impact of sensor-based laboratory learning in biology, incorporated with predict-observe-explain (POE) pedagogy on students' perceptions of self-efficacy, perceived usefulness, and perceived ease of use. The 21-item Likert-scale questionnaire was administered to 44 eleventh-grade students both before and after the intervention. The results indicated that the POE-based ubiquitous learning in the real context of plant photosynthesis could promote students' self-efficacy and perceived ease of use on the learning. However, there was no significant effect on perceived usefulness of the intervention. As such, this study also provided the discussion for promoting the students' perceived usefulness after participating in the POE-based tablet.

Keywords: Ubiquitous learning, POE, ubiquitous computing, biology education, wireless networks

1. Introduction

Recent development in science and technology applications have influenced on learning process. In 21st century learning, computer and wireless network technologies have greatly affected the delivery of learning and capacitated people to convenient. The widespread of mobile devices, such as Tablet, PC, PDA and smart phone, has transformed learning modes from e-learning to m-learning (Minami, Morikawa and Aoyama, 2004). These technologies provide learning opportunity to everyone, anywhere and anytime. Mobile technology provides opportunities to support science learning both inside and outside classroom. With mobile technology, the learning environment can go with the student to outdoor, the laboratory and other beyond.

Chen, Chang and Wang (2008) suggested that ubiquitous learning is context sensitive anytime, anywhere learning using ubiquitous devices such as, Tablet, Web, PDA's in indoor, outdoor, individual and groups. Schroerder and Haskell (2011) described u-learning as social media plus mobile learning. Some of these characteristics are applicable to here and now learning (Huang, Chiu, Liu and Chen, 2011): (i) Hastening of learning need (on time suitable) (ii) Initiative of knowledge occupancy (providing information to students immediately) (iii) Situation of learning activity (flow of everyday activities) (iv) Context awareness (interaction controlled by context such as location, time, activity etc) (v) Self-regulated learning (students control their learning process)

Cause many parts of science topics difficult to understanding and learning in many area teaching by textbook which is low motivation and better to learn which should to learn in real phenomena. The prevalence of outdoor education has increased considerably in recent years (Bloom, Holden, Sawey and Weinburgh, 2010). Students can use those outdoor learning experiences to understand and establish new knowledge and concepts regarding the topic being studied (Auer, 2008, Upadhyay and DeFranco, 2008). On the other hand, Teachers can incorporate knowledge regarding ecology that students gained through outdoor learning into formal classroom instruction to improve students perception (Eick, 2012). Thus recent u-learning should be appropriately used to support authentic learning in real phenomena as science subject.

2. Purpose of the Study

Based on the abovementioned rational, this study aims to investigate students' perceptions delivered in sensor-based laboratory learning environment incorporated predict-observe-explain (POE) pedagogy for biology learning of photosynthesis. Specifically, the following questions were answered:

- Do the students engaged in sensor-based laboratory learning environment incorporated predict-observe-explain (POE) perform significantly better by students' perceptions of self-efficacy?
- Do the students engaged in sensor-based laboratory learning environment incorporated predict-observe-explain (POE) perform significantly better by students' perceptions of perceived usefulness?
- Do the students engaged in sensor-based laboratory learning environment incorporated predict-observe-explain (POE) perform significantly better by students' perceptions of perceived ease of use?

3. Literature review

3.1 A Context-aware Ubiquitous Learning

Recent progress in wireless and sensor technologies has lead to a new development of learning environments, called context-aware ubiquitous learning environment, which is able to sense the situation of learners and provide adaptive supports based on radio-frequency identification (RFID), wireless network, embedded handheld device, and database technologies. Many researchers have been investigating the development of such new learning environments. For examples, a context-aware mobile learning system was used as a sensing device for nursing training courses and they found that students' learning outcomes were notably improved by utilizing the mobile learning system (Chen and Huang, 2012). Shih, Chuang and Hwang (2010) study with fifth grade students with the inquiry-based mobile learning system. They investigate by pre- and post-test with observations and interview focus group. The finding showed significant positive results for students' learning. In an addition, e-library activity worksheets were developed by Hung, Lin and Hwang (2010) that helped students focus their outside ecology observation tasks, and results indicated that most students demonstrated substantial improvements and extended their inquiry skill. These evidence indicated potentials of the novel learning environment of context-aware ubiquitous learning in teaching and learning.

3.2 Inquiry- based Learning

Kuhn, Black, Keselman and Kaplan (2000) descript that inquiry-based learning is one of primarily pedagogical based on the investigation of questions. By the process of investigation and collection of science data, inquiry activities provide a valuable context for learners to acquire, clarify, and apply an understanding of science concepts (Edelson, Gordin and Pea, 1999). Furthermore, many researchers try to develop learners' investigation skills, data analysis and critical thinking using inquiry-based learning. They adopt activities related to the natural world to allow students to observe events and objects in the physical world from various facets, and to develop an understanding of how scientists explore the natural world (Hmelo-Silver, Duncan and Chinn, 2007). The advantages of inquiry learning are that it can lengthen the retention period of new knowledge, increase problem solving flexibility and creativity, and increase student learning motivation (Lord and Orkwiszewski, 2006). When inquiry learning is used in science subjects, it shows great potential for increasing students' understanding of scientific knowledge and their engagement in science.

POE strategy is a constructivist-based pedagogy and many researchers employed to facilitate learner's conceptual change process in inquiry-based learning. This kind of settings may provide a powerful learning environment for students where they have opportunities to construct scientific conceptual understanding that is durable over time (de Jong, 2005). POE can provide students to work on tasks collaboratively in group. Thus, it encourages a cooperative learning environment where students can share their knowledge and discuss with others in their group (Küçüközer, 2008, Tao and Gunstone, 1999).

4. Method

4.1 Study Participants

A total of 44 student-respondents in their eleventh grade, age ranging from 16 to 17 years in a local public school at the northeastern region of Thailand participated in this study. They were attending a biology course for basic education level. They have no experience yet using sensor-based laboratory in biology learning. This implied that they are heterogeneous before interacting with the experimental study.

4.2 Learning Materials and Activity

To engage student into context-aware ubiquitous learning, this study employed wireless microcomputer-based laboratory by Vernier and software technology. Vernier LabQuest-2 is a standalone interface used to collect sensor data with its built-in graphing and analysis application. The large, high-resolution touch screen makes it easy and intuitive to collect, analyze, and share data from experiments in class, field, and anyplace. Its wireless connectivity encourages students' collaboration and personalized learning. For this study, students were provided opportunity to conduct an investigation of photosynthesis in plant. With wireless features of the Vernier laboratory, the researchers designed students' laboratory experience with measurement of rate of oxygen (O_2) and carbon dioxide (CO_2) released by plants in the real context by gas sensors. The Vernier CO_2 and O_2 gas sensors used to measure gaseous carbon dioxide by monitoring the amount of infrared radiation absorbed by carbon dioxide molecules, and gaseous oxygen levels in a variety of environment, respectively. With the use of wireless LabQuest-2 data logger, experimental data were obtained, processed, and then shared to tablets in classroom via a server, as seen in Figure 1.

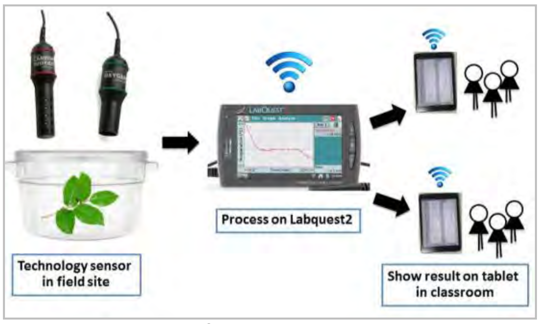


Figure 1. Materials of tablet-based laboratory learning.

To promote students' self-efficacy and acceptance toward the learning intervention, constructivist POE (predict-observe-explain) sequential learning steps were designed to foster biology learning by inquiry. In this study, instructor used 100 minutes class session to provide the u-learning experience for biology class. Student took overall 30 minutes for pre-test and post-test, and another 70 minutes for their POE learning process. Before performing the learning process, the instructor gave an orientation for working with the Vernier laboratory in 5 minutes and then 65 minutes for prediction (5 minutes), observation (30 minutes), and explanation (30 minutes). Figure 2 represented the prediction step based on an open-ended inquiry question provided by instructor, "what will happen if we measure rate of CO_2 and O_2 from plans outside the classroom?''. Then, students predict the graph of rate of O_2

and CO_2 when plant live in light and dark area on a work sheet. This warm-up activity was designed to provide the students with basic knowledge, and stimulate their motivation to learn science.

Figure 2. Prediction Stage. Students drawing graph on work sheet.

In the next learning step, a volunteer group of students was setting up by instructor and assign them to investigate rate of photosynthesis of plants using the Vernier gas sensor and LabQuest-2 data logger at the field for 30 minutes, regarding the variations in amount of carbon dioxide and oxygen. Figure 3 illustrates the observation activity in the real context.

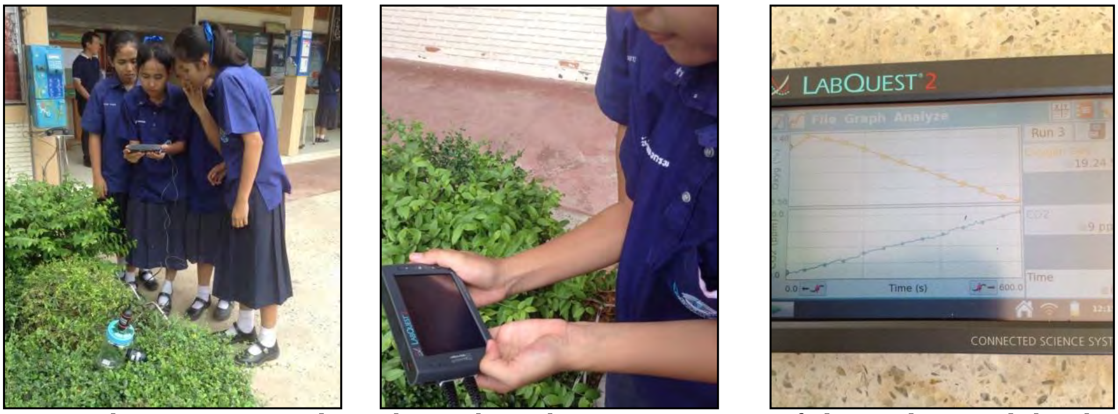


Figure 3. Observe Stage. Students observe by authentic measure rate of plants release and absorb O_2 and CO_2 in real area and process result by Labquest2 and share data on tablet in classroom by wireless connection.

At the same time with the observation, another student viewed the real-time experimental data of rate of plant photosynthesis through tablet connected wireless internet system in classroom. Figure 4 illustrates the use of tablet to monitor results of photosynthesis experimentation.

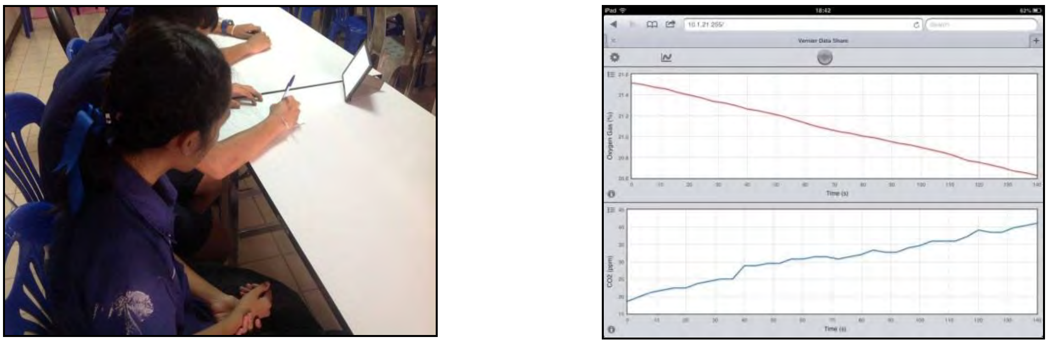


Figure 4. Observe Stage. The results were shared on tablet in classroom by wireless connection. Students compared their predictions and real observe.

For the last learning step of explanation, the volunteer group of students came back to the classroom and then the instructor conducted a forum discussion with peers to collaborative explain the plant photosynthesis phenomenon, emphasizing comparison of their prediction and explanation (See Figure 5).

Figure 5. Explain Stage. Students discuss and explain their prediction and real observe.

4.3 Instruments

A 21-item Likert-scale questionnaire was developed to use in this study for examining students' self-efficacy, perceived usefulness, and perceived ease of use toward the context-aware ubiquitous learning of plant photosynthesis. There were eight items of self-efficacy obtained from Wang and Hwang (2012), and six and seven items of perceived usefulness and perceived ease of use, respectively, from Hwang, Yang and Wang (2013). To develop a Thai version of the questionnaire, the original English version (See Table 1) was translated identically in Thai language. One expert was recruited to identify communication validity of the items. On each item, respondents were assigned to rate how much the respondent agree with into five scale, from 1-strongly disagree to 5-strongly agree). The reliability for self-efficacy, perceived usefulness, perceived ease of use, overall items was 0.76, 0.47, 0.73, and 0.85 (N=40), respectively.

Table 1: Items used to measure Self-efficacy, Perceived Usefulness and Perceived Ease of use.

Scale	Description of the scale	Example of items
Self-efficacy	Self-efficacy, or students' beliefs regarding their capability to execute actions necessary to achieve designated outcomes	• I am confident that I can learn the basic concepts well of this work. • I am confident that I can finish this work well.
Acceptance of the technology or learning approach	To better understand the students' perceptions of tablet-based laboratory learning approach, the students' ratings for the ''perceived usefulness,'' and ''perceived ease of use''	Perceived Usefulness • The learning mechanisms provided by the learning system smoothed the learning process. • The learning approach is more useful than the conventional computer-assisted learning approaches. Perceived Ease of use

		• I felt that the interface of the learning system was easy to use. • To sum up, the learning system adopted in this learning activity was easy to learn and use.

4.4 Data collection and Analysis

Figure 6 shows the procedure of the experiment. Before the learning activity, the students took the pre-test questionnaire. During the learning activity, stage 1 teacher provide inquiry question ‘‘what will happen if we measure rate of CO_2 and O_2 from plans now?’’ that students predict graph of result on activity sheet in each group (15 minutes). Then stage 2, students observe in field site by material learning (Vernier laboratory). The result will be show on tablet in each group inside classroom (30 minutes). Stage 3 student discuss in their group, compare the result of prediction and observation and explain (30 minutes)

After the learning activity, the students took the post-test questionnaire for comparing the pretest and the improvements learning.

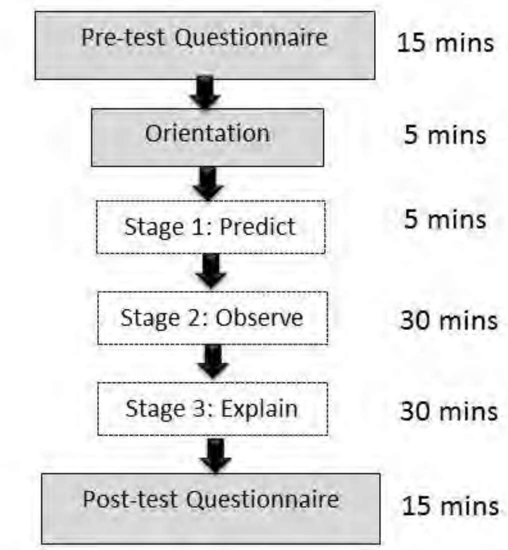


Figure 6. Diagram of experiment design.

The statistical data analysis techniques selected for this study were pair sample t-test. When there are measurement and nominal variables. The paired t-test in SPSS is used to compare the values of means from two related samples in a before and after learning.

5. Results and Discussion

In order to explore effects of the sensor-based laboratory learning environment incorporated predict-observe-explain (POE) for biology learning of photosynthesis, Table 2 shows results on students' self-efficacy, perceived usefulness, and perceived ease of use.

Table 2. Statistical results of paired t-test on students' perception scores.

Variable		N	Mean	SD	t
Self-efficacy	Pre-test	44	27.41	.555	2.846*
	Post-test	44	29.80	.665	

Perceived usefulness	Pre-test	44	21.93	.358	1.335
	Post-test	44	22.75	.512	
Perceived ease of use	Pre-test	44	23.68	.426	4.310*
	Post-test	44	26.39	.535	

*$p < 0.05$

The results show that student' self-efficacy and perceived ease of use delivered in the context-aware u-learning were statistically significant difference between pretest and posttest (t = 2.846, p<0.05 and t = 4.310, p<0.05, respectively). The result indicated also that the students' perceptions on self-efficacy and perceived ease of use showed significant improvement after participating with the context-aware u-learning. These means that learning with the sensor-based laboratory learning in biology, incorporated with predict-observe-explain (POE) pedagogy was not difficult and they were able to perform the investigation of plant photosynthesis with peers to accomplish the predetermined learning task in biology class. The result is consistent with the research findings that students with a strong sense of efficacy are more likely to challenge themselves with difficult tasks and be intrinsically motivated (Margolis and McCabe, 2006). In an addition, this finding consistent with Shih, Chuang and Hwang (2010), which reports that students felt more liberating and relaxing than learning in the classroom with the inquiry-based learning experience in field trip. Also, Hwang, Yang and Wang (2013) reported that students showed positive feedback after participating with the context-aware u-learning. However, the result indicated that there was no significant difference on students' perceived usefulness between pretest and posttest (t = 1.335, p<0.05). This result expressed that the students did not perceived the usefulness of the sensor-based laboratory learning incorporated with predict-observe-explain (POE) pedagogy for their biology learning. The reason might be that they did not familiar with the use of tablet technology in science learning and never use sensor-based laboratory learning for doing science before. Therefore, they might have no idea how this kind of learning environment would be benefits to them. According with Hwang, Tsai and Yang, (2008) descript that there are several levels of individualized guidance, to support learning with a context-aware u-learning environment, for naive learners, adaptive supports and guidance for real-world operations or observations should be provided for learners who have different backgrounds and experiences.

6. Conclusions

Although ubiquitous learning or context-aware learning environments is still available in Thailand. this paper is an effort to design, develop, and implement context-aware ubiquitous learning experience for biology learning in school science. This study demonstrates how instructions using context-aware ubiquitous learning with a sequential learning process of predict-observe-explain promoted students' perceptions on self-efficacy and perceived ease of use in teaching and learning about photosynthesis of plant phenomena. The results suggested that student can perform science investigation with sensor-based learning technology effectively. However, to promote students' perceived usefulness of the ubiquitous learning experience, the intervention in this study may need a revise to contribute fostering their perception of usefulness following Tsai, Tsai and Hwang (2011), which explored students' conceptions of context-aware ubiquitous learning and they found that students perceived the learning environment as followings: application of technology in the learning process; a convenient way to attain information to achieve their goals; a timely guide to apply the mobile devices to provide directions for learning; increase of knowledge; and a way of allowing them to engage in inquiry practices, such as allowing open-ended exploration for the learning topic.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Auer, M. (2008). Sensory perception, rationalism and outdoor environmental education. *International Research in Geographical and Environmental Education*, *17*(1), 6-12.

Bloom, M., Holden, M., Sawey, A., & Weinburgh, M. (2010). Promoting the use of outdoor learning spaces by K-12 inservice science teachers through an outdoor professional development experience. *Springer*, 97-110.

Chen, G., Chang, C., & Wang, C. (2008). Ubiquitous learning website: Scaffold learners by mobile devices with information-aware techniques. *Computers & Education*, *50*(1), 77-90.

Chen, C. C., & Huang, T. C. (2012). Learning in a u-museum: developing a context-aware ubiquitous learning environment. *Computers & Education*, *59*(3), 873–883.

de Jong, T. (2005). Engagement in learning: designing learning environments that are really used. In L. Verschaffel, E. De Corte, G. Kanselaar, & M. Valcke (Eds.), *Powerful environments for promoting deep conceptual and strategic learning* (pp. 45–59). Leuven, Belgium: Leuven University Press.

Edelson, DC, Gordin, D.N., & Pea, R.D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, *8*(3/4), 391-450.

Eick, C. (2012). Use of the outdoor classroom and nature-study to support science and literacy learning: a narrative case study of a third-grade classroom. *Journal of Science Teacher Education*, *23*(7), 789-803.

Hmelo-Silver, C., Duncan, R., & Chinn, C. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99-107.

Huang, Y., Chiu, P., Liu, T., & Chen, T. (2011). The design and implementation of a meaningful learning-based evaluation method for ubiquitous learning. *Computers & Education*, *57*(4), 2291-2302.

Hung, P.H., Lin, Y.F., & Hwang, G. J. (2010). Formative assessment design for PDA integrated ecology observation. *Educational Technology & Society*, *13*(3), 33–42.

Hwang, G., Tsai, C., & Yang, S. (2008). Criteria, strategies and research issues of context-aware ubiquitous learning. *Educational Technology & Society*, *11*(2), 81-91.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013), A concept map-embedded educational computer game for improving students' learning performance in natural science courses, *Computers & Education*, 69, 121-130.

Küçüközer, H. (2008). The effects of 3D computer modelling on conceptual change about seasons and phases of the moon. *Physics Education*, *43*(6), 632.

Kuhn, D., Black, J.B., Kesselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*. *18*, 495-523.

Lord, T. & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *The American Biology Teacher*, *68*(6), 342–345.

Margolis, H., & McCabe, P. (2006). Improving self-efficacy and motivation what to do, what to say. *Intervention in School and Clinic*, *41*(4), 218-227.

Minami, M., Morikawa, H., & Aoyama, T. (2004). The design of naming-based service composition system for ubiquitous computing applications, 304-312.

Schroeder, B., & Haskell, C. (2011). Micro-cycles: course design models for mobile learning. In I. L. Chen, & T. Kidd (Eds.), *Ubiquitous learning: A survey of applications, research, and trends.*

Shih, J.-L., Chuang, C.-W., & Hwang, G.-J. (2010). An inquiry-based mobile learning approach to enhancing social science. learning effectiveness. *Educational Technology & Society*, 13(4), 50-62.

Tao, P., & Gunstone, R. (1999). The process of conceptual change in force and motion during computer-supported physics instruction. *Journal of Research in Science Teaching*, *36*(7), 859-882.

Tsai, P., Tsai, C., & Hwang, G. (2011). College students' conceptions of context-aware ubiquitous learning: A phenomenographic analysis. *The Internet and Higher Education*, *14*(3), 137-141.

Upadhyay, B., & DeFranco, C. (2008). Elementary students' retention of environmental science knowledge: Connected science instruction versus direct instruction. *Journal of Elementary Science Education*, *20*(2), 23-37.

Wang, S. L., & Hwang, G. J. (2012). The role of collective efficacy, cognitive quality, and task cohesion in computer-supported collaborative learning. *Computers & Education*, *58*(2), 679-687.

How Self-Efficacy Affects Students' Performance and Pace in Self-Directed Learning with ICT

Andrew C.-C. LAO[a*], C.-L. HUANG[a], O. KU[b] & T.-W. CHAN[a]
[a]*Graduate Institute of Network Learning Technology, National Central University, Taiwan*
[b]*Research Center for Science and Technology for Learning*
*ccl.andrew@gmail.com

Abstract: Due to the fact that not only most Self-Directed Learning (SDL) studies remains discussed at the undergraduate or adult level, but also ICT in education has been positively proved as a compelling factor for the children, a SDL with ICT scenario is described in this paper. In SDL, with or without the assistance or guidance from the classroom teacher, students learn by their own during the learning activity, in which the students will identify their needs, set their goals, take learning missions, and reflect their learning outcome by reviewing their finished learning missions and performance. However, although students' self-efficacy affects how the students own the perception on self-management or goal accomplishments (Ormrod, 2006), limited discussion is found for exploring self-efficacy in SDL with ICT. Therefore, in this study, students are pre-categorized into high and low self-efficacy groups for exploring the effect of self-efficacy for SDL with ICT. As a result, the result shows that students' self-efficacy might not be a crucial factor that affect students' performance or pace, with only a slightly higher gain score on performance is found in this study.

Keywords: Self-directed learning, SDL, self-efficacy, ICT

1. Introduction

Self-Directed Learning (SDL) is believed as an essential andragogy in promoting students' individualization (Knowles, 1975). However, the original design of SDL lies within adults or adolescents (Gibbons, 2003; Knowles, Holton, Swanson, 2012), and consequently limited discussions were found at the elementary level. The reasons to the phenomenon, on the one hand, might be due to the children's maturity or cognitive engagement problems (Guthrie, 2004; Pressley, 2006; Taylor, Pearson, Peterson, & Rodriguez, 2003). On the other hand, the rationale for most of the public elementary curriculum restricts the flexibility for individual students' development, even though every student holds different capability in learning. Nevertheless, with regards to the aforementioned problems, compare to children in decades before, most of the children nowadays, could be more independent from the parents or teachers, and they could be more capable to learn individually (Glaubman, Glaubman, & Ofir, 1997; Philips & Stipek, 1993). Therefore, an increasing number of researchers pay attention to the development of SDL at the elementary level (Teo, Tan, Lee, Chai, Koh, Chen et al., 2010), but only a few studies discussed the ICT adoption in public schools with SDL. Followed by the master plan from the government, Teo et al. explored the possibilities of SDL with technology readiness in terms of learning goals, resource allocation, planning, monitoring, and reassessment of learning strategies.

Moreover, most studies have empirically proved the positive effects for students' cognition or affection by using ICT in education. In other words, ICT shows potential in public classrooms for education. Therefore, it could be concluded that the computer technology not only allowed students to be more engaged, but also led to a significant improvement on learning performance. In addition, since students' self-efficacy affects how the students manage to finish the learning goals for the learning activity (Kim, 2009), the effect for the self-efficacy in SDL should be considered. Hence, with regards to the importance of SDL and the positive effect of ICT, this study designs a SDL scenario that based on the rationale by various researchers related to SDL, and provides a preliminary analysis for the effect of different levels of self-efficacy that related to students' performance and pace.

2. Related Work

2.1 The research related to Self-Directed Learning

Self-Directed Learning (SDL) helps promote students individuality in learning and illustrate how the students' intention and action for learning individually (Tough, 1971; Knowles, 1975; Knowles, Holton & Swanson, 2012). The original idea for SDL could be traced to Tough's study, which was called "self-planned learning". In Tough's study, he found that most students or learners frequently applied the 3Ws (what, where, and when) in learning. The 3Ws could be considered as guidance for designing the SDL scenario, where students should be able to determine what to learn, where to start, and when to finish the leaning activity. Besides, researchers like Tough and Knowles both shared the similar concept in SDL, but Knowles had later popularized the term "Self-Directed Learning" and offered many learning resources for students and teachers to apply SDL in classrooms. For example, Knowles and his colleagues believed that, in SDL, students are responsible for their decisions in the learning activity, while students would set and develop their personal goals, revise and reflect their own pace or learning experience. In addition, Knowles (1975) also proposed six steps for the implementation of SDL in classrooms:

- Setting the classroom environment
- Learning needs
- Learning goals
- Identifying learning resources
- Applying suitable learning strategies
- Evaluating the learning outcome

The six steps presented above played an essential role for the adoption of SDL in the classrooms, because the steps emphasized on the transformation of the classroom learning from teacher-centered to student-centered, and focused on not only the knowledge acquisition, but also the personal reflection for the learning outcome. Therefore, followed by the SDL concept by Tough or Knowles, a researcher such as Gibbons (1994, 2002) provided specific criteria for SDL in practice. More specifically, to facilitate SDL in classrooms, from Gibbons' two studies, he suggested various criteria from two perspectives that are related to the development of the SDL framework, in which adolescents learned under the guidance by the school teachers. (See Table 1).

Table 1: Gibbons' two studies for facilitating SDL in classrooms.

Teacher's perspectives (Gibbons, 1994)	Students' perspectives (Gibbons, 2003)
● To let the students acquire knowledge from the school teacher ● To let the students learn how to teach oneself ● To let the students learn how to direct their learning individually	● To develop students' skill ● To achieve best performance with additional challenges ● To be self-managed ● To be self-motivated and be able to assess the learning outcome on their own

From the two perspectives mentioned in Table 1, the study by Gibbons (1994) emphasized on how to let the students acquire certain skills from the teacher, while his later study (Gibbons, 2003) highlighted the students' personal development, such as self-management or self-assessment. This could be referred that the weight for the feasibility of SDL in classrooms might be transformed from teachers' perspectives to the students', but we believed that there existed certain reasons that both the teacher and students' perspectives should be taken into consideration. For that reason, the design of this study would enhance the students' individuality and teacher's management for students' learning experience.

Accordingly, based on the idea of SDL by various researchers, a few studies discussed their application for SDL with ICT. For example, Robertson (2011) applied a blog-based system for students to learn in a self-directed way. In addition, followed by the master plan for education in the country, Teo et al. (2010) developed a questionnaire for SDL with ICT readiness by two pilot studies. Tan, Divaharan, Tan, & Cheah (2011) viewed SDL as a natural learning process, in which students' ownership, teachers' monitoring and management for students' learning were discussed at the elementary level, and they provided practical examples or experience for overcoming the adoption on SDL with ICT in education. Consequently, using the applications in SDL with ICT as references, the design of this study would be further discussed in Section 3.

2.2 Self-efficacy and SDL

Since students have to set their goals and strive for self-directness in SDL, students' self-efficacy, which indicates students' personal perception to success (Coutinho & Neuman, 2008), was believed as a positive relation with SDL (Kim, 2009). For the studies related to self-efficacy, most of the studies were derived from Bandura's Social Learning Theory (Bandura, 1997), in which self-efficacy refers to one's belief and capabilities for the learning outcome. Also, many researchers examined the difference for self-efficacy between high and low achievement students (Shell, Bruning, & Colvin, 1995; Stephenson, Poissant, & Dade, 1999; Stipek, 2002). In their studies, the researchers concluded that different levels of self-efficacy would result in a difference on performance expectations, because students with high self-efficacy would set a higher goal, task persistence, apply effective learning strategies, and time management than the students with low self-efficacy.

Besides, as the emergence of ICT in education show positive outcome for effective learning, studies discussed the self-efficacy with ICT. For example, Teo (2009) examined the technology acceptance by the pre-service teachers on self-efficacy, while the self-efficacy was believed to have a direct effect related to the ICT readiness, which might affect the learning outcome for the learning activity. The other study, Liaw (2008) also provided evidence that the self-efficacy is also a critical element that affects students' usage with ICT. Therefore, it is important to examine students' self-efficacy in SDL. With regards to the discussion above, this study will explore the effect of different levels of self-efficacy, in terms of students' pace and performance.

3. Design

The essence of SDL emphasizes on students' individualization. To this end, this study designs a SDL environment based on the design principles originally proposed by Knowles (1975). With regard to the discussion in Section 2, this study designs mechanisms for SDL used in classrooms. In figure 1, the three components: students, the teacher, and the system would be described as follows:

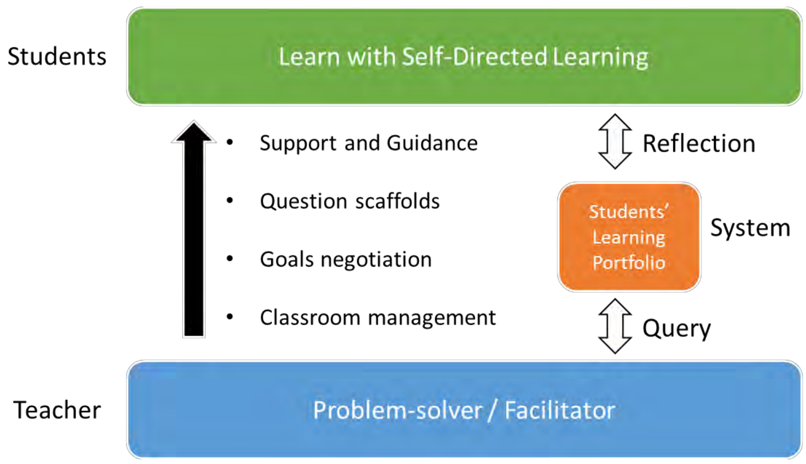


Figure 1. The design framework in this study.

3.1 Students: learn in a self-directed way

Students learn in SDL way, in which interactive learning missions (Chen, Liao, Cheng, Yeh, & Chan, 2012) would be provided as the major learning material during the learning activities. Most of the learning missions are designed based on the spiral curriculum, and these learning missions), along with some drill and practice games (Ku, Chen, Wu, Lao, & Chan, 2014), could be formulated by the concept understanding and procedural fluency (Kilpatrick, Swafford, & Findell, 2001). In the first 40-minute class of every week, students would first review their finished learning missions, then they have to decide how many learning missions and the performance should reach during the learning activity. During this learning process, students would be able to review their needs as well as reflect their learning experience, and set an appropriate goal.

3.2 Teacher: become a facilitator or a problem-solver

For the teacher, s/he is required to look into students' needs, and offer suitable directions for students to solve the problem in learning. For example, since most students could be able to work on their learning missions individually, a few students, especially low-achievers or those are easily distracted, are needed to be "coached" one-by-one. Moreover, according to the information provided by the system, the teacher would be able to support or provide guidance to the students, offer questions scaffolds that help resolve the mathematical problems, review and negotiate students' learning goal via the learning system, and maximizing the classroom management (such as offering bonus).

3.3 System: play as a learning portfolio for fulfilling the needs

The computer technology would make the SDL adoption in classrooms easier. Compares to the past, teachers or instructors were needed to review students' learning process one-by-one. This was quite time-consuming and the classroom teacher were barely to evaluate all students' learning carefully. Therefore, the design of this study takes the advantage of ICT, by offering real-time database inquiry for both students and the teachers. On the one hand, students can retrieve the status of their learning experience, which helps review and reflect their learning. On the other hand, teachers could be support by the system, in which the every students' learning status (such as accomplished learning missions, goals, and performance) can be accessed via the teachers' monitor.

4. Method

This section describes how the study is conducted. Students were randomly distributed in every class, and most of the students (> 90%) brought their parent-purchased tablet PC to the school. Every time when the mathematics class starts, students would turn on their table PCs and visit the web-based learning platform on their own. Nonetheless, due to the fact that most students own different self-efficacy in learning, it is needed to understand the different levels of self-efficacy among the students. To this end, since various studies applied the Motivated Strategies Learning Questionnaire (MSLQ, Pintrich, 1990) for understanding students' difference, this study adopted the self-efficacy for learning & performance scale in MSLQ. However, the original manuscript of MSLQ is used for adults and is in English, this study applied a localized version by Hsin, Lin, Yeh (2005), and the Cronbach alpha is 0.72.

Before the activity starts, the localized MSLQ was distributed to the students. Thirty-one effective samples were returned from two Grade 2 classes in a public school (See Table 2). As a result, to classify the effects of different levels of self-efficacy, students were categorized into two groups: high and low level of self-efficacy. Students with the above average score would be considered as high self-efficacy group, while the students with below average would be assigned to the low self-efficacy group. In addition, in order to understand the effect on self-efficacy, the performance and the pace were collected in both the pre-test and the post-test.

Table 2: Gibbons' two studies for facilitating SDL in classrooms.

Group of students	Class A	Class B	Total
HiSE group	10	6	16
LoSE group	7	8	15
Total	17	14	31

5. Results

In order to understand how students' self-efficacy affects students' performance and pace, this section describes preliminary analysis for both the pace and performance with independent t-test. To this end, a group of 31 students were divided into two groups based on the quantitative result by the MSLQ questionnaire (see Section 3), i.e. high and low self-efficacy groups (HiSE and LoSE), separately. For the students who own an above-average result, will be assigned to HiSE group (n=16), while the students who own a below-average result, will be the LoSE group (n=15). In this study, the result demonstrates that HiSE students do not have a significant difference with the LoSE students on both performance ($t = -1.105$, $p>.05$) and pace ($t = .984$, $p>.05$). It implies that the students' differences on different levels of self-efficacy would not be a factor that might affect students' learning performance and pace. The reason to this phenomenon might due to the limited flexibility for individuals in the public classrooms, and consequently no significant difference was found on students' performance and pace.

Table 3: The comparison of pace between high and low self-efficacy groups with independent t-test.

Group of students	Mean (number of missions)	SD	d.f.	t	p
HiSE	240.13	11.05	29	-1.105	.278
LoSE	245.67	9.20			

First, as stated in Table 3, the result for students' pace shows that the HiSE group has no significant difference with the LoSE group, and the HiSE group (mean: 240.13) is slightly lower than the LoSE group (mean: 245.67). This implies that HiSE students would not be beneficial by the difference on the self-efficacy with LoSE students. It might be due to the fact that some of the students in LoSE group have reached the maximum pace (eight students, the maximum pace: 256), and while most of the students in HiSE group is comparatively slower. Second, for the performance (See Table 4), no significant difference is found between the two groups of students. This implied that students with high self-efficacy would not result a difference with the low self-efficacy students. Nevertheless, we notice that an increase on the gain score for the performance of HiSE students, but a decrease for the performance of LoSE students. It could be explained that although no significant difference is found for the gain score, students with different levels of self-efficacy might lead to a slightly difference for the learning performance.

Besides, the average pace and performance in the samples for high self-efficacy have a slightly lower score than the low self-efficacy students. We believed that it might be due to the ceiling effects or some outliners existed among the students.

Table 4: The comparison of performance between high and low self-efficacy groups with independent t-test.

Group of students	Mid-term (Mean)	Final (Mean)	Gain score (Mean)	s.d.	d.f.	t	p
HiSE	86.50	87.06	0.56	4.163	29	.984	.333
LoSE	90.73	89.87	-0.81	3.907			

6. Conclusion

With regard to the emerging importance for SDL in the education, students learning capabilities on different perspectives (such as individual differences, learning habit) should be considered. Therefore, this study provides a preliminary exploration on the self-efficacy for both the performance and the pace in the SDL environment. Hence, in order to understand the effects between students with different levels of self-efficacy, this study addresses a SDL scenario into the public classrooms in Taiwan. As a result, students with different levels of self-efficacy did not show significant differences on both performance and pace. This indicates that the self-efficacy would not be an index for understanding students' difference for learning in SDL. However, this study is only a preliminary study and analyze for self-efficacy in SDL. Additional exploration should be carefully addressed. For example, the sample size of this study is too small (HiSE group = 16, LoSE group = 15), and additional samples are needed to be done. Also, Further, this study provides a preliminary evidence and analysis on self-efficacy for SDL with ICT, and could be a reference on the adoption of SDL with ICT in education.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (MOST 101-2511-S-008 -016 -MY3, MOST 100-2511-S-008 -013 -MY3) and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Chen, Z. H., Liao, C. C. Y., Cheng, H. N. H., Yeh, C. Y. C., & Chan, T. W. (2012). Influence of game quests on pupils' enjoyment and goal-pursuing in math learning. *Educational Technology & Society*, 15(2), 317-327.

Coutinho, S. A., & Neuman, G. (2008). A model of metacognition, achievement goal orientation, learning style and self-efficacy. *Learning Environments Research*, *11*(2), 131-151.

Gibbons, M. (1994). *The Self-directed Learning Sourcebook: Ideas, Methods and Materials for Teaching SDL*. Hawker Brownlow Education.

Gibbons, M. (2003). *The self-directed learning handbook: Challenging adolescent students to excel*. Wiley.com.

Glaubman, R., Glaubman, H., & Ofir, L. (1997). Effects of self-directed learning, story comprehension, and self-questioning in kindergarten. *The Journal of Educational Research*, *90*(6), 361-374.

Guthrie, R. V. (2004). *Even the rat was white: A historical view of psychology* . Pearson Education.

Hsin, C.-I., Lin, S. S.-J., Yeh, C.-C. (2005). Pilot analysis of motivational strategies in learning Calculus among Technology College students Motivational Strategies in Learning Calculus, *Journal of National University of Tainan, 39*, 65-82.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding+ it up: Helping children learn mathematics*. National Academies Press.

Kim, Y. H. (2009). Relationship of self efficacy, self-directedness and practice satisfaction to clinical practice education in nursing students. *Journal of Korean Academy of Fundamentals of Nursing*, *16*(3), 307-315.

Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Englewood Cliffs: Prentice Hall/Cambridge.

Knowles, M. S., Holton III, E. F., & Swanson, R. A. (2012). *The adult learner*. Routledge.

Ku, O., Chen, S. Y., Wu, D. H., Lao, A. C. C., and Chan, T. W. (2014). The effects of game-based learning on mathematical confidence and performance: High ability vs. low ability. *Educational Technology & Society*.

Liaw, S. S. (2008). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, *51*(2), 864-873.

Ormrod, J. E. (2006). *Educational psychology: Developing learners (5th ed.).* Upper Saddle River, N.J.: Pearson/Merrill Prentice Hall.

Phillips, D. & Stipek, D. (1993). Early formal schooling: Are we promoting achievement or anxiety?. *Applied and Preventive Psychology*, *2*(3), 141-150.

Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).

Pressley, M. (2006). *Reading instruction that works: The case for balanced teaching*. Guilford Press.

Robertson, J. (2011). The educational affordances of blogs for self-directed learning. *Computers & Education*, *57*(2), 1628-1644.

Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology*, *87*(3), 386.

Stephenson, R., Poissant, H., & Dade, M. O. (1999). *Third Graders' Self-Regulation and Self Efficacy in a Concept Formation Task: Differences between Low and High Achievers.* ERIC Clearinghouse.

Stipek, D. J. (2002). *Motivation to learn: From theory to practice (4th edition).* Needham Heights, MA: Allyn & Bacon.

Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal*, 3-28.

Tan, S.C., Divaharan, S., Tan, L.L.W, & Cheah, H.M. (2011). *Self-directed Learning with ICT: Theory, Practice and Assessment*(PP. 62). Singapore: Educational Technology Division, Ministry of Education.

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302-312.

Teo, T., Tan, S. C., Lee, C. B., Chai, C. S., Koh, J. H. L., Chen, W. L., & Cheah, H. M. (2010). The Self-Directed Learning with Technology Scale (SDLTS) for young students: An initial development and validation. *Computers and Education, 55*(4), 1764-1771

Tough, A. (1979). *The adult's learning projects: A fresh approach to theory and practice in adult learning*. Toronto: Ontario Institute for Studies in Education.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Effects of Game-based Peer Response on Writing Quality: High-ability vs. Low-ability

Jen-Hang WANG[a*], Sherry Y. CHEN[c], Oskar KU[b] & Tak-Wai CHAN[c]
[a]*Department of Computer Science and Information Engineering, National Central University, Taiwan*
[b]*Research Center for Science and Technology for Learning, National Central University, Taiwan*
[c]*Graduate Institute of Networking Learning Technology, National Central University, Taiwan*
*harry@CL.ncu.edu.tw

Abstract: In this study, we proposed a game-based peer response to enhance student writing. In addition, we also examined how students with different levels of ability react to the game-based peer response. The results revealed that such a game-based peer response could enhance students' writing quality and narrow the gap between the high-ability students and the low-ability students. Moreover, the effects of the game-based peer response on the improvement of high-ability students' writing quality was limited and did not reach to a significant level. Conversely, the game-based peer response was able to significantly improve the writing quality of low-ability students, regardless the overall quality or each individual aspect of writing quality.

Keywords: Game-based learning, individual differences, peer response

1. Introduction

Peer response (Elbow, 1973; DiPardo & Freedman, 1988) is also known as peer review, which refers to a collaborative activity, in which learners work together to improve the quality of their works by providing comments for each other. Recently, a lot of researchers investigated the effects of peer response and found its benefits can be classified into four aspects: social, cognitive, affective and linguistic (Rollinson, 2005; Min, 2006). Regarding the social aspect, negotiations used in the process of peer response could enhance students' communication and collaboration skills (Mendonça & Johnson, 1994). Regarding the cognitive aspect, peer response could not only facilitate students to develop critical and analytical skills for writing (Stoddard & MacArthur, 1993), but also make them have a greater awareness of audience (Lockhart & Ng, 1993). Regarding the affective aspect, peer response could help students reduce apprehension and increase confidence (Leki, 1990) and develop a greater sense of the ownership of the work (Tsui & Ng, 2000). Regarding the linguistic aspect, peer response could help students gain more new ideas and different points of view (Lockhart & Ng, 1993) and improve the quality of their works (Cho & Schunn, 2007).

Due to the aforementioned benefits, peer response has been widely applied to enhance student writing. For example, Sims (2001) attempted to use peer comments to improve children's expressive writing, and found that peer response enhanced students' writing fluency. Subsequently, Boscolo and Ascorti (2004) attempted to apply peer response to assist children to improve the clarity of their narrative writing. They found that peer response fostered students' abilities to detect information gaps or inconsistencies in writing. Additionally, Tuzi (2004) also used electronic peer feedback to support the revisions of academic writing. He found that electronic peer feedback assisted students to understand how to structure an essay.

The aforementioned studies demonstrated the effectiveness of peer response. Nevertheless, all of these peer response approaches are still implemented in a traditional education context. As suggested by Tüzün, Yılmaz-Soylu, Karakuş, İnal, & Kızılkaya (2009), students in a traditional educational context may have low motivation. In particular, young students have a limited attention span (Moreno-Ger, Martinez-Ortiz, Sierra, & Fernandez-Manjon, 2008). Therefore, there is a need to use a mechanism that can catch students' attention and increase their motivation. Among various mechanisms, digital games have transformed the way people learn and make learners have enjoyable experience (Marsh, 2011). Furthermore, some researchers found that digital games can motivate learners because they raise curiosity and allow learners to be in control of their own learning (Dickey, 2007; Huizenga, Admiraal, Akkerman, & ten Dam, 2009; Papastergiou, 2009). This may be the reason why game-based learning (GBL) emerged in the past ten years.

GBL possesses many positive effects on student learning (Pivec, 2007). In particular, previous research found that GBL could enhance students' learning motivation. For example, a study by Liu and Chu (2010) compared GBL and non-GBL in ubiquitous context. The results demonstrated that students with ubiquitous games could have better learning motivation than those with a non-gamed method. Thus, game-based learning can be a potential approach to address the problems of students' low motivation and short attention span. To this end, this study attempts to develop game-based peer response by incorporating GBL into peer response.

However, it is unknown whether such game-based peer response can be appreciated by all learners. This is due to the fact that game-based learning includes a variety of multimedia elements, which may cause cognitive overload. In other words, the game-based peer response delivers feedback via multiple information sources, which may increase students' cognitive load (Fried, 2008). In particular, learners are diverse so not all of learners have such a capacity to handle cognitive overload, which is usually happened when it is beyond the learners' capacity (Ang, Zaphiris, & Mahmood, 2007). Accordingly, there is a need to consider whether all learners can cope with such cognitive load. In other words, individual differences become an important issue. Among various individual differences, the diversity in learning abilities greatly affect students' perceptions (Cheng, Lam, & Chan, 2008), which, in turn, will influence their learning outcomes. Thus, such ability differences may affect how students react to this game-based peer response. Therefore, there is a need to examine the impacts of students' abilities on their reactions to the proposed game-based peer response.

To this end, the aims of this study are two folded. One is to develop game-based peer response while the other is to examine the effects of the ability differences on students' reactions to the game-based peer response. To correspond to the aforementioned two aims, two research questions are examined in this study:

(a) What are the effects of the game-based peer response on students' writing quality?
(b) How do high-ability students and low-ability students react differently to the game-based peer response?

The answers to these two research questions can contribute to develop a deep understanding of how to undertake game-based peer response that can accommodate students' individual differences. By doing so, both high- and low-ability students can benefit from game-based peer response.

2. Methodology Design

This study was conducted in an elementary classroom. To correspond to the aforementioned research questions, an empirical study was conducted to evaluate the effects of game-based peer response. The details are described in this section, including the implementation of the game-based peer response, participants, a pre-test and a post-test, pedagogical activities, measurement of writing quality, and data analysis.

2.1 The Implementation of Game-based Peer Response

We developed a game-based peer response, where peer response was conducted with various game elements, including game activities and rewards. Regarding game activities, participants need to play as a head of a publisher and manage their own publisher by completing various game activities required by different departments of the publisher (Figure 1), such as editing drafts, giving feedback to their peers' works, evaluating feedback received, revising their own work, and publishing and promoting their completed works to other peers (Figure 2). The purpose of such a series of game activities was to extend students' attention span in learning, and, in turn, facilitate themselves to complete target learning tasks. This is because each of the aforementioned game activities is associated with a target task, which has a clear sub-goal to be achieved. In other words, a complex writing process was decomposed into a series of tasks. When students complete a target task, they also achieve its sub-goal and can move to pursue the next sub-goal. Through the accumulation of these sub-goals, the ultimate goal is, accordingly, reached.

Figure 1. A Snapshot of the Game-based Peer Response (the Publisher)

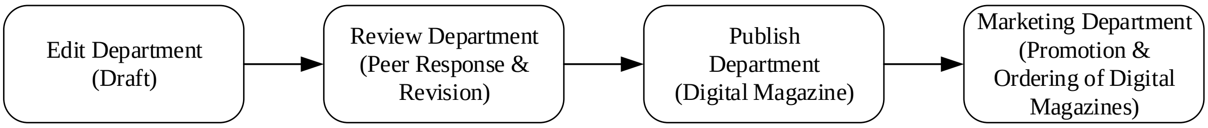


Figure 2. The Main Activities of the Game-based Peer Response

Further to a series of game activities, various reward mechanisms were also employed to extend students' attention span and increase their motivation and participation in learning activities, including virtual currency, leaderboards, and trophies. The virtual currency was used to reward students' behavior for their level of participation and hard working in each game activity and the earned virtual currency can be used to buy marketing tools for promoting users' published magazines or to order other peers' published ones. On the other hand, leaderboards and trophies were used in the responding tasks of the Review department. To ensure the responsibility and motivation of users in participating in responding activities, leaderboards and trophies were introduced as levels of social reputation. More specifically, students would receive various levels of trophies based on their ranking on the leaderboard for their feedback performance. In summary, the intention of developing this game context was to enhance students' engagement in peer response so that their writing quality can be improved.

2.2 Participants

Twenty-one third-grade elementary students participated in this study. They were aged 8-9 years and were recruited from the same class. In other words, they were taught by the same curriculum and were given the same writing assignments and instruction. Furthermore, they had no experience of peer response prior to taking parting in this study.

2.3 Pre-test and Post-test

To evaluate students' writing ability, participants needed to take a pre-test and a post-test at the beginning and the end of the experiment, respectively. More specifically, the pre-test was applied to examine students' prior writing abilities while the post-test was employed to assess their writing ability after taking the game-based peer response. The pre-test and post-test were represented as a composition test, where participants were given a theme-based topic and they were required to complete a narrative composition within an 80-minute period.

2.4 Pedagogical Activities

To help participants know how to act as the providers and recipients of peer response, instructions were given to them based on two guidelines: (a) the interaction between readers and writers proposed by Elbow (1973) and (b) the guidance for peer response proposed by Hansen and Liu (2005). By doing so, the participants could undertake peer response with proper attitudes and procedures. Moreover, they were introduced how to complete writing and responding tasks with the gamed-based peer response.

Then, all participants were evenly re-allocated into small peer response groups of four or five students. Furthermore, participants conducted a series of activities: (a) to receive writing instruction from their teacher, (b) to make drafts individually with a tablet laptop, (c) to receive feedback instruction from their teacher, (d) to read group-mates' drafts and give feedback, (e) to evaluate how useful the feedback received, (f) to revise their own drafts based on the feedback from their peers, (g) to collect completed works and publish them as a digital publication, (h) to make promotion for their published works to have opportunities to present their works to more audience. During this process, students would obtain various rewards, depending on their performance in the aforementioned target tasks.

2.5 Measurement of Writing Quality

An assessment mechanism proposed by Yang, Ko, and Chung (2005) was adopted to assess students' writing quality because it was designed for elementary students, and then appropriate for our participants. This assessment covers five items: (1) elegant words, (2) clear paragraph, (3) coherence, (4) title consistence, and (5) new & original ideas. A five-point rating scale was used for each item. Thus, the total score for a composition was between the lowest score (5 points) and the highest score (25 points). Two raters were recruited to independently evaluate the participants' writing quality so each student's final score was defined based on the mean of scores by the raters, of which the inter-rater reliability was found to be Kappa = 0.728 ($p < .001$). In other words, a substantial level for the measure of agreement between the raters was reached.

2.6 Data Analysis

In this study, we investigated how high-ability students and low-ability students react differently to the game-based peer response. Therefore, students were classified into the high-ability and low-ability students based on the mean scores of the pre-test. Then, an Independent Samples t-test, which is suitable to test "the difference between the means of two independent groups" (Howell, 2007), was used to examine differences between the high-ability students and the low-ability students for the pre-test scores and the post-test scores. On the other hand, Paired Samples t-test, which is appropriate to test the difference between the means of paired samples (Howell, 2007), was employed to inspect differences between the pre-test scores and the post-test scores for the high-ability students and the low-ability students. These aforementioned analyses were undertaken by using SPSS for Windows (version 16.0). The level of significance was set at $p < .05$ for all comparisons.

3. Results and Discussion

This section is divided into three subsections. The first subsection is to present the writing quality of the high ability students and the low-ability students in the pre-test and the post-test. The second subsection is to describe how the high-ability students and the low-ability students performed differently before and after interacting with the game-based peer response. Then, the third subsection is to present a discussion of why the high-ability students and low-ability students reacted differently to the game-based peer response.

3.1 High-ability students vs. Low-ability students

3.1.1 Pre-test Scores

Regarding the pre-test scores, significant differences existed between the scores from the high-ability students and those from the low-ability students. More specifically, the former significantly outperformed the latter, not only in the aspect of overall quality, but also in the aspects of elegant words, clear paragraph, coherence, and new & original ideas (Table 1). These findings suggested that the high-ability students had a better lever of prior writing ability than the low-ability students. Accordingly, the former were more capable to use appropriate words and phrases and organize their paragraph structures, and also were better able to express thoughts in a distinctive ways.

On the other hand, the high-ability students did not perform significantly differently from the low-ability students, in the aspect of title consistence ($p > .05$). This finding implied that both of them had a similar level of prior ability in this aspect.

Table 1: Writing quality of the pre-test (high-ability vs. low-ability).

	HA (n = 11)	LA (n = 10)	Independent samples t-test
	M (SD)	M (SD)	*t*
Overall	15.27 (1.90)	11.50 (.85)	5.96***
Elegant words	2.73 (.79)	1.8 (.42)	3.41**
Clear paragraph	3.09 (.70)	2.20 (.63)	3.05**
Coherence	2.82 (.87)	2.10 (.32)	2.55*
Title consistence	4.00 (.00)	3.90 (.57)	.56
New & original ideas	2.64 (.81)	1.50 (.53)	3.77**

* $p < .05$, ** $p < .01$, *** $p < .001$

3.1.2 Post-test Scores

Regarding the post-test scores, no significant differences ($p > .05$) existed between the scores from the high-ability students and those from the low-ability students. This implied that both groups had a similar level of posterior writing ability, regardless the overall quality or the other aspects aforementioned after they undertook the game-based peer response (Table 2). In other words, the gap between these two groups was minimized.

However, it is still unclear why the aforementioned gap between the high-ability students and the low-ability students has been narrowed or who can benefit from this game-based peer response. Therefore, it is needed to further explore how the high-ability students and the low-ability students react differently to this game-based peer response. To address this issue, we conducted a comparison between students' pre-test scores and post-test scores, of which the results are presented in the subsection below.

Table 2: Writing quality of the post-test (high-ability vs. low-ability).

	HA (n = 11)	LA (n = 10)	Independent samples t-test
	M (SD)	M (SD)	*t*
Overall	16.18 (2.23)	16.80 (2.49)	-.60
Elegant words	2.73 (.47)	2.70 (.48)	.13
Clear paragraph	2.91 (.83)	3.40 (.84)	-1.34
Coherence	3.18 (.87)	3.10 (.74)	.23
Title consistence	4.55 (.69)	4.60 (.52)	-.20
New & original ideas	2.82 (.75)	3.00 (.67)	-.58

3.2 Pre-test vs. Post-test

3.2.1 High-ability students

Regarding how high-ability students react to the game-based peer response, small differences were found between their post-test scores and pre-test scores in the aspects of writing quality (Table 3).

However, such differences did not reach a statistically significant level ($p > .05$), apart from title consistence ($p < .05$). In other words, the writing quality of high-ability students was not greatly improved. These findings implied that the effect of the game-based peer response on high-ability students' writing ability was not obvious enough to improve their writing quality.

Table 3: Writing quality of the pre-test and the post-test (high-ability students).

	Post-test	Pre-test	Paired samples t-test
	M (SD)	M (SD)	*t*
Overall	16.18 (2.23)	15.27 (1.90)	1.61
Elegant words	2.73 (.47)	2.73 (.79)	.00
Clear paragraph	2.91 (.83)	3.09 (.70)	-.61
Coherence	3.18 (.87)	2.82 (.87)	1.00
Title consistence	4.55 (.69)	4.00 (.00)	2.63*
New & original ideas	2.82 (.75)	2.64 (.81)	.80

* $p < .05$

3.2.2 Low-ability students

Regarding how low-ability students react to the game-based peer response, significant differences were found between their post-test scores and pre-test scores, not only in the aspect of overall quality but also in the aspects of elegant words, clear paragraph, coherence, title consistence, and new & original ideas (Table 4). In other words, the game-based peer response is beneficial for the low-ability students in all aspects of writing quality. These findings suggested that the game-based peer response could significantly help low-ability students improve their writing quality.

Table 4: Writing quality of the pre-test and the post-test (low-ability students).

	Post-test	Pre-test	Paired samples t-test
	M (SD)	M (SD)	*t*
Overall	16.80 (2.49)	11.50 (.85)	5.62***
Elegant words	2.70 (.48)	1.8 (.42)	5.01**
Clear paragraph	3.40 (.84)	2.20 (.63)	3.67**
Coherence	3.10 (.74)	2.10 (.32)	3.35**
Title consistence	4.60 (.52)	3.90 (.57)	3.28*
New & original ideas	3.00 (.67)	1.50 (.53)	4.39**

* $p < .05$, ** $p < .01$, *** $p < .001$

3.3 Discussion

The results of independent samples t-tests presented in the previous subsections suggested that the game-based peer response could narrow the gap between the high-ability students and the low-ability students. More specifically, the change of this gap was from a significant level to a non-significant level. Furthermore, the aforementioned results of paired samples t-tests revealed that the effect of the game-based peer response on the improvement of high-ability students' writing quality was limited and did not reach to a significant level. Conversely, the game-based peer response was able to significantly improve the writing quality of low-ability students, regardless the overall quality or each individual aspect of writing quality. In brief, the low-ability students, but not the high-ability students, benefited from the game-based peer response. Due to such benefits, the former could demonstrate a similar level of writing quality as the latter at the post-test finally.

The fact that the high-ability students and the low-ability students reacted differently to the game-based peer response may be caused by the levels of ability that students possess. In general, high-ability students can obtain the sense of achievement from their works so they enjoy undertaking peer response from intrinsic motivation. Therefore, the impacts of the game context on the enhancement of high-ability students' motivation in peer response were limited. That is why there are no significant differences between the high-ability students' pre-test scores and post-test scores in most aspects of writing quality.

On the other hand, the low-ability students with the game-based peer response showed significantly better performance in all the aspects of writing quality, including elegant words, clear paragraph, coherence, title consistence, and new & original ideas. It may not be easy for the low-ability students to get the sense of achievement from their works so there is a need to drive them by stimulating their extrinsic motivation. This may be the reason why the low-ability students had significant improvement in writing quality after interacting with the game-based peer response. In other words, the game-based peer response played as a mechanism that can stimulate their extrinsic motivation.

More specifically, such a playful mechanism used various game elements to motivate students. For example, the game-based peer response used points to help students get rewards when they accomplished peer response activities. The other game element is the leaderboards, which show how useful students' comments to their classmates are. By doing so, students' efforts in providing helpful commentary and criticism could be well recognized. In other words, these game elements could not only enhance low-ability students' motivation, but also let them have a stronger sense of achievement. Accordingly, the low-ability students demonstrated better progress with the gamed peer response in all the aspects of writing quality.

4. Conclusions

We proposed a game-based peer response to enhance student writing and investigated how students react to the game-based peer response. In addition, the levels of ability that students possess were also considered in this investigation. Therefore, two research questions were examined in this study. Regarding the first research question, i.e., *what are the effects of the game-based peer response on students' writing quality,* the results revealed that the game-based peer response could help students enhance their writing quality. Moreover, such a game-based peer response could reduce the gap between the high-ability students and the low-ability students from a significant level to a non-significant level.

Regarding the second research question, i.e., *how high-ability students and low-ability students react differently to the game-based peer response*, the results suggested that the game-based peer response was significantly beneficial to the low-ability students, but not to the high-ability students, in terms of writing quality. These aforementioned results are interesting but this study is conducted with a small-scale sample. Therefore, future work needs to be undertaken with a large-scale sample to provide more evidence.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (MOST 101-2511-S-008 -016 -MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

Ang, C., Zaphiris, P., & Mahmood, S. (2007). A model of cognitive loads in massively multiplayer online role-playing games. *Interacting with Computers, 19,* 167–179.

Boscolo, P., & Ascorti, K. (2004). Effects of collaborative revision on children's ability to write understandable narrative texts. In G. Rijlaarsdam, L. Allal, L. Chanqouy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes. Studies in writing, Vol. 13* (pp. 157-170). Boston, MA: Kluwer.

Cheng, W., Lam, S., & Chan, C. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project- based learning. *British Journal of Educational Psychology, 78*(2), 205-221.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education, 48*(3), 409-426.

Dickey, M. D. (2007). Game design and learning: A conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development, 55*(3), 253-273.

DiPardo, A., & Freedman, S. W. (1988). Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research, 58*(2), 119-149.

Elbow, P. (1973). *Writing without teachers*. London, England: Oxford University Press.

Fried, C. B. (2008). In-class laptop use and its effects on student learning. *Computers & Education, 50*(3), 906-914.

Hansen, J. G., & Liu, J. (2005). Guiding principles for effective peer response. *ELT Journal, 59*(1), 31-38.

Howell, D.C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Learning, Wadsworth Publishing.

Huizenga, J., Admiraal, W., Akkerman, S., & ten Dam, G. (2009). Mobile game-based learning in secondary education: engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning, 25*(4), 332-344.

Leki, I. (1990). Coaching from the margins: Issues in written response. *Second language writing: Research insights for the classroom*, 57-68.

Liu, T.-Y., & Chu, Y.-L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*(2), 630-643.

Lockhart, C., & Ng, P. (1993). How useful is peer response. *Perspectives, 5*(1), 17-29.

Marsh, T. (2011). Serious games continuum: Between games for purpose and experiential environments for purpose. *Entertainment Computing, 2*(2), 61-68.

Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly, 28*(4), 745-769.

Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118-141.

Moreno-Ger, P., Martinez-Ortiz, I., Sierra, J. L., & Fernandez-Manjon, B. (2008). A Content-centric development process model. *Computer, 41*(3), 24-30.

Papastergiou, M. (2009). Digital game-based learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1-12.

Pivec, M. (2007). Editorial: play and learn: potentials of game-based learning. *British Journal of Educational Technology, 38*(3), 387–393.

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal, 59*(1), 23-30.

Sims, D. (2001). *Improving elementary school students' writing using reading and writing integration strategies.* Illinois. (ERIC Document Reproduction Service No. ED454502).

Stoddard, B., & MacArthur, C. A. (1993). A peer editor strategy: Guiding learning-disabled students in response and revision. *Research in the Teaching of English, 27*(1), 76-103.

Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing, 9*(2), 147-170.

Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition, 21*(2), 217-235.

Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., İnal, Y., & Kızılkaya, G. (2009). The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & Education, 52*(1), 68-77.

Yang, J. C., Ko, H. W., & Chung, I. L. (2005). Web-based interactive writing environment: Development and evaluation. *Educational Technology & Society, 8*(2), 214-229.

The Effects of Mini-Games on Students' Confidence and Performance in Mental Calculation

Oskar KU[a*], Denise H. WU[b], Andrew C.-C. LAO[c], Jen-Hang WANG[d] & Tak-Wai CHAN[c]
[a]*Research Center for Science and Technology for Learning, National Central University, Taiwan*
[b]*Institute of Cognitive Neuroscience, National Central University, Taiwan*
[c]*Graduate Institute of Network Learning Technology, National Central University, Taiwan*
[d]*Department of Computer Science & Information Engineering, National Central University, Taiwan*
*oskar@cl.ncu.edu.tw

Abstract: Low confidence toward mathematics is one of the critical issues that diminish students' will to learn mathematics. Several studies indicated that game-based learning (GBL) might be a potential approach to address this issue. In addition, mental calculation, a fundamental mathematical skill, was considered to be a potential skill that may build students' confidence toward mathematics. Thus, this study attempted to promote students' mathematical confidence by incorporating mini-games, a sub-genre of games, into mental calculation learning. A preliminary study was conducted to investigate how students reacted to learn mental calculation in a mini-game environment. The results of the preliminary study suggested that the mini-game-based approach was popular to students. However, it was observed that students' with different levels of academic ability might benefit diversely from the mini-game environment. Thus, an experiment was conducted to investigate how students with different levels of academic ability react to learn mental calculation with mini-games, in terms of two important aspects of learning, i.e., performance and confidence. A control group participated in a paper-based learning approach to provide a baseline of comparison. The results indicated that The mini-games helped both high- and low-ability students gain significant improvement on their mathematical confidence. In addition, the low-ability students learned with mini-games gained more improvement on mental calculation than their paper-based peers did.

Keywords: mini-games, confidence, math learning, mental calculation, learning performance

1. Introduction

Mathematics is an important and fundamental skill taught in schools. However, it is also a difficult subject in students' mind (Stodolsky, Salk, & Glaessner, 1991). Such a negative perception may result from students' low confidence toward mathematics, which, in turn, may diminish students' will to learn mathematics (Brown, Brown, & Bibby, 2008). The importance of self-confidence can be seen in many aspects (Linnenbrink & Pintrich, 2003; Maclellan, 2014). For example, a person's self-confidence toward a subject can predict not only how much effort he/she will pay to learn the subject but also his/her expectation of learning outcomes (Schunk, 1990). In addition, high self-confidence may lead students to engage in a learning task actively (Gushue, Scanlan, Pantzer, & Clarke, 2006) and attain better learning outcome (Kleitman, Stankov, Allwood, Young, & Mak, 2013). Thus, helping students to build their confidence toward mathematics is an urgent issue.

Digital games, as an alternative learning approach, may be a possible solution to the problem of students' low confidence toward mathematics. There have been a number of studies reporting the positive effects of GBL (Pivec, 2007; Chang, Wu, Weng, & Sung, 2012), such as enhancing students' learning motivation (Klawe, 1998; Nussbaum, 2007) and improving students' learning performance (Girard, Ecalle, & Magnan, 2012). Importantly, several studies indicated that GBL might enhance students' confidence (Cunningham, 1994; Radford, 2000). On the other hand, mental calculation, a relatively basic mathematical skill, may also help students shape their confidence toward learning mathematics (Rubenstein, 2001).

To this end, this study aims at investigating whether embedding learning content in digital games can enhance students' confidence and performance in mental calculation, especially for low confident students. Among various genres of digital games, mini-games are chosen as the learning environments in this research due to their simplicity. Regarding learning material, mental calculation was chosen as the learning content for its importance to mathematics. In addition, how students with different levels of ability react to the mini-games-based learning approach was also examined, because the issue of individual difference is more and more important to the design of digital learning. Thus, the research questions of this study are: (1) Can embedding mental calculation learning in mini-games enhances students' confidence and learning performance toward mental calculation and their future math learning? (2) Do students with different levels of academic ability react similarly to mini-game-based learning?

2. Related Work

2.1 Mathematical Confidence

Confidence influences students' performance (Al-Hebaish, 2012) and effort (Bandura, 1982; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). For example, confidence was also found to affects one's math performance (Stankov, Lee, Luo, & Hogan, 2012). Besides, it also affects students' will to enroll mathematics courses (Metie, Frank, & Croft, 2007).

On the other hand, the recent reports of the Trends in International Mathematics and Science Study (TIMSS) indicated that there were a large proportion of Asian students possessing low confidence toward learning mathematics (Mullis, Martin, Gonzalez, & Chrostowski, 2004; Mullis, Martin, & Foy, 2008). Such a phenomenon was getting more and more serious with the students' age. This is a critical issue that needs to be noticed because low confidence may make students feel mathematics difficult (Brown, Brown, & Bibby, 2008), which, in turn, may make student avoid facing mathematics. Thus, there is a need to help students build up their mathematical confidence.

2.2 Game-based Learning

Past studies on GBL revealed many positive effects brought by applying digital games to learning environments. For example, GBL enhanced students' learning motivation (Dickey, 2007) and improved students' engagement in learning (Huizenga, Admiraal, Akkerman, & ten Dam, 2009). In addition, digital games were found to have the potential to enhance students' confidence (Cunningham, 1994). For example, Straker et al. (2011) investigated whether playing virtual reality (VR) electronic games helped children with developmental coordination disorder (DCD) gain motor confidence. Their results indicated that playing VR electronic games enhanced DCD children's confidence in performing motor skills. Such research demonstrated that digital games might have the potential to help students build their confidence toward the course that they are learning. In this vein, digital games may be a possible solution to address the issue of students' low confidence toward learning mathematics. Therefore, embedding math learning into digital games may be a possible solution to enhance students' self-confidence toward learning mathematics.

2.2.1 Mini-games

Among various game genres, a mini-game (causal game) is a kind of relatively simple games designed for players who do not want to spend much time and effort on playing games. The main purpose of mini-games is providing people a period of relaxation time between two formal tasks. In contrast with large-scale games, mini-games are simpler and easier to be operate. Even so, they still possess the characteristics of a game. In most computer games, players have to devote significant attention on playing games. However, in mini-games, players can keep more concentration on learning tasks rather than playing games; games are just assistants that engage students in learning tasks. On the other hand, because mental calculation requires a high level of concentration for students to produce an answer, there is a need to reduce students' cognitive load. To this end, mini-game is adopted as the design of the games with exchangeable learning materials for this study. More specifically, mini-games were adopted majorly due to their simplicity, which might cause less cognitive load for learners than complex games did. Additionally, mini-games are content-independent, which makes it possible for the mini-games to be associated with different learning materials.

3. Research Design

Two studies were conducted to investigate the effects of mini-games on students' confidence and performance in mental calculation. In addition, a framework is presented to illustrate the relationship between the two studies.

3.1 Preliminary Study

Since not all GBL produced positive results in past studies, there is a need to examine whether such a pedagogy is beneficial for building students' confidence toward mental calculation and enhancing their performance. Thus, a preliminary study was conducted to examine whether learning in mini-game environments brings positive effects on students' confidence and performance in mental calculation. The purposes of the preliminary study included the following aspects:

1) Examining the effects of mini-game-based mental calculation learning. If students' attitude and performance on mental calculation were improved due to the intervention of mini-game-based learning (mGBL), a study with experimental design would be conducted to investigate how the mGBL approach builds students' confidence and meanwhile enhances their performance subsequently.
2) Collecting students' learning behavior. Students' reactions to the mini-game-based mental calculation learning will be analyzed to serve as the design principles for improving the game design in the subsequent study.

3.2 Main Study

Based on the results of the preliminary study, a main study was designed and conducted to verify whether the game-based approach significantly build students' confidence toward learning mental calculation and improve their mental calculation performance. Importantly, the main study further investigates whether the mini-game-based approach brings different learning effects on students with different levels of learning ability.

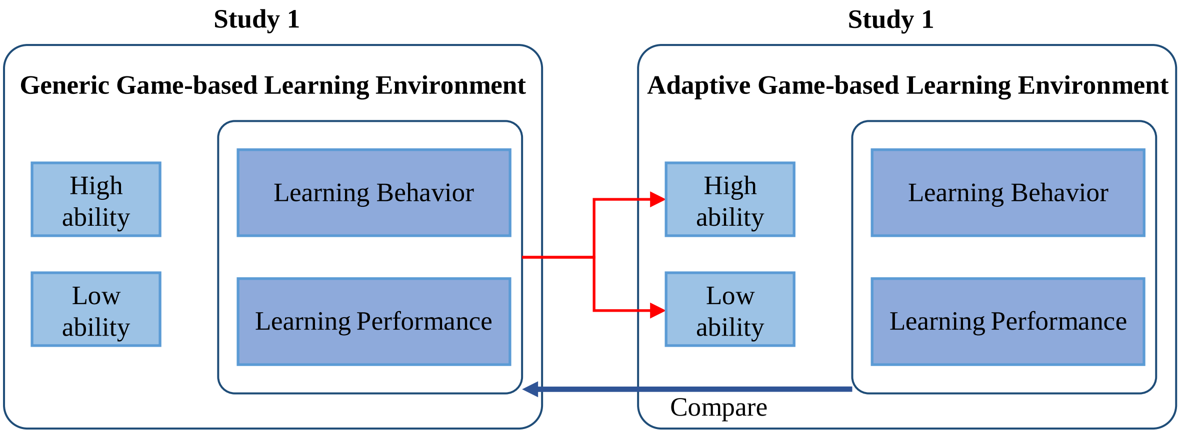


Figure 1. The relationship between Study 1 & Study 2.

3.3 Research Framework

Figure 1 illustrates the relationship between the preliminary and the main studies. The preliminary study explored the possible effects of game-based mental calculation learning while the main study further investigated how the effects were produce and whether the effects acted equally on students with different levels of ability.

4. Study One: Preliminary Study

4.1 Participants

The participants of the preliminary study were fourth grade students from an elementary school in northern Taiwan. Students in the school were normally distributed to the classes of the school based on their academic performance. Thus, a class with 28 students was randomly selected to participate in the preliminary study.

4.2 Instruments

4.2.1 Mini-games

Four mini-games (Fig. 2) were developed as the learning environments for students to learn the mental calculation skills.

Game A Game B Game C Game D

Figure 2. Screenshots of mini-games

The four mini-games were designed as two multiple choice questions games (A & B) and fill-in questions (C & D). The detail descriptions of the mini-games are shown in Table 1.

Table 1: The descriptions of the mini-games.

Game	Name	Description
A	Space Traveler	Problems are shown at the top of game screen. Students have to shoot the asteroid with correct answer. Otherwise, their spacecraft will crash.
B	Forest Protector	Students have to shoot the invader who holds the correct answer to the question displayed at the top of game screen.
C	Light City	Students have to enter the correct answer to the question shown at the bottom of game screen to prevent the bomb cart from exploding.
D	Panda Math	Students have to help the panda in the screen answer the question thrown by the enemy before the question block hit the panda.

4.2.2 Materials

In order to prevent the interference from regular math courses, this study adopted mental calculation, which was not taught in regular math courses, as students' learning materials.

There are thousands of mental calculation strategies to simplify computational problems. In addition, students can also invent their own strategies. In this study, several patterns of mental calculation skills were selected according to students' prior knowledge. The goal of this study is helping students gain more "number sense", so that they might get more sense of analyzing numbers and consequently formulate an algorithm to simplify the problem they met.

4.3 Procedures

In order to investigate students' reactions to learning mental calculation within the mini-games, a set of instruction was delivered to the participants twice a week. The set of instruction lasted for nine weeks. The Procedures of each session are depicted in Figure 3. In the beginning of each session, the teacher conducted a mini lecture to introduce a mental calculation skill for five minutes. Then, the teacher used five minutes to ensure whether students understood the content of the mini lecture. Finally, students practice in mini-games for 20 minutes.

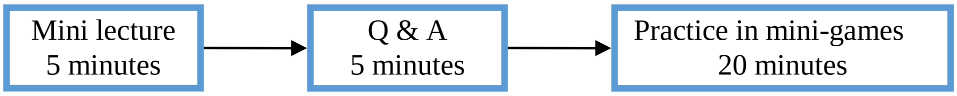


Figure 3. The procedures of a session

4.4 Results and Discussion

According to students' math scores of their school exam, students were divided into high- and low-ability groups (Table 2) to investigate whether they react differently to the intervention of mini-game-based learning. In order to investigate the impact of the mini-games in this study, the accuracy of students with different math abilities and the response time (RT) of students with different math abilities were analyzed.

Table 2: The numbers of the participants in the low and high ability groups.

High ability	Low ability	Total
13	15	28

4.4.1 Accuracy

Figure 4. The change of accuracy

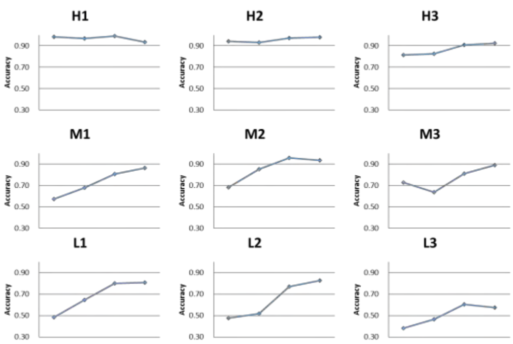


Figure 5. The average accuracy of students

The result of the Mann-Whitney test revealed a significant difference of the gain score of accuracy between the high- and the low-ability students ($U = 52.000$, $p = .036$). Such a result may imply that the low-ability students gained more improvement than high-ability students in terms accuracy (Figure 4).

Furthermore, according to students' mathematical achievements, the learning profile of the first three, the middle three, and the last three students were chosen to be analyzed. Figure 5 illustrates the average accuracy of the nine students. High math achievement students gained little improvement in accuracy; they performed very well from the beginning of the learning sessions. The middle and left-behind group students gained quite steady improvement on the accuracy of answering questions.

4.4.2 Response Time

Students' change of response time is shown in Figure 6. The result of the Mann-Whitney test indicated that no significant difference of the gain score of response time was found between the high- and the low-ability students ($U = 76.000$, $p = .322$).

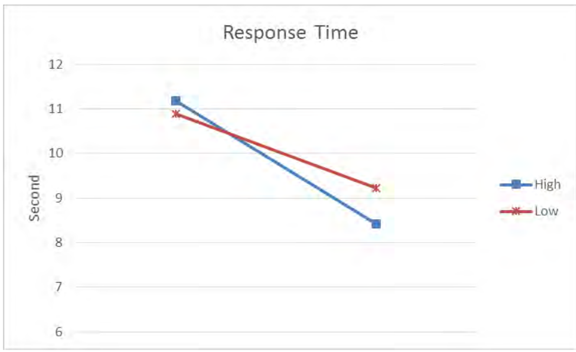


Figure 6. The change of response time

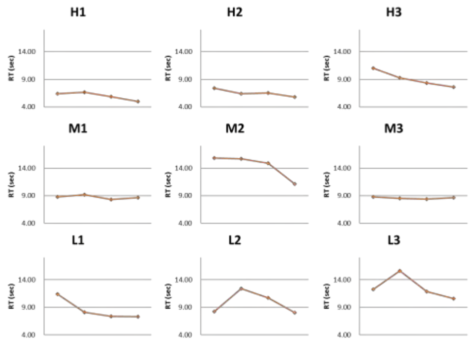


Figure 7. The average response time of students

However, when looking into the response time of the first three, the middle three, and the last three students, we can find that students showed quite different patterns of the change of response time. Figure 7 depicts students' RT. In the high achievement group, all the three students gained improvement in RT steadily. In the middle and the left-behind groups, the improvement seemed unstable. The differences among all sessions were large. Sometimes they may performed very well in a session but very poor in the next time. Maybe it was due to the fact that the learned skills in left-behind student were not consolidated enough; thus, those student performed quite unstable in their response time on answering questions.

Although several positive effects were obtained in this study, an important issue emerged with the conduction of this preliminary study should be further investigated: students with various abilities demonstrated different levels of improvement in mental calculation skills. A further study needs to be done to investigate whether such a result was caused by GBL.

5. Study Two: Main Study

The main study was conduct to investigate the ability issue raised in the preliminary study. More specifically, since individual difference is a key factor when designing individual learning environments, this study further investigates how students with different levels of academic ability react to mGBL on mental calculation in terms of two important aspects of learning, i.e., performance and confidence.

5.1 Experimental Design

A quasi-experiment was conducted to examine the effects of mGBL. Participants were fourth–grade elementary school students from Northern Taiwan. The participants were 59 students (N=59), aged 10-11 years old. Two classes of students were randomly selected and then assigned as an Experimental Group (EG, 14 males and 17 females) and a Control Group (CG, 12 males and 16 females). The experimental group received computer game-based learning, while the control group received paper-based learning.

In addition, students in both groups were further divided into two subgroups—high achievement and low achievement—to investigate the effects of mGBL on high and low achievement students.

5.2 Instruments

The learning materials used in the main study were the same as those in the preliminary study. Students in the EG learned in mini-game environments while CG students learned with paper handouts and worksheets.

5.2.1 Mini-games (EG)

Two mini-games were adopted as the learning environments for EG students to interact with learning materials. The two mini-games were implemented as a board game and a sports game (Fig. 4). The rules of both games are simple so that players can get started with the games without difficulty. Students learned and practiced in the mini-game environments. Students can replay the electronic learning materials at any time.

Figure 4. The board game and the sports game

5.2.2 Paper Handouts and Worksheets (CG)

Every student in the CG received a copy of learning materials printed on A4 paper. After receiving a short instruction from their teacher, students were ask to practice the newly learned mental calculation strategies with their worksheets. They could review the learning material whenever they have problems during practice. They could also ask their teacher to explain the mental calculation strategy individually. The teacher can pause students' practice and re-explain the strategy to the whole class if he found too many students encountered the same problem.

5.3 Procedures

The procedure of this study is presented in Figure 5. The pretest was administrated a week before the experimental intervention. The posttest was administrated a week after the last session of the experimental intervention. During the experiment, nine learning sessions were conducted for five consecutive weeks; each session lasted for 25 minutes. In each session, the teachers in both groups taught students the key features of a new strategy or reviewed previous strategies for five minutes. Students then preceded individual learning for 20 minutes.

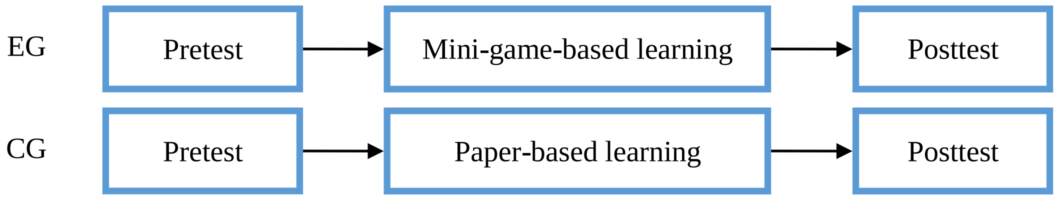


Figure 5. The procedures of a session

5.4 Results and Discussion

In order to examine whether students with different levels of ability benefited equally from the mini-game-based approach, participants were divided into four sub-groups in terms of their academic abilities (Table 3).

Table 3: The numbers of students grouped by their math ability.

	EG	CG
High ability	14	12
Low ability	12	13

5.4.1 Confidence

As shown in Table 4, low-ability CG students demonstrated a negative trend of confidence change, while the other three sub-groups demonstrated a positive change of confidence. In other words, low-ability CG students' confidence dropped after experiment.

Table 4: The mean scores of confidence toward mathematics grouped by ability.

		Pretest		Posttest	
	Ability	**Mean**	**SD**	**Mean**	**SD**
CG	Low	3.19	0.57	2.75	0.76
	High	3.55	0.67	3.69	0.88
EG	Low	3.14	0.48	3.45	0.56
	High	3.57	0.54	3.86	0.81

Regarding the mini-game-based (EG) condition, the Wilcoxon's matched-pairs signed-ranks tests revealed that both high- ($Z = 2.165$, $p = .030$) and low-ability ($Z = 2.156$, $p = .031$) students gained significant improvement on confidence. In addition the Mann-Whitney test indicated no significant difference of the gain score of confidence between the two sub-groups ($U = 81.000$, $p = .876$). Conversely, both high- and low-ability CG groups did not demonstrate significant improvement on their confidence.

For low-ability students, such a result may be caused by the positive feedback (scores, reward, and correct signs) provided by mini-games. Frequently receiving a message about completing a learning task successfully may encourage low-ability students to build their confidence toward similar tasks in the future (Pajares, 2006).

As for the high-ability EG students, there may be another source for their improvement of confidence—challenge. High-ability students may tend to expect challenge (Jones & Spooner, 2006). When completing challenging tasks, students may obtain a sense of achievement (Dickey, 2007), which, in turn, may enhance their confidence (Hammond, 2004).The mini-games adjust the game challenge dynamically; thus, high ability students can always receive adequate challenge during gameplay.

5.4.2 Mental Calculation Performance

Table 5 presents students' mental calculation performance. The Wilcoxon's matched-pairs signed-ranks tests indicated that students in each sub-group gained significant improvement of their mental calculation performance.

Table 5: The mean scores of confidence toward mathematics grouped by ability.

		Pretest		Posttest			
	Ability	**Mean**	**SD**	**Mean**	**SD**	Z	*P*
CG	Low	10.85	2.94	36.08	20.89	2.971	.003
	High	33.58	25.29	76.08	25.49	3.061	.002
EG	Low	18.58	9.92	62.75	24.63	3.061	.002
	High	24.64	22.20	75.21	22.08	3.297	.001

For the CG students, as expected, the result of the Mann-Whitney test demonstrated a significant difference of the gain score of mental calculation performance between the high-ability and the low-ability students ($U = 40.000$, $p = .039$). In other words, the high-ability students in the CG gained more improvement than their low ability peers in the same setting.

On the other hand, although the high-ability EG students did not performed as well as their high-ability CG peers, they attained a comparable level of performance in the posttest. As for low-ability students, the result of the Mann-Whitney tests demonstrated a significant difference of the gain score between the EG low-ability and the CG low-ability students ($U = 38.500$, $p = .032$) while the difference of pretest between the EG low-ability and the CG low-ability students was not found ($U = 46.500$, $p = .085$). This result might imply that the low-ability EG students gained more improvement than those in the CG.

The results demonstrated in this section indicated that the mini-game-based learning might provide greater benefit to low-ability students in terms of the learning gain. Such a result might be caused by the immediate feedback for error correction in the mini-games, which helped learner consolidate knowledge (McDaniel, Roediger, & McDermott, 2007).

6. Conclusions

Confidence plays an important role in students' learning process. However, a lack of confidence toward mathematics is many students' common problem, which prevents them from pursuing advanced mathematical knowledge. This research introduced mini-games as an approach to motivate students to learn a fundamental mathematical skill, mental calculation and investigated whether such an approach can enhance students' confidence toward mathematics.

The results from both the preliminary and the main studies indicate that students' confidence and performance in mental calculation were improved after receiving a mini-game-based mental calculation course. Furthermore, the experimental results obtained from the main study indicate that students with the mini-game-based environment gained significant improvement on confidence toward mathematics; both high and low ability students gained significant improvement on their mathematical confidence. In contrast, students learned with the paper-based setting did not demonstrate similar change on their mathematical confidence. On the other hand, students in all conditions gained significant improvement for mental calculation performance. Importantly, the EG students gained more improvement than their CG peers.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this study under contract numbers NSC 100-2511-S-008-014-MY3 and NSC 100-2511-S-008-013-MY3.

References

Al-Hebaish, S. M. (2012). The Correlation between general self-confidence and academic achievement in the oral presentation course. *Theory and Practice in Language Studies, 2*(1), 60-65.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.

Brown, M., Brown, P., & Bibby, T. (2008). 'I would rather die': Reasons given by 16-year-olds for not continuing their study of mathematics. *Research in Mathematics Education, 10*(1). 3-18.

Chang, K. E., Wu, L. J., Weng, S. E., & Sung, Y. T. (2012). Embedding game-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education, 58*(2), 775-786.

Cunningham, H. (1994). Gender and computer games. *Media Education Journal 17*, 13-15.

Dickey, M. D. (2007). Game design and learning: A conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development 55*(3), 253-273.

Girard, C., Ecalle, J., & Magnan, A. (2012). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning, 29*(3), 207-219.

Gushue, G. V., Scanlan, K. R., Pantzer, K. M., & Clarke, C. P. (2006). The relationship of career decision-making self-efficacy, vocational identity, and career exploration behavior in African American high school students. *Journal of Career Development*, *33*(1), 19-28.

Hammond, C. (2004). Impacts of lifelong learning upon emotional resilience, psychological and mental health: fieldwork evidence. *Oxford Review of Education*, *30*(4), 551-568.

Huizenga, J., Akkerman, S., Admiraal, W., & ten Dam, G. (2009). Mobile game-based learning in secondary education: Engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning*, *25*(4), 332-344.

Jones, G., & Spooner, K. (2006). Coaching high achievers. *Consulting Psychology Journal: Practice and Research*, *58*(1), 40-50.

Klawe, M. (1998). When does the use of computer games and other interactive multimedia software help students learn mathematics? *Proceedings of NCTM Standards 2000 Technology Conference*. Arlington, VA.

Kleitman, S., Stankov, L., Allwood, C. M., Young, S., & Mak, K. K. L. (2013). Metacognitive self-confidence in school-aged children. In M. M. C. Mok (Ed.), *Self-directed learning oriented assessment in the Asia-Pacific* (pp. 139-153). New York, NY: Springer.

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly: Overcoming Learning Difficulties 19*(2), 119-138.

Maclellan, E. (2014). How might teachers enable learner self-confidence? A review study. *Educational Review, 66*(1), 59-74.

McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14*(2), 200-206.

Metie, N., Frank, H. L., & Croft, P. (2007). Can't do maths—Understanding students' math anxiety. *Teaching Mathematics and its Applications, 26*(2), 79-81.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades.* Boston, MA: International Association for the Evaluation of Education Achievement, Boston College.

Mullis, I. V. S., Martin, M. O., and Foy, P. (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades.* Boston, MA: International Association for the Evaluation of Education Achievement, Boston College.

Nussbaum, M. (2007). Games, Learning, Collaboration and Cognitive Divide. OECD. Retrieved May, 20, 2010 from http://www.oecd.org/dataoecd/43/39/39414787.pdf

Pajares, F. (2006). Self-efficacy during childhood and adolescence. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 339-367). Greenwich, CT: Information Age Publishing.

Pivec, M. (2007). Editorial: Play and learn: potentials of game-based learning. *British Journal of Educational Technology, 38*(3), 387-393.

Radford, A. (2000). Games and learning about form in architecture. *Automation in Construction*, *9*(4), 379–385.

Rubenstein, R. N. (2001). Mental Mathematics beyond the Middle School: Why? What? How?. *The Mathematics Teacher*, *94*(6), 442-446.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*, 71-86.

Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences*, *22*(6), 747-758.

Stodolsky, S., Salk, S., & Glaessner, B. (1991). Student Views About Learning Math and Social Studies. *American Educational Research Journal, 28*(1), 89-116.

Straker, L. M., Campbell, A. C., Jensen, L. M., Metcalf, D. R., Smith, A. J., Abbott, R. A., ... & Piek, J. P. (2011). Rationale, design and methods for a randomised and controlled trial of the impact of virtual reality games on motor competence, physical activity, and mental health in children with developmental coordination disorder. *BMC Public Health*, *11*, 654-665.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202-248.

Chan, C. H. & Chen, S. Y. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Interface Design of Electronic Journals via Mobile Devices: A Cognitive styles Perspective

Chu-Han Chan [a] **& Sherry Y. Chen** [b*]
[a] *Graduate Institute of Network Learning Technology National Central University, Jhongli Taiwan*
[b] *Graduate Institute of Network Learning Technology National Central University, Jhongli Taiwan*
* sherry@cl.ncu.edu.tw

Abstract: With the advancement of information technology, combining with electronic journals and mobile devices would produce ubiquitous electronic journals. However, there is a need to consider the usability evaluation because usability is a strong predictor of design issues. To satisfy individual needs, the effects of cognitive styles on usability inspection are investigated in this study. To this end, this study aimed to examine how different cognitive style groups perceive the interface design of an electronic journal. More specifically, Nielsen's ten heuristics (Hs) were applied to investigate user' perceptions. The results show that H8 was considered the most important heuristic by all users. The results also demonstrate that Holists who perceive excessive advertising may strongly need previous/next buttons while Serialists who feel this electronic journal provides too many advertising may consider that too much information is presented in the home page. The findings can be applied to support the development of individualized mobile electronic journals.

Keywords: Electronic journals, Mobile devices, Cognitive styles, usability, Nielsen's heuristics

1. Introduction

Digital learning refers to utilize digital technologies to support student learning (Chan, et. al., 2006). Among a variety of digital technologies which can be applied to implement learning materials, mobile devices particularly offer many advantages, e.g., convenience, flexibility and ubiquitous information access (Jacob & Issac, 2008). Among these advantages, the portability is a major advantage that leads to the other two. Regarding flexibility, portability can facilitate users to access information anytime (Liu and Carlsson, 2010). Regarding ubiquity, portability removes geographic boundaries so users can locate information at any locations (Looney et al., 2004). Due to these advantages, there are on-going interests to use mobile devices to support teaching and learning recently (Morris, 2010; Petrova and Li, 2009). For instance, Wurst, Smarkola and Gaffney (2008) compared ubiquitous mobile learning with a traditional lecture-based course in higher education. The results from their study suggested users with mobile learning showed significantly more satisfaction than those in traditional classrooms. More recently, Cavus and Uzunboylu (2009) used the mobile devices to develop a mobile learning system and they found both users' attitudes toward the mobile devices and their creativity were improved significantly at the end. In summary, mobile learning does indeed become a mainstream method of education in 21st Century (Peters, 2007).

Further to mobile devices, electronic journals are another useful digital technology widely used in educational settings because they can facilitate to disseminate scientific information (Ollé and Borrego, 2010). By doing so, students can effectively acquire new information to enhance their understandings. In addition to disseminating scientific information, the electronic journals also provide other benefits, including the speed of access and the ability to download, print, and send articles (Tyagi, 2011). Due to the widespread use of electronic journals, research into this issue has mushroomed. In an early period, Bar-Ilan and Fink (2005) conducted a study to examine the use of printed and electronic journals in a science library. The results showed more than 80% of the respondents frequently used and preferred an electronic format. Later on, Prabha (2007) tracked journal subscription and format data for 515 journals in the Association of Research Libraries (ARL) university member libraries. The findings showed journals subscribed in print only decreased to one-third of the journal collections while, concurrently, access to electronic journals increased to one-third of the collections.

The aforementioned studies demonstrated electronic journals played an important role in scholarly communication. Such importance increases the use of electronic journals in various countries. For instance, Kurata et al., (2007) examined the position of electronic journals in scholarly communication based on Japanese researchers' information behavior. The results showed Japanese researchers used electronic journals for information access as a matter of course. Recently, Bravo and Díez (2011) examined the models of consumption of the academic communities of five Spanish universities. Their study revealed the overall totals for downloads at the universities showed constant growth from 2002 onward. In other words, there was an upward trend in the consumption of scholarly information in electronic formats in the Spanish academic communities.

The aforesaid results demonstrated electronic journals are popular academic tools. In other words, there are an increasing number of users to access electronic journals. On the other hand, great diversities exist among such users, who may have heterogeneous backgrounds, in terms of their knowledge, skills and needs (Chen and Macredie, 2010). Thus, it is necessary to examine relationships between individual differences and the use of electronic journals. Among various individual differences, previous studies mainly focused on examining how users' subject background affected their information seeking behavior (Talja and Maula, 2003). In addition to subject background, other human factors are also essential, e.g., cognitive styles, which refer to a person's information processing habits, capturing an individual's preferred mode of perceiving, thinking, remembering, and problem solving (Messick, 1976). Previous research found cognitive styles are key determinants to affect users' information seeking (Clewley et al., 2010). Thus, it is necessary to examine how different cognitive style groups react to the use of electronic journals.

Among various dimensions of cognitive styles, Pask's Holism/Serialism has been received attention recently. Jonassen and Grabowski (2012) describe Holists as preferring to process information in a 'whole-to-part' sequence. In contrast, Serialists are described as preferring a 'part-to-whole' processing of information. Holists and Serialists have different characteristics. Due to such differences, recent works examined how Holists and Serialists behave differently. For instance, Clewley et al., (2011) found Serialists and Holists have different preferences for their navigational styles. The former prefer to follow a linear pattern by having a suggested route or looking at the subject content step-by step with back/forward buttons. Conversely, the latter tend to take a non-linear pattern by 'jumping' between different levels of subject contents with hypertext links. Furthermore, Chen and Chang (2014) investigated how member grouping affects users' reactions to mobile collaborative learning from a cognitive style perspective. The results suggest there is a need to provide Serialists with additional help when they use mobile collaborative learning.

In addition to the effect of the cognitive styles, the interface design of the electronic journals is also important because user interface may be thought of as a 'window' through which users interact with electronic journals so the design of user interface may affect how users access electronic journals. In other words, the user interface formulates the working environment of electronic journals so it is critical that the working environment is friendly enough to accommodate users' different preferences. As such, the usability evaluation of electronic journals becomes paramount because it can provide concrete prescriptions for developing electronic journals that are able to align to diverse users' needs. A number of methods can be used to evaluate usability. Among them, Nielsen's heuristic approach is most commonly used because it can be used effectively by novices and experts alike and can be performed at any stages of the development lifecycle (Nielsen, 1994a). Nielsen's Heuristics were first formally described in presentations in the Human–Computer Interaction conference through papers published by Nielson and Molich (1990). Since then, they have refined the heuristics based on a factor analysis of 249 usability problems to derive a revised set of heuristics with maximum explanatory power. Table 1 presents the detail of the revised set of 10 heuristics (H).

These ten heuristics are concise and simple to learn so they are widely applied to evaluate the user interface of a variety of applications. Petrie and Power (2012) assessed the usability of six complex, highly interactive websites based on Nielsen's heuristics. The results of their study showed there were 935 usability problems found in the evaluation. Recently, Hsieh, Su, Chen and Chen (in press) also used Nielsen's ten heuristics to assess the usability of a robot-based learning companion. Based on the results of the assessment, they developed three versions of robot-based learning companion. Due to such popularity, the study presented in this paper also assesses the usability of a game-based learning system with Nielsen's ten heuristics.

Table 1: Nielsen's ten heuristics (1994b).

Heuristics	Explanations
H1:Visibility of system status	The system should always keep user informed about what is going on by providing appropriate feedback within reasonable time
H2:Match between system and the real world	The system should speak the user's language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order
H3:User control and freedom	Users should be free to develop their own strategies, select and sequence tasks, and undo and redo activities that they have done, rather than having the system do these for them
H4:Consistency and standards	Users should not have to wonder whether different words, situations, or actions mean the same thing and the system should follow platform conventions.
H5:Error prevention	Even better than good error messages is a careful design, which prevents a problem from occurring in the first place.
H6:Recognition rather than recall	Make objects, actions, and options visible. The users should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
H7:Flexibility and efficiency of use	Allow users to tailor frequent actions. Provide alternative means of access and operation for users who differ from the ''average'' user (e.g., physical or cognitive ability, culture, language, etc.)
H8:Aesthetic and minimalist design	Dialogues should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
H9:Help users recognise, diagnose and recover from errors	Error messages should precisely indicate the problem and constructively suggest a solution. They should be expressed in plain language.
H10:Help and documentation	Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

The aforementioned studies demonstrate the usefulness of Nielsen's heuristic evaluation. However, paucity of studies uses Nielsen's heuristics to assess the user interface of electronic journals, i.e., the ScienceDirect. In particular, there is a lack of studies to investigate Holists and Serialists' reactions to electronic journals in the context of mobile devices. To this end, we address this issue. In brief, the aim of this study is to examine how different cognitive style groups perceive the interface design of an electronic journal.

2. **Methodology**

2.1 Participants

As indicated by Nicholas et al. (2009), the majority users of digital resources were students. Thus, the participants (N=23) were recruited from master students from the Department of Computer Science and Information Engineering at National Central University in Taiwan. In other words, the participants had a similar subject background so that the effects of prior knowledge could be minimized. In addition, a request was issued to students in lectures, and further by email, making clear the nature of the study and their participation. All participants had the basic computer and Internet skills necessary to use the electronic journals.

2.2 ScienceDirect

Among various electronic journals, this study adopted the ScienceDirect (Figure 1) to reach the aim described in Section 1. This is because the ScienceDirect covers various topics, such as life sciences, chemistry, and physics. Furthermore, the ScienceDirect also provides multiple search mechanisms: (1) Basic Search, (2) Advanced Search and (3) Expert Search, which differ with respect to the complexity of their interface design and search mechanisms. More specifically, the Expert Search and Advanced Search were considered as an example of complex search design whereas the Basic Search was appreciated by its simplicity. Having such varieties in interface design and search mechanisms provides a wider range of choices, which can help to identify users' preferences.

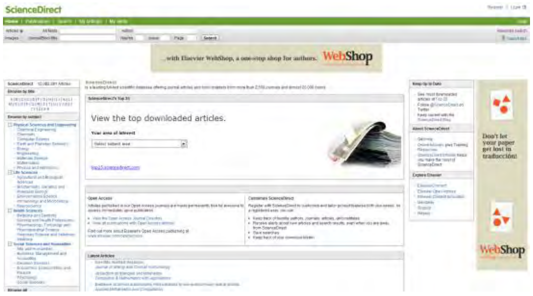


Figure 1. The homepage of the ScienceDirect.

2.3 Questionnaire

To investigate how users with different cognitive styles perceived the interface design of the ScienceDirect. A paper-based questionnaire was developed and it included two parts. In the first part, which included 10 three-point Likert-scale questions ("disagree", "general" and "agree"), users were asked to describe the degree of their satisfaction with the ScienceDirect on the basis of each heuristic. The internal consistency for the overall scale is 0.58 by Cronbach's alpha, which indicates an adequate satisfaction of the questionnaire. In the second part, which consisted of 30 questions, users were requested to check whether the interface design of the ScienceDirect met the criteria of each heuristic.

2.4 Experimental Procedures

To achieve the aim of this study, the procedure included three steps (Figure 2). Initially, all participants were required to fill out their personal information and the SPQ. According to the results of the SPQ, our participants consisted of 12 Holists and 11 Serialists. Subsequently, all of the participants were trained to learn the principles of Nielsen's heuristics so that all of the participants had the understandings of how to conduct the usability assessment. Then, they were required to interact with the ScienceDirect via tablet PCs. Finally, the participants needed to evaluate the usability of the ScienceDirect based on Neilson's ten heuristics. Such evaluation was conducted via the questionnaire described in Section.

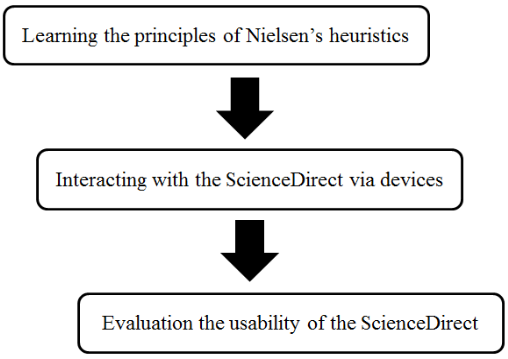


Figure 2. The Experimental Procedure.

2.5 Data Analyses

Traditional statistics were applied to conduct data analyses from both macro and micro views in Study Two. The macro view covers two aspects: (a) relationships between the satisfaction of each heuristic and (b) relationships between each criterion in all heuristics. The micro view is obtained by further examining the aforementioned relationships. Spearman's correlations, which could be used to interpret the strengths of a statistical relationship between two random variables (Stuart et. al, 1991), were applied to find the aforesaid macro view and micro view. Such analyses were undertaken by using Statistical Package for the Social Sciences (SPSS) for Windows (release 18.0). A significance level of $p<0.05$ was adopted for this study.

3. Results and Discussions

3.1 Overall

The satisfaction of H8 is negativity related to that of H1($r=-.458$, $p<.05$) and positively related to H3($r=.492$, $p<.05$) and H6 ($r=.492$, $p<.05$). The results indicated users with high satisfaction with H8 would show low satisfaction with H1 whereas they would show high satisfaction with H3 and H6. In other words, the users' satisfaction with H8 plays an important role. Thus, this study also conducted detailed analyses for questions related to H8, H1, H3 and H6. As displayed in Table 2, H8 includes three items, i.e., Q24, Q25 and Q26. Q24 is associated with Q18 belonged to H6. Q25 is linked with Q3 and Q8, which are belonged to H1 and H3, respectively. Q26 is connected with Q18 belonged to H6. These findings suggest Q25 is an important issue, which is related to Q3 and Q8. More specifically, too many advertisements may let users feel that it is difficult to identify where the Expert Search is and that there is a need to provide previous/next buttons. The other important issue is Q18, which is related to Q24 and Q26. In other words, presenting too much information in the home page may also make users feel that this electronic journal provides too many functions and too much information. This finding suggests displaying too much information in the home page may cause users' cognitive overload so they cannot appreciate the value of information and function provided by the electronic journal. In brief, there is a need to pay enough attention to Q18 and Q25, which are essential for the interface design of electronic journals.

Table 2: The variables of Nielsen's Heuristics (The whole sample).

		H8		
		Excessive functions (Q24)	Excessive advertising(Q25)	Overall Excessive information(Q26)
H1	Highlighted Keywords(Q1)	.233	.042	.215
	Lack of detailed instruction (Q2)	-.094	-.094	-.210
	Hard to find the location of the Expert Search (Q3)	-.342	**-.533****	-.151
H3	Lack of undo/redo functions (Q7)	.279	.058	.128
	Lack of previous/next buttons (Q8)	-.086	**.509***	-.066
	Provisions of multiple search.(Q9)	-.350	-.163	-.302
H6	Too many subject categories (Q17)	.387	.147	.250
	Excessive information in the Home page (Q18)	**.707****	.311	**.691****
	Clear text icons (Q19)	.042	.042	.032

Keys: * $p < .05$, ** $p < .01$

3.2 Cognitive styles

Further to the aforesaid findings for the whole sample, how each cognitive style group reacted to each Nielsen's heuristic is also analyzed. Holists and Serialists share some similarities but several differences also exist between them.

3.2.1 Similarities

The satisfaction of H8 was positively related to H3 for Holists(r=.622, p<.05). On the other hand, the satisfaction of H8 was positively related to H6 (r=.777, p<.01) and negativity related to H1 (r=-.712, p<.05) for Serialists. These results indicated Holists and Serialists who showed high satisfaction with H8 would show high satisfaction with H3 and H6, respectively but Serialists would also show low satisfaction with H1. In other words, the satisfaction with H8 plays an important role for both Holists and Serialists. Thus, this study also conducted detailed analyses for questions related to H8, H1, H3 and H6. As displayed in Table 3, H8 includes three items, i.e., Q24, Q25 and Q26. Regarding Holists, Q24 and Q25 are associated with Q9 and Q8 belonged to H3. The findings from Holists are similar to those from the whole sample. More specifically, Holists who perceived excessive advertising may strongly need to use previous/next buttons. Additionally, Holists who perceived excessive functions may not need the provision of multiple search. This may be due to the fact excessive advertising and functions increase their cognitive overload already so they do not need multiple search but they need previous/next buttons to facilitate their navigation in hyperspace.

Regarding Serialists, Q25 is related to Q18 belonged to H6 and Q26 is connected with Q1 and Q18 belonged to H1 and H6, respectively. These findings suggest Q26 and Q18 are important issues. Regarding Q26, highlighted Keywords in search results and too much information displayed in the home page may let Serialists feel overwhelmed. Regarding Q18, presenting too much information in the home page may also make Serialists feel this electronic journal provides too many advertising and information. This finding is consistent with the results from 3.1 which claim too much information displayed in the home page may cause users' cognitive overload. Such a problem may be more serious to Serialists because they only use the options that are relevant to their current tasks (Clewley et al., 2010), which, in turn, they cannot appreciate the value of rich information provided by the electronic journal. In brief, Q18 and Q26 are essential factors for designing the interface of electronic journals for Serialists.

3.2.2 Differences

Regarding Serialists, the satisfaction of H1 was negatively related to H6 (r=-.969, p<.05). Regarding Holists, the satisfaction of H8 was negatively related to H5 (r=-.32, p<.05) and the satisfaction of H1 was positively related to H7 (r=.853, p<.01). In other words, the users' satisfaction with H1 plays an important role. Thus, this study also conducted detailed analyses for relationships between questions belonged to H1 and those belonged to H6 and H7. However, no significant relationships were found for Serialists. Conversely, some significant relationships were discovered for Holists. As displayed in Table 4, H1 includes three items, among which both Q1 and Q3 are associated with Q22 belonged to H7. In other words, Q22 is an important issue. Regarding Q1, the highlighted keywords in search results may be enough for Holists so that they do not need different types of font size to enhance the visual clue. Regarding Q3, it is difficult to find the location of the Expert Search for Holists so they may need to change the font size to help them find where the Expert Search is

Table 3: Findings similar to the whole sample.

		H8		
		Excessive functions (Q24)	Excessive advertising(Q25)	Overall Excessive information(Q26)
Holists				
H3	Lack of undo/redo functions (Q7)	.529	-.316	.447
	Lack of previous/next buttons (Q8)	-.239	**.625***	-.354
	Provisions of multiple search.(Q9)	**-.657***	-.120	-.507
Serialists				
H1	Highlighted Keywords(Q1)	.542	-.039	**.671***
	Lack of detailed instruction (Q2)	-.194	.418	-.289
	Hard to find the location of the Expert Search (Q3)	-.463	-.571	-.311
H6	Too many subject categories (Q17)	.542	.386	.261
	Excessive information in the	.542	**.810****	**.671***

	Home page (Q18)			
	Clear text icons (Q19)	-.149	-.311	-.467

*Keys: * p < .05, ** p < .01*

Table 4: Findings different from the Whole sample.

		H1		
		Highlighted Keywords(Q1)	Lack of detailed instruction (Q2)	Hard to find the location of the Expert Search (Q3)
Holists				
H7	Only English version(Q20)	-.029	-.239	-.169
	Provision of three different search mechanisms(Q21)	.507	-.354	1.0
	Provision of three different types of font size (Q22)	**-.598***	.250	**.625***
Serialists				
H6	Too many subject categories (Q17)	.542	.516	-.463
	Excessive information in the Home page (Q18)	.083	-.194	-.463
	Clear text icons (Q19)	-.559	-.289	.069

Keys: * p < .05, ** p < .01

4. Conclusion

This study aims to examine how different cognitive style users response differently to the interface design of the electronic journal. The major results of our research showed most of the students thought H8 was the most important heuristic. However, there are some differences between Holists and Serialists. More specifically, Holists who perceive excessive advertising may strongly need previous/next buttons while Serialists who feel this electronic journal provides too many advertising may feel too much information presented in the home page. Such differences between Holists and Serialists reveal that cognitive styles do play an important role. Accordingly, cognitive styles should be considered for the development of individualized mobile electronic journals. However, this study has several limitations. Firstly, the sample is small so further works need to use a larger sample to verify the findings presented in this study. Additionally, there is also a need to conduct further research to examine how other human factors, such as gender differences or prior knowledge, influence learners' responses to the usability inspection of the electronic journals in the mobile context.

Acknowledgements

We would like to thank all the people who prepared and revised previous versions of this document.

References

Bar-Ilan, J. , & Fink, N. (2005). Preference for electronic format of scientific journals: A case study of the Science Library users at the Hebrew University. *Library& Information Science Research, 27*(3), 363−376.

Bravo, B. R., & Díez, M. L. A. (2011). An analysis of the use of electronic journals in a Spanish academic context: Developments and profitability. *Serials Review, 37*(3). 181-195.

Cavus, N. , & Uzunboylu, H. (2009). Improving critical thinking skills in mobile learning. *Procedia-Social and Behavioral Sciences, 1*(1), 434-438.

Chan, T.-W., Roschelle, J., Hsi, S., Kinshuk, Sharples, M., Brown, T., et al. (2006). One-to-one technology-enhanced learning: An opportunity for global research collaboration. *Research and Practice in Technology-Enhanced Learning, 1*(1), 3-29.

Chen, S. Y. , & Macredie, R. D. (2010). Web-based Interaction: A Review of Three Important Human Factors. *International Journal of Information Management, 30*(5), 379-287.

Chen, S. Y. , & Chang, L. P. (2014). The influences of cognitive styles on individual learning and collaborative learning. *Innovations in Education and Teaching International, (ahead-of-print)*, 1-14.

Clewley, N. , Chen, S. Y. , & Liu, X. (2010). Cognitive styles and search engine preferences: Field dependence/independence vs holism/serialism. J*ournal of Documentation, 66*(4), 585-603.

Clewley, N. , Chen, S. Y., & Liu, X. (2011). Mining Learning Preferences in Web-based Instruction: Holists vs. Serialists. *Educational Technology & Society, 14*(4), 266-277.

Hsieh, Y. Z., Su, M. C. , Chen, S. Y., & Chen, G. D. (2013). The development of a robot-based learning companion: a user-centered design approach. *Interactive Learning Environments, (ahead-of-print),* 1-17.
Jacob, S. M. , & Issac, B. (2008). Mobile technologies and its impact–An analysis in higher education context.
Jonassen, D. H., & Grabowski, B. L. (2012). Handbook of individual differences learning and instruction. Routledge.
Liu, Y., Li, H. , & Carlsson, C. (2010). Factors driving the adoption of m-learning: An empirical study. *Computers & Education*, *55*(3), 1211-1219.
Kurata, K., Matsubayashi, M., Mine, S., Muranushi, T., & Ueda, S. (2007). Electronic journals and their unbundled functions in scholarly communication: Views and utilization by scientific, technological and medical researchers in Japan. *Information processing & management, 43*(5), 1402-1415.
Looney, C. A. , Jessup, L. M., & Valacich, J. S. (2004). Emerging business models for mobile brokerage services. *Communications of the ACM, 47*(6), 71-77.
Messick, S. (1976). *Individuality in learning*, Jossey-Bass San, Francisco.
Morris, T. A. (2009). *Anytime/anywhere online learning: Does it remove barriers for adult learners*. Online education and adult learning: New frontiers for teaching practices. Hershey, PA: IGI Global.
Nielsen, J. & Molich, R. (1990). Heuristic evaluation of user interfaces. In Proceedings of CHI 90,249-256.
Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. Proceedings of CHI'94 conference. 152-158.
Nicholas, D. , Huntington, P., Jamali, H. R., Rowlands, I., & Fieldhouse, M. (2009). Student digital information-seeking behaviour in context. *Journal of Documentation*, *65*(1), 106-132.
Ollé, C. , & Borrego, Á . (2010). A qualitative study of the impact of electronic journals on scholarly information behaviour. *Library & Information Science Research*, *32*(3), 221-228.
Petrie, H. , & Power, C. (2012). What Do Users Really Care About? A Comparison of Usability Problems Found by Users and Experts on Highly Interactive Websites. CHI '12 Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems. 2107-2116.
Prabha, C. (2007). Shifting from print to electronic journals in ARL university libraries. *Serials Review, 33*(1), 4-13.
Stuart, A. , Kendall, M. G., & Ord, J. K. (1991), *Classical Inference and relationship*. Oxford University Press, New York.
Talja, S. , & Maula, H. (2003). Reasons for the use and non-use of electronic journals and databases: a domain analytic study in four scholarly disciplines. *Journal of documentation, 59*(6), 673-691.
Tyagi, S. (2011). Use of E-Resources by Engineering Faculties in Selected Universities of Western Uttar Pradesh, India: A Survey. *Asian Journal of Information Science and Technology, 1*(1), 56-62.
Wurst, C. , Smarkola, C., & Gaffney, M. A. (2008). Ubiquitous laptop usage in higher education: Effects on student achievement, student satisfaction, and constructivist measures in honors and traditional classrooms. *Computers & Education, 51*(4), 1766-1783.

Enhancing Metacognition through Weblog in Physics Classroom Thai Context

Jirutthitikan PIMVICHAI, Chokchai YUENYONG* and Sakanan ANANTASOOK

Science Education Program, Faculty of Education, Khon Kaen University, Thailand
* ychok@kku.ac.th

Abstract: This study examines the effects of using weblog technologies to create argumentative activities and enhance students express the metacognition on the situations' discussion. The Toulmin's diagram of Arguments Pattern (TAP) was adopted to interpret the students' metacognition in Thai rural physics classroom context. The participants were 33 Grade 10 students in the first semester of 2014 academic year, from a rural school in the Khon Kaen province of Northeast Thailand. The finding shows that rural students are very satisfactory to learn through the blog and the blog support them to post the comment for discussion. Their writing express metacognitive awareness, knowledge and control depend on each TAP's coding schema. Consequently, this study provides effective weblog argumentative strategies for enhancing student's metacognition in physics classroom Thai context.

Keywords: metacognition, argumentation, physics weblog

1. Introduction

A weblog (blog) is a web-based technology that is one of the popular web 2.0 teaching tools in this century (Namwar & Rastgoo, 2008). Blog can help to break down the classroom walls and increase students' motivation to learn about science (Barlow, 2008). It is provide both teachers and learners with the ability to extend discussions away from the traditional face-to-face classroom. Blogs provide teachers as a user-friendly online format that can be used to emphasize strategies, introduce new topics, ask questions, review concepts, evaluate for tests, have argumentation and provide enhancement opportunities (Barlow, 2008). Students learn how to share their thoughts and communicate opinions which reflecting on real-world issue through comment on the blog (Duplichan, 2009). In addition, blogs can also give the "silent student" a voice, they are not very comfortable show their ideas in classroom. The blog can offer a "safe place" for their voices to be heard in a lower pressure environment (Luehmann and Frink, 2009). Students felt blog was "fun" and "helpful" and made them interested in using these technologies (Barlow, 2008; Columbo, 2007; & Erickson, 2009). It allows students the chance to take charge of their own learning (Luehmann and Frink, 2009).

Since 2010, The Technology for Teaching and Learning Center, under the Thai Ministry of Education every years training Thai teachers to create blog content through "wordpress" free blog service (URL: http://www.wordpress.com). Anantasook (2014) research study show that the students achievement after learning by using a weblog that applying social media entitle "basic C program language" was higher than before learning at a significance level of .05 and the satisfied of students who learning academic achievement with lesson on the blog in the high level as 4.29 of 5 score and standard based as 0.86. Another study suggests that teaching with lessons on the blog could decrease the differences between individuals. Students can study base on their own or preparing to study before attends in class. They can learn a lesson at anytime, anywhere when they comfortable and they can share ideas with friends via the comments box lessons. These can decrease the problem of students who not dare ask the teacher, answer the question or discussion in the classroom (Sudprakone, 2012).

According to the section 24 of the National Education Act of B.E. 2542 (1999) (Office of the National Education Commission; ONEC, 1999), provide the important idea of learning process, educational institutions shall provide training in thinking process, management, how to face various situations and application of knowledge for obviating and solving problems. The teacher must be aware of the importance of effective strategies for teaching and learning science (physics).

Physics is a scientific subject that is important to study in order to understand phenomena that occur and an introduction before study in higher education (Rosnow & Rosenthal, 1989; Aikenhead & Ryan, 1992). Most students do not like it, they think physics is a difficult subject, difficult to grasp and understand especially on the calculation content. Therefore, it must be improvements in the instructional model for teaching and learning factors that affecting the increase academic achievement. Yuenyong, Jones and Sung-ong (2011) suggested that metacognition is essential tool that supports students' learning physics using higher-order thinking processes and situated cognition ideas are the arguments that show the influence of culture on students' cognitive processes. When students learn how to create a scientific argument that there is a reason able to integration of the thinking skills with specific knowledge related to social issues. Students are able to use better reasoning in the support of the issue manually in order to promote the issues in argument are conflicting (Lin and Mintzes, 2010). They can develop confidence in making decisions in their lives and to participate as a responsible citizen in the social responsibility and democratic (Driver, et al., 2000).

However, from the previous researcher study revealed that Thai students from both urban and rural school are often shy, afraid to argue with friends and are afraid of express their own opinions in classroom. Particularly the students in rural school, they show a few comment and they do not attempt to describe the situation, discussion or arguments by their own. It seems like they lack metacognitive awareness, knowledge and control. From above advantage of weblog to support learning, the researcher interesting to create this technology as a metacognition strategy for enhance argumentative teaching and learning in rural physics classroom under Thai context.

2. Aims of the research

This research study examines the effect of weblog upon enhancing argumentation and metacognition in Thai rural physics classroom context.

3. Methodology

This study concern with qualitative approach, the findings stem from interpretive the students responded; how the effect of weblog upon enhancing argumentation and metacognition in Thai rural physics classroom context. The Toulmin's diagram of Arguments Pattern (TAP) (Toulmin, 1958) were used to interpret the students' metacognition that from six argumentation situations which shown in the comment wall on the blog.

Participants

The sample for the case study consisted of 33 Grade 10 students in the first semester of 2014 academic year that from Dongmon Wittayakom school, a rural small size school in the Khon Kaen province of Northeast Thailand.

Weblog as the instrument

The weblog [http://jirutthitikan.wordpress.com/] which concern in this study was created by the researcher. The objective's blog setting was provided for teaching and learning of argumentation in physics classroom at the force, work and energy Unit. This blog consist of five main menus; (1) Introduction of the Unit, (2) Lesson on the Unit, (3) Argumentation situation, (4) Participants, and (5) Mind Map.

The first introduction the unit menu includes the introduction on the physics area in Thai science curriculum and the objective of learning in the force, work and energy unit. The second lesson on the unit menu that include pre-post test and 6 contents of the force, work and energy unit which student must learn in class that are: : (1) Force and Work (2) Energy, Kinetic Energy, Potential Energy (3) Law of Energy Conservation, (4) Apply Law of energy Conservation, (5) Power, and (6) Mechanical. The third argumentation situation menu, that include six argumentation situations which relevant with six contents in the previous menu. The fourth participant menu that provide the participant information in this study and the final mind map menu will show students' task after they finish learning the unit.

4. Data collection and data analysis

Five students' groups completed argumentation in the six argumentative questions with their friends. They also express evidence of their conceptual understanding "thinking of thinking" (metacognition) by post or comment discussion. The metacognition on student discussion could found through Toulmin's diagram of Arguments Pattern (TAP) (Toulmin, 1958). It consists of 6 categories that include: Data, Claim, Warrants, Qualifiers, Rebuttals and Backing that shown in Table 1. The interpretive of students' argumentation typically looks at three dimensions of metacognition (metacognitive awareness, knowledge and control).

Table 1: Coding scheme from Toulmin's diagram of Arguments Pattern (TAP)

Discourse Move	Definition
Data	Students can use facts or evidence to prove their argument.
Claim	The statement being argued. Students' principle comment or an assertion made by students' brainstorm in groups. Arguments which are a simple claim versus a counterclaim or a claim versus a claim. Argument which may have several claims and counterclaims.
Warrants	A student has arguments consisting of a claim versus a claim with data, warrants or backings but do not contain any rebuttals. The general, hypothetical (and often implicit) logical, statements that serve as bridges between the claim and the data.
Qualifiers	Statements that limit the strength of the argument or statements that propose the conditions under which the argument is true.
Rebuttals	Counter arguments or statements indicating circumstances when the general argument does not hold true. Sometime students are subject to argumentation displays and extended argument with more than one rebuttal or argumentation has arguments with a series of claims or counterclaims with data, warrants or backings with the occasional weak rebuttal. Although argumentation shows arguments with a claim, with a clearly identify able rebuttal.
Backing	Statements that serve to support the warrants (i.e., arguments that don't necessarily prove the main point being argued, but which do prove the warrants are true).

According to Table 1, TAP reflects the students' thinking that are considered in the metacognition includes: 1) Metacognitive awareness: *students know, use, and can interpret scientific explanations of the methodology of force, work and energy*; 2) Metacognitive knowledge: *students understand the nature and development of scientific knowledge of force, work and energy;* and 3) Metacognitive control: *students can generate and evaluate scientific explanations and arguments of force, work and energy.*

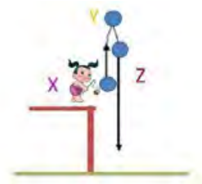
5. Finding and Discussion

The dialogue of students' discussion in six argumentative situations was shown in the blog. The information was then searched to identify episodes of metacognition from dialogical argument. Dialog of the rural students who attend in the discussion or debate comments on various issues of argument on the blog was frequent and there also diverse missed opportunities for giving the quality of the feedback to students. The evidence on all aspects of teaching through blog investigated that the arguments are relied on the use of metacognition for discussion that shown in each component of Toulmin's diagram of Arguments Pattern (TAP) as the following:

5.1 Data (D) Component

The students' comment during weblog learning activities provided the data corpus on the students' argument engagement. Even though, some answer is correct, the claim or explanation of situation is not enough that could be found in the dialogue of the Law of Energy Conservation content in table 2.

Table 2: The dialogue from the comment blog wall about the Law of Energy Conservation content.



Teacher	We are studying the law of energy conservation. This is the question: Throw up an object at the position X vertically until it reaches the highest position Y. Then, the objects drop down to position Z (Figure). Compare the kinetic energy of the object at the position X, Y, and Z?
Group 1: Student 3	I think Y has Kinetic Energy (Ek)=0 because the velocity (Vy)=0
Group 2: Student 1	I think when the object dropped, its velocity increases until it has constant velocity (V) at the position Z.
Group 3: Student 2	I think the velocity at X is higher than at Z (Vx>Vz) and velocity Y is Zero (Vy=0)?
Group 4: Student 3	I think the highest kinetic energy stay at Z and zero energy at Y.
Group 5: Student 1	I think the lowest kinetic energy stay at Y and Z more than X position

The students' scientific conception, scientific attitude, logic and reasoning come from their daily context and previous experience in the study of self-knowledge. Students usually use the data from textbook and the content on the blog in an argument with friends in the classroom. The teachers motivate students to participate in learning activities using the score as encourages them to raise the learners' concentration in classroom. Therefore, it makes students eager to find out the comment to get the most points.

Data (D) Component and the metacognition

The question and answers are carefully chosen to reflect the relationship between the teacher and students in the physics blogging classroom using "Data" component found that; students have used the fact to prove the comment of argument. The student's uses data which articulate and expand the idea in each comment. They post their ideas from other sources than the main blog to refer to their claim. In this situation, the role of students is very highly use metacognitive awareness, knowledge, control because they had many source of data in order to answer the question.

5.2 Claim (C) Component

In the claim component, students have few claims on all questions because some groups of students tend to agree with the claims of friends. Sometimes they would not dare to claim on their own; students do not want to comment the questions themselves because they fear of giving wrong comments. Some comment like as agree with their friend that rarely show the scientific explanation. Therefore, students will agree with the conclusion that the majority believes in and assert it as the true claim. Sometimes students may have several claims and counter claims. Students can claim by explaining the reasons with data and claims attributed to comment questions. Teachers compared the claims or comments of each group at the end of the argument to conclude which of them is correct. This could be seen in the following dialogue: Apply the Law of energy conservation (table 3).

Table 3: The dialogue from the comment blog wall about Apply the Law of energy conservation content.

Teacher

What kind of energy is the bungee jumper involve in? Describe it at each of the following stages:

A. Before Jumping

B. While the bungee cord isn't stretched

C. While the bungee cord is stretched the most

Group 3

(*Discussion in group*)

- I think (A) has only gravitational potential energy of jumper.
- I think same too! and then what do you think about B?
- Hmm !!! Figure (B) if viewed in the figure B "I think in the first period, gravitational potential energy of jumper is decreased, kinetic energy is increased." Anyone agrees with my idea or not?
- I agree it should be possible and figure (C)? Who can explain and how do you know?
- Figure (C) "I think this period has the elastic potential energy of bungee cord and gravitational potential energy of jumper.
- And then how will you explain about gravitational potential energy.
- (C) this period has the elastic potential energy of bungee cord and gravitational potential energy of jumper. Everyone agrees! Ok?

Group 3 conclusion (claim)

- We're thinking (A) has only gravitational potential energy of jumper, (B) in the first period, gravitational potential energy of jumper is decreased, kinetic energy is increased and (C), this period has the elastic potential energy of bungee cord and gravitational potential energy of jumper.

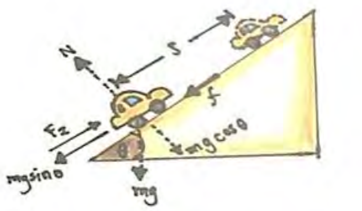
Claim (C) Component and the metacogmition

The questions and answers chosen to reflect on the relationship between teachers and students in the physics blogging classroom using "Claim" component found that; the students' arguments in blogging classroom are interesting. Students comment on each issue with many claims in the force, work and energy content while some groups, students tend to agree with the claims of their friends instead of making claim themselves. As a result, the class mocked or blamed the situation so that the student is the audience more than the claimer. In the dialogs of claim component, students' metacognitive awareness, knowledge and control are very slightly use by owns self. They are flowing groups and teachers more than trust owns self.

5.3 Warrants (W) Component

In the warrants component, when students are into claim and discussion, Students often have evidence for the claim and are used for reliable data to support their comment questions that are about the topic. The students when it is claimed by warrants it consists of reasoning, assumptions and source of comment. Some students are drawing their own on the book and show in class (it is difficult to post the picture in the blog at that time) to warrants their claims or may be more of the comment to other groups. This can be seen from the following dialogue about force and work content (table 4).

Table 4 : The dialogue from the comment blog wall about force and work content.

Teacher	Pushing a toy car on a flat surface (case A) and push it up the slope (case B) with equal speed, distance, and acting force. Which one use more work? And why ?
Group 4	We're think case B cost more work than case A; FB > FA"
	
Group 4 Student 2	A student affirms the claims by drawing (that shown in the figure).

Warrant (W) Component and the metacognition

The questions and answers chosen to reflect on the relationship between teachers and students in the physics blogging classroom using "Warrant" component found that; when students prepare to make claim and discussion, they often have evidence for the claim and use reliable data to support their comments. The students claimed that warrants consist of reasoning, assumptions and source of comment. Some students draw pictures by themselves to warrant their claims. The comment on this physics arguments, most students would agree with or justify, combine ideas organize the discussion. Students' metacognitive knowledge in the warrant component was shown more than metacognitive awareness and control. Student usually used knowledge from their old experience and the teacher teach through a blog to support their reason.

5.4 Qualifiers (Q) Component

The results from six argumentative activities indicated that students use qualifiers that propose the conditions under which the argument is true in physics blogging classrooms rather than propose the limitation of their friend's arguments. In this dialog of qualifiers component, students' metacognitive awareness, knowledge and control couldn't shown metacognition since they are not aware of qualifiers for decide to answer the argumentative question. The students' debate based on their experience and teachers though then they never think of the condition or limitation while discussion in blogging classroom. However, when students need to use picture explanation, they wrote on the book and they could not post on the blog. It could be mention that shown the limitation of arguments via the blog.

5.5 Rebuttals (R) Component

In the rebuttals component, students agree rather than rebut with the hero who can explain or answer questions. Most students who are not proficient enough or have not experience usually consider their friends as heroes without rebuttals. They are maybe afraid to comment because they have that notion that the comment which receives rebuttals is wrong. The student replied that if the comment is wrong, everyone can saw that is a shameful thing for them. This can be seen from the following dialogue about power content (table 5).

Table 5: The dialogue from the comment blog wall about power content.

Teacher	Could you try to solve this problem? The different way to the waterfall at the top of mountain; first is tortuous way and less slope, another is a straight line but more slope. Which way are easy and spent little time?
Group 2 Student 1	I think the first way is easier because it has less inclined slope and uses lesser power.
Group 2 Student 2	I agree with you. Anyone have any answer other than that? *(No comment from other)*
Group 2	Claim that "I think the first way is easier because it has less inclined slope and uses lesser power."

Table 5: The dialogue from the comment blog wall about power content (continue).

Group 3 Student 1	I think the first way is easier and save energy because the driver can slow drive. While the second way need high accelerate for release the high friction force.
Group 3 student 2	I agree with you. Anyone have any answer other than that? *(No comment from other)*
Group 3	Claim that “the first way is easier and save energy because the driver can slow drive. While the second way need high accelerate for release the high friction force.”

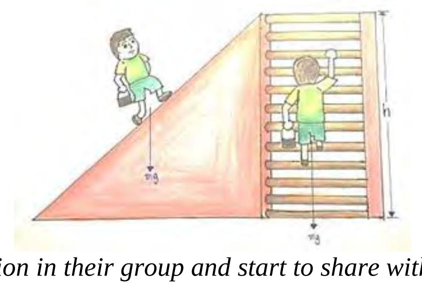
Rebuttals (R) Component and the metacognition

Students rarely make rebuttals to the argument, because most students who are not proficient often wait for ideas from their “hero” friends. They are sometime afraid to comment because they consider an argument that has rebuttals is wrong. The student replied that if the comment is wrong, everyone can seen that is a shameful thing for them and may be they thought that their friends make appropriate explanation. Therefore, students show a little of metacognitive control because they do not like arguments to express their idea, they usually wait scientific explanations from their “hero” friends. However, students who can debate that mean they expressed more metacognitive awareness and metacognitive knowledge which both kind of metacognition appropriate for rebuttals.

5.6 Backing (B) Component

In the backing component, students were subjected for argumentation and had arguments with a series of claims with data, warrants and backings obtained from the references of the theory of energy (kinetic energy and potential energy). Even though reference is a reliable reason to support the comment of them, the most students tend to support ideas from a group of other friends rather than their own comments. Sometime they tend to wait for the hero group started comment and followed it although the answer was incorrect. This can be seen from the following dialogue energy content (table 6).

Table 6 : The dialogue from the comment blog about energy (kinetic energy and potential energy) content

Teacher	A and B carry the basket with the same size and weight. A climbing up the vertical ladder, B climbing up the slope ladder. At the same height, who make higher energy on the basket?  *(After students discussion in their group and start to share with other group)*
Group 2	The energy to carry the basket A and B up to the top have to be equal because they are at the same height level.
Teacher	*(Teacher repeated answer Group 2)* “Group 2 comment that the necessary energy to carry the basket A and B up to the top have to be equal because they are at the same height level” Do you have anything different from group 2 idea?
Group 3	Group 3 agree with Group 2 because energy of the basket A and B are potential Energy (Ep). Due to Ep= mgh the question fix the same m, h, and g, then EpA=EpB
Group 1	Group 1 agrees with group 3 and group 2 because energy of the basket A and B were caused by additive height from potential energy (E_p)
Group 4	Group 4 agrees and support the claim from group 1 because of the equal height and equal weight due to E_p=mgh; $E_pA=E_pB$
Teacher	Concluded that the questions of A and B on the wall, these two cases are the same height. Energy of the two baskets will both increase equally with the equation mgh.

Backing (B) Component and the metacognition

Students are subject for argumentation and had arguments with a series of claims with data, warrants and backings obtained from the references of the theory. Even though reference that is a reliable reason to support their comments. In this "backing" component, it is found that most students tend to support idea from their friends' group rather than their own. Some students tend to wait for another group makes the comments. Sometime they tend to wait for another group, which is a hero group that usually comments, to give ideas and agree with them. From these situation, it seem like students show a little metacognitive awareness, knowledge and control. They do not attempt to give describe the situation or comment by their own.

6. Conclusion and Suggestion

This research study revealed that learned by weblog which provide the scientific content and the argumentative activities could be motivate students to express their own ideas, opinions on situation's discussion. It expressed the students' metacognitive thinking on the situations that demonstrate on the post wall. The rural students claim on various issues commonly learning in the class. The students claim (C) on the Data (D) and warrants (W) component. Other elements comprise of the rebuttals (R) and backing (B) in which students made was based on their experience and content knowledge from the blog and their teacher. For the qualifiers (Q), students usually do not claim in each contents. In metacognition dimension, students express very highly metacognitive awareness, knowledge and control in the use many source of data (D) for debate. While they had very slightly metacognitive awareness, knowledge and control for claim (C) their ideas. Students' metacognitive knowledge in the warrant (W) component show more than other kind of matacognition. In addition, students who can debate that mean they had metacognitive awareness and metacognitive knowledge for rebuttals (R). However, students do not show metacognition in the qualifiers because they never think of the condition or limitation while discussion in blogging classroom.

Interestingly, students prefer to post the comment for discussion, this maybe they never learn with the blog, learning through blog is the new approach for them. Learning on the blog could encourage students to higher thinking from argumentation and it could reduce the problem from students who not dare ask teacher, they could share the understanding on learning with friends via the comments box lessons. However, sometime they are aware that their comment maybe wrong, that make them more serious and lack the confident before share some ideas on the blog. In addition, the study also reveal that students who learned through the lesson on web blog very satisfactory. They were exciting and funning with learning because lessons' blog are interesting, easy to use and quickly access lessons.

These findings suggest that the teacher can create the scientific content blog within the argumentative activities in Physics for enhancing student's metacognition in Thai context.

7. Acknowledgements

This research study was supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission, through the Cluster of Research to Enhance the Quality of Basic Education.

8. References

Aikenhead, G.S, & Ryan, A. G.(1992). The development of a new instrument: "View on science technology-society" (VOSTS). *Science education*, *76*(5), 447-491.

Anantasook, R. (2014). *The effect of Weblog instruction applying social media entitle "Basic C Program language" upon achievement learning and satisfaction of Grade 9 students*. Classroom research. Surin : Rattanaburi school.

Barlow, T. (2008). Web 2.0: Creating a classroom without walls. *Teaching Science*, *54*(1), 46-48.

Driver, R., H. Asoko, J. Leach, E. Mortimer, & P. Scott. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, *23*(7): 5-12.

Erickson, D. (2009). *The effects of blogs versus dialogue journals on open-response writing scores and attitudes of grade eight science students*. Unpublished professional paper. University of Massachusetts Lowell.
Colombo, M. & Colombo, P. (2007). *Blogging to improve instruction in differentiated science classrooms*. Phi Delta Kappan, 60-63.
Duplichan,S. O. (2009). Using weblogs in the science classroom. *Science Scope*. *33*(1), 33-37
Lin, S-S & Mintzes, J.J. (2010).Learning argumentation skills through instruction in socio scientific issues: the effect of ability level. *International Journal of Science and Mathematics Education*, *8*(6), 993-1017.
Luehmann, A. L., & Frink, J. (2009). How Can Blogging Help Teachers Realize the Goals of Reform-based Science Instruction? A Study of Nine Classroom Blogs. *Journal of Science Education and Technology*, *18*(3), 275-290.
Namwar, Y., & Rastgoo, A. (2008). *Weblog as a learning tool in higher education. Turkish Online Journal of Distance Education*, 9(3). Retrieved August 8, 2014, from http://tojde.anadolu.edu.tr/tojde31/articles/article_15.htm
Office of the National Education Commission (ONEC). (1999). *National Education Act B.E. 2542 (1999)*. Bangkok: Office of the National Education Commission (ONEC).
Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge inpsychological science. *American Psychologist*, 44, 1276-1284.
Sudprakone, S. (2012). *The effect of Web-base instruction applying social media entitle "Basic Computer" upon achievement learning of mattayousuksa 1*. Master of Education Thesis, Graduate School, Khon Kean University.
Toulmin, S. E. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.
Yuenyong, C., Jones., A. & Sung-ong, S. (2011). "Student energy conception : empirical results from Thailand and New Zealand." In M. M. H. Cheng, and W. W. M. So. (Eds.). *Science Education in International Contexts*. Science Publisher, 3-16.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Knowledge Propagation in Practical Use of Kit-Build Concept Map System in Classroom Group Work for Knowledge Sharing

Toshihiro NOMURA[a*], Yusuke HAYASHI[a], Takuma SUZUKI[b] & Tsukasa HIRASHIMA[a]
[a]*Graduate School of Engineering, Hiroshima University, Japan*
[b]*Junior High School, Ryogoku Junior High School, Japan*
*nomura@lel.hiroshima-u.ac.jp

Abstract: This study proposes a design of class in which learners collaboratively organize and share what they learned in lessons. Concept mapping is a well-known technique that can help students to create visual representations of the structure of their understanding. This study uses a special kind of concept mapping for knowledge sharing, Kit-Build method. The characteristic of this method is to provide parts to build a concept map for learners. A teacher makes the parts and thus he/she can manage what students should learn and assesses answers automatically. This paper reports the result of classes conducted with Kit-Build concept map and analyzed the knowledge propagation in the collaborative learning. Through collaborative Kit-Build mapping among students the concordance rate of learner maps went up and the score of the maps improved. This indicates that students have shared and corrected their knowledge through the activity. In this study the teacher could realize the understanding of students and taught them what did not understand well based on the feedback of KB map system.

Keywords: kit-build, concept map, collaborative learning

1. Introduction

Development of creativity is one of the important goals in education (Griffin, 2012). On the other hand, Jonassen (Jonassen, 1992) and Scardamalia et al. (Scardamalia, 2012) point out a stage prior to develop creativity. Jonassen defines three stages of knowledge acquisition; introductory, advanced, and expert. At the introductory stage, learners acquire the knowledge that is in well-structured domains and it has only a correct answer for a question. At the advanced stage, learners acquire more advanced knowledge, that is in ill-structured domain and it could have various answers for one question. Expertise is the final stage where experts have more internally coherent and more richly interconnected knowledge structures. Scardamalia specifies two levels of knowledge construction; entry and high level. These levels are the equivalent of Jonassen's introductory and advanced stage, respectively.

This study focuses on knowledge construction for the basis of creative tasks. The aim of this study is to design learning activities in classroom intended to encourage knowledge construction prior to development of creativity. In this study, according to Jonassen and Scardamalia et al., the requirements for the activities are that students can build their own knowledge inductively and check correctness of it in interaction with other students. They need to correct their knowledge as necessary and finally they share correct knowledge. This shared and correct knowledge is expected to lead their next activities for development of creativity. From the viewpoint of the definition of collaborative learning by Dillenbourg (Dillenbourg, 1999) the required elements are symmetry of action and status of students, interaction among them and inductive thinking.

To satisfy the requirement this study adopts collaborative learning with Kit-Build concept map (Yamasaki, 2010). Kit-Build concept map is a kind of closed concept map construction system in which learners make concept maps from parts provided from a teacher. Although this is similar to Expert skeleton concept map, the difference from it is that Kit-Build concept map requires learners to completely rebuild a concept map from fragmented pieces of concepts and links. Learners relive the consideration of relation among concepts in the target domain. This method is also effective in diagnosis of learners' understanding. The concept map represents the structure of what the teacher want students to understand. A learner's understanding can be evaluated as the concordance rate between a map made by the learner and one made by the teacher. General concept mapping uses open system in which there is no restrictions to build concept maps and constructors use any nodes and links. In this case, it is difficult to diagnose concept maps (Herl, 1999). There are studies show the effectiveness of Kit-Build concept map in evaluation of and feedback to learners based on it (Sugihara, 2012) (Yoshida, 2013).

Maldonado et al. proposes a system to support this individual and collaborative creative knowledge building by using concept map and ICT literacy (Maldonado, 2012). He tracked and analyzed the flow of knowledge that is created as a result of individual pre and post-concept maps at the personal computer and group concept maps construction at the multi-touch tabletop. As mentioned above, ICT literacy can offers the possibilities to extend the support to students and moreover can be an analytical tool by which teachers support students. Also, the purpose of this study is to support knowledge building by using ICT literacy. However, our study intends to different level from his study.

This paper reports practical use of Kit-Build concept map system in classroom group work for knowledge sharing and knowledge propagation among them in the learning. The remainder of this paper is organized as follows. The next section explains Kit-Build map system that is the core technology of this study. Section 3 presents the design of classes the authors conducted. These classes are conducted in a junior high school as a part of usual classes by one of the co-author. Section 4 shows the result of the class and analysis of the data. Last section concludes this paper.

2. Kit-Build Concept Map

Concept maps (Novak, 2006) are graphical tools for organizing and representing knowledge or understanding. A concept map includes concepts and relationships between concepts indicated by a connecting line linking two concepts. Two concepts linked with a relation represent a proposition. Concept map is effective for facilitating learning and for enabling learners to create visual representations of their comprehensive structure. This has enormous significance in enabling to evaluate and share learners' knowledge.

Kit-Build concept map (KB map) is a framework to build and diagnose concept maps (Yamasaki, 2010). In this method, learners build concept maps (learner map) by assembling provided parts. The parts are generated by decomposing an ideal concept map (goal map) that is prepared by a teacher as the goal of his/her teaching. The characteristic of KB maps is that it is possible to compare and overlap them because both of the maps, learner maps and goal maps are composed of the same components. Comparing them, system can calculate the concordance rate of the map as the score of learner maps automatically. KB map provides the teacher with information about learners' comprehensions and differences of their thoughts on the map that overlaps all the learner maps (group map). Also, because of providing the same parts among learner and goal maps, it is easy for students to compare their maps in discussing and negotiating understanding with maps and KB map clarify the differences between a learner's and others' understanding in creating the map collaboratively.

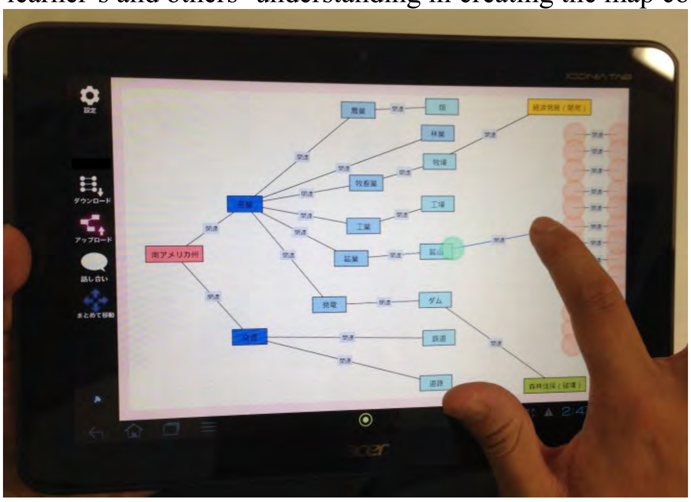


Figure 1. KB map editor.

There is a system to realize interaction based on Kit-Build method. This system is called KB map System. This system is a web application with two client systems: KB map Editor and KB map Analyzer. KB map Editor provides an environment to make a learner map. This system can be used with tablet pc (Sugihara, 2012). Figure 1 show the picture of the KB map editor. This permits students to carry their maps and to show their maps to the other students for discussing. KB map Analyzer has function to overlap learner maps and teachers understand his/her students' comprehension with this system.

3. Outline of Practice in a Junior High School Japanese Social Studies Class

3.1 the Purpose and the Design Principle of the Class

This study designed and conducted the lessons in which students use KB map. This is collaboration with the junior high school teacher that is one of the authors of this paper. He has a desire to make students enhance their understandings with both of collaboration and teaching.

In the lessons what students do is the following two things: to organize their own knowledge on KB map inductively from materials and what they have learned in the previous lessons and to compare and correct their knowledge represented on KB map through discussion. After that, the teacher explains the correct answers compared with students' KB maps. Through this process, this lesson expects that students having incorrect knowledge learn correct one from the others and that the teacher identifies students' misunderstandings still remained after discussion and teach them carefully.

3.2 Design of the Class

In this study, the procedure of the class was the following: Firstly the teacher does a review of the topics in the previous classes. The teacher shows some pictures related topics and explain it. Secondary, students make KB map (pre-map) individually. In this phase they organize their knowledge inductively from materials and what they have learned in the previous lessons. Thirdly, they go into a small group and work together. They discuss the difference among their maps and make a KB map of the group (collaborative-map) collaboratively. Fourthly, they make modify their own KB map (post-map) individually again if necessary. Lastly, the teacher gives feedback about the topic with the collaborative-maps students have made. In this procedure, pre-maps underlie collaborative-maps, and post-maps reflect collaborative-maps. The procedure of the class is shown in Figure 2.

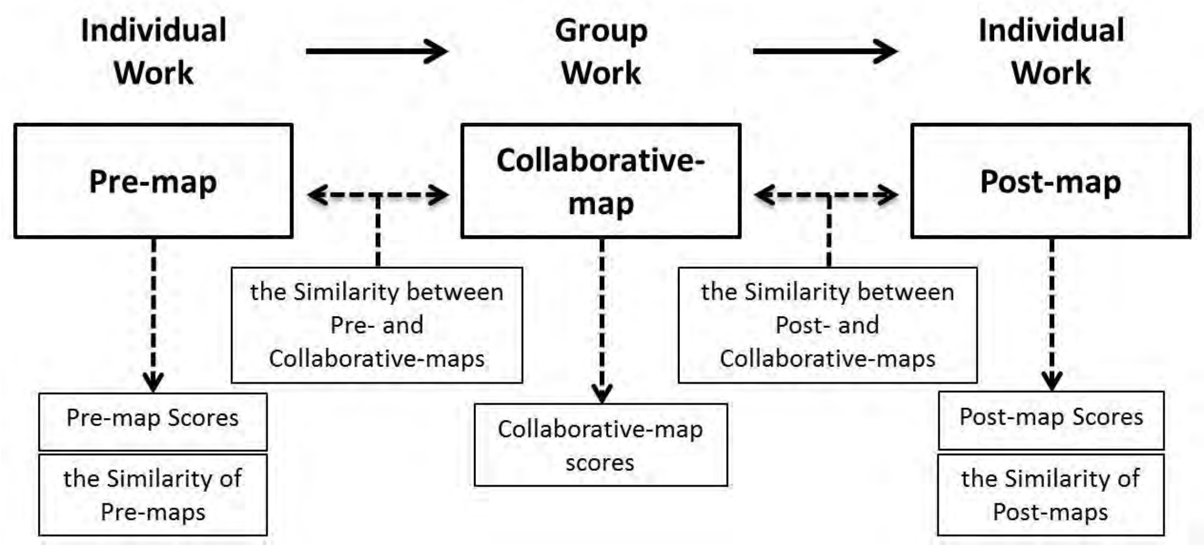


Figure 2. The Flow of This Class and the Data from the Maps.

3.3 Hypothesis

In this research, three hypotheses were formed. These are the followings:

Hypothesis 1: After group work, students have much more common understanding about the topic than before. (in a group the similarity of post-maps is higher than the one of pre-maps)
Hypothesis 2: After group work, students have much more correct understanding than before. (post-map scores are higher than pre-map scores)
Hypothesis 3: The teacher can realize the understanding of students based on group map that is made by KB map system.

3.4 Participants and Procedure of the Practice

Three lessons were conducted for three first grade classes in a junior high school. The participants are 76 Japanese students who are 12 or 13 old in total from three classes. The number of students in the class A is 26, in the class B is 25 and in the class C is also 25. In the classes, the teacher regularly uses concept maps for teaching social studies and thus, the students were used to make concept maps in learning it. In addition to that, they also have ever used KB map system in studying social studies lesson. Therefore, it is not so difficult for them to make concept maps and to use KB map system.

The subject domain of the class conducted in this study was the unit of South America in geography. The theme was the dilemma of economic development and deforestation. The students had mainly studied economic development in South America and had had knowledge about deforestation within the bounds of common sense.

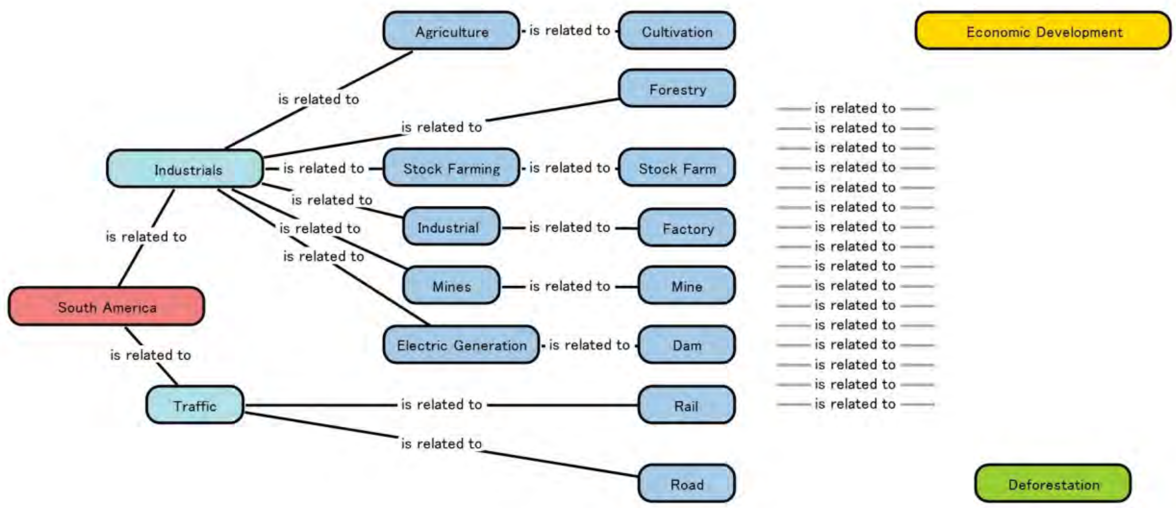


Figure 3. Kit of This Class.

In this practice, firstly the teacher makes a concept map about industrials and traffics in South America on the blackboard with the students. This map making is a task to recall what they studied in the previous classes. They made the parts of the map in several lessons. Secondly the teacher provided tablet computers with students. The kit of this class is shown in Figure 3. With a tablet, each student made a KB map including the map on the blackboard. The part of the map that is on the blackboard is already constructed on the KB map and the students are required to link the rest of it with the separated concepts and links. The separated concepts are "economic development" and "deforestation." The students are required to consider which concepts connect to economic development and/or deforestation. The kit includes enough unfixed links to connect all the concepts about industrials and traffics with both of them. In the goal map, all the separated concepts connected to both of economic development and deforestation. However, the teacher did not clearly tell the students that they did not have to use all the links. Therefore, the students must have considered whether each concept was connected with economic development and/or deforestation or not. In the previous class, the students had learned economic developments and industrial, however, they had not learned deforestations and traffics. The teacher gave the students some documents about deforestations, and students could refer it when they made KB map. Thus, this assignment consisted of both recall and generation. Making the connection to economic development is a recall task because the teacher explained in the previous lessons. On the other hand making the connection to deforestation is a generation task because they need to consider with their pre-existing knowledge and the provided pictures. As the first step of the map making, the students make a map individually as a personal opinion. This map is called “pre-map,” here. After that, they went into a small group of four or five students and discussed the difference among their maps and made a map collaboratively as an agreement. This map is called “collaborative-map,” here. Next, they could modify their map individually if they need after group work. This map is called “post-map.” Finally, the teacher explained the correct answer in comparison with the collaborative-maps. Here, the teacher used the group map that was made by overlapping every collaborative-map.

4. The Result and Consideration

4.1 Testing the Hypotheses

Hypothesis 1: After group work, students have much more common understanding about the topic than before (in a group the similarity of post-maps is higher than one of pre-maps)

Figure 4 and table 1 show the comparison among the average concordance rates between the pre- or post-maps in each group. It is calculated by the following equation:

$$\frac{\text{(the number of the pairs of students who connected the same link to the others)}}{\text{(the number of the propositions in goal map)} \times \text{(the number of group members)}}$$

The average concordance rate among post-maps is higher than the one among pre-maps. There is a significant difference between them in every class (in the class A: two-sample t-test, $t(5) = 4.8906$, $p<0.01$, in the class B:two-sample t-test, $t(5) = 4.0865$, $p<0.01$ and in the class C:two-sample t-test, $t(5) = 6.3571$, $p<0.01$). This indicates that students shared their understanding and built a consensus in each group in some way.

Furthermore, Figure 5 and table 1 show the comparison among the average concordance rates between pre- or post-maps and the collaborative-map in each group. This is calculated by the following equation:

$$\frac{\text{(the number of the propositions that are common between the individual and collaborative map)}}{\text{(the number of the propositions in the goal map)}}$$

The concordance rate between post- and collaborative-maps is higher than the one between pre- and collaborative-maps. There is a significant difference between them in every class (in the class A: Wilcoxon signed-rank test, $n=26$, $V=228$, $p<0.01$, in the class B: Wilcoxon signed-rank test, $n=24$, $V=153$, $p<0.01$, in the class C: Wilcoxon signed-rank test, $n=25$, $V=253$, $p<0.01$). This indicates many students changed their individual-maps following their collaborative-maps after discussion.

Hypothesis 2: After group work, students have much more correct understanding than before (post-map scores are higher than pre-map score).

Figure 5 and table 2 show the comparison between the average scores of the pre-maps and the post-maps. A score of the learner map indicates degree of similarity between the learner map and the goal map. It takes the value of 0 to 1. If the score is 1, it means the learner map is completely same as the goal map. The score is calculated by the following equation:

$$\frac{\text{(the number of the correct propositions in a learner map)}}{\text{(the number of the propositions in the goal map)}}$$

The average scores of the post-maps are higher than ones of the pre-maps in every class. There is a significant difference (in the class A: Wilcoxon signed-rank test, n=26, V=210, p<0.01, in the class B: Wilcoxon signed-rank test, n=24, V=148.5, p<0.01, in the class C: Wilcoxon signed-rank test, n=25, V=253, p<0.01) between them. This indicates that the understanding of students was improved through making the KB map in a group.

Moreover, this study compares the score of the pre-map to the score of the collaborative-map using the Wilcoxon rank sum test. A significant difference was found between the score of the pre-map and the score of the post map in every class (in the class A: n=26, U=31.5, p<0.05, in the class B: n=24, U=34, p<0.01, in the class C: n=25, U=34.5, p<0.01). The scores of the collaborative-maps are higher than the scores of the pre-maps. This finding suggests that the knowledge created collaboratively is more correct than the average knowledge created individually.

Consequently, because individual learner maps came closer to collaborative-maps, the degree of the similarity among group members' maps increased, and because the score of the collaborative-map was higher than the score of the pre-map, the students' knowledge improved.

Consequently, the summary of the results is the followings:

(1) the concordance rate of individual-maps in groups increased through group work, and
(2) the concordance rate between individual-maps and collaborative-map also increased through group work, and
(3) the score of individual-map is improved.

These can be considered as the followings:

(1') the students share their understanding in their group, and
(2') they made a collaborative-map as their shared understanding, and
(3') their shared understanding improves their individual-understanding.

The authors carried out a questionnaire survey to clarify how students had made the collaborative-maps and what is the reason if students had changed their maps after group work. This questionnaire includes two questions about decision making in group work and improvement of individual-map after it: "How did you make your collaborative-map?" and "How did you change your individual-map?" Students could choose from three options: "(1) by majority vote", "(2) agreement on the others' opinion" and "(3) by just following others' opinion". The 60-70 percentages of the students chose the second option on both questions. This suggests that students place importance on the agreement with the opinion by discussing when they change their opinion and make their collaborative knowledge. In addition to that, the relation between pre-maps and collaborative-maps shows data supporting the suggestion. Table 3 indicates the relation between pre-maps and collaborative-maps. In each group, the proposition in the pre-maps is categorized according to the number of the students that they have the same propositions in their pre-maps. Moreover, the propositions are categorized by the correctness of the propositions in their pre- and collaborative-maps. Note that students did not always choose the proposition by majority vote. Of course, they chose by majority vote, however, there are some cases that they chose from minority.

Hypothesis 3: The teacher can realize the understanding of students based on group map that is made by KB map system.

In this class, the teacher made a group map from the collaborative-maps, and compared the group map with the goal map to show how many groups did not make the correct links. Figure 7 is the group map composed of the lacking links in the collaborative-maps this study obtained in the class-A. According to the group map, the teacher could get the information about the propositions that the groups had a lack of understanding toward and were divided on the opinion of. The teacher supposed that the students had lack of understanding toward the propositions about the new concepts, railways and roadways. However, according to the group map, there was the lack of understanding toward the various propositions that the teacher did not suppose. The greatest numbers of the propositions that students had the lack of understanding about were, in class A "Forestry relate to economic developments." and "Factory relate to deforestations.", in class B "Farm relate to economic developments." and "Farm relate to deforestations.", in class C "Forestry relate to economic developments." and "colliery relate to deforestations.". In this class, teacher could explain mainly about these propositions. After the lessons the teacher said that this is the first time to get information about understanding of students on time in classroom. Although he had tried to investigate individual thought of students during group work, it needed the help of other many teachers and it is difficult to organize the result during the lesson. This time he satisfied the information and give instruction based on it during the lesson.

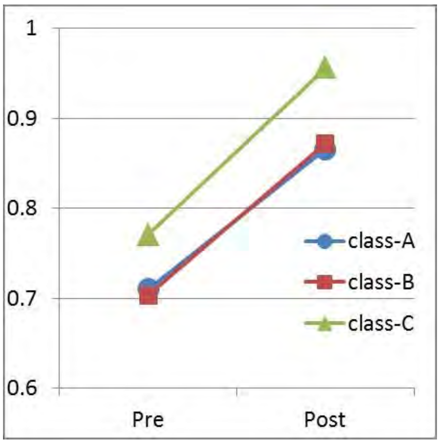


Figure 4. The Average Concordance Rate among Individual Maps in the Groups.

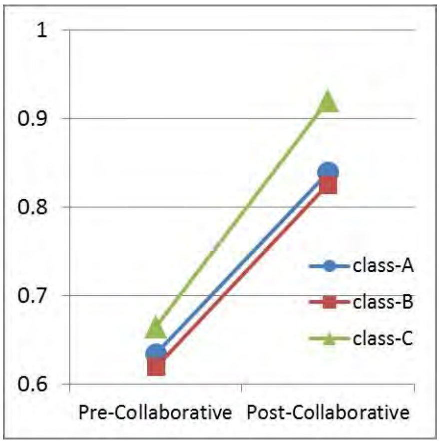


Figure 5. The Average Concordance Rate between the Individual and Collaborative Map.

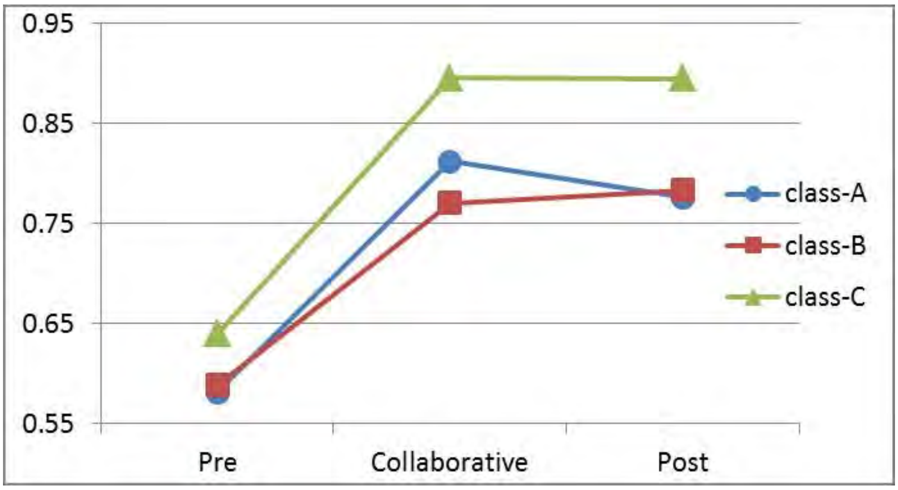


Figure 6. The Average Map Scores in three classes.

Table 1: Concordance Rates.

	Pre-maps in the group	Post-maps in the group	Pre-Collaborative	Post-Collaborative
Class-A	0.7104(*SD*=0.0399)	0.8661(*SD*=0.0462)	0.6346(*SD*=0.2170)	0.8389(*SD*=0.1563)
Class-B	0.7031(*SD*=0.0733)	0.8724(*SD*=0.0858)	0.6198(*SD*=0.1861)	0.8255(*SD*=0.1403)
Class-C	0.7708(*SD*=0.0696)	0.9563(*SD*=0.0296)	0.6650(*SD*=0.2064)	0.9200(*SD*=0.0875)

Table 2: Map Scores.

	Pre-map	Collaborative-map	Post-map
Class-A	0.5817(*SD*=0.2206)	0.8125(*SD*=0.1936)	0.7764(*SD*=0.1922)
Class-B	0.5885(*SD*=0.2395)	0.7708(*SD*=0.1407)	0.7839(*SD*=0.1758)
Class-C	0.6400(*SD*=0.2226)	0.8958(*SD*=0.1164)	0.8950(*SD*=0.0967)

Table 3: the Relation between the Pre-Maps and Collaborative-Maps.

The type of result	Unanimous accord				Majority				Even	
Which is common in pre-maps	Correct agreement		Incorrect agreement		Major correct opinion		Major Incorrect opinion		Even	
Correct or Incorrect in collaborative-maps	C	I	C	I	C	I	C	I	C	I
The number of Propositions	51	2	3	7	82	6	35	23	67	12

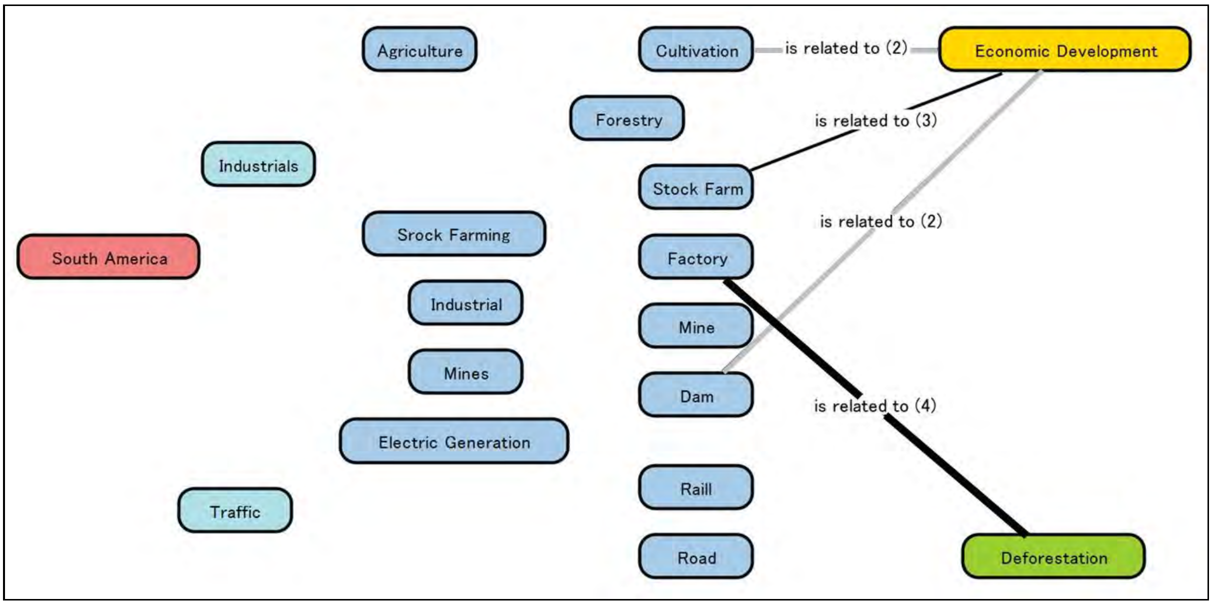


Figure7. Group Map in the Class-A.

4.2 Analysis of the Knowledge Propagation in Collaborative Learning

As mentioned above, the students' knowledge improved on the whole. However, this result is just average. Now let's take a look at the several students. The students did not improve all of their misunderstanding. Although, some students corrected their misunderstanding of some items through group work, other students went wrong in understanding of some items after group work even though they had had correct understanding. About these phenomena, this study examined the knowledge propagation in more detail.

Table 4 shows the patterns of change of understanding and the distribution of them. There are eight patterns of change among pre-, collaborative- and post-maps by the correctness of the propositions in each map. The distribution is derived from about tallied the 1200 propositions in the pre-, collaborative- and post-maps.

The most common pattern is Pattern-A: students had had correct understanding of a proposition from the beginning and then had kept it at the end. This pattern accounts for over half of all propositions. Because the average score for the pre-map was about six out of ten and it improved in post map, this may be the reason why their map score was kept.

The second most common pattern is Pattern-E: students had had incorrect understanding of a proposition at the beginning and then they changed it to correct one following the agreement in their group. This pattern accounts for over two out of ten propositions and accounts for about half of the propositions that the students got the wrong answer about in their pre-maps. This may provide significant share of the reason why their score for pre-map improved.

The third most common pattern is Pattern-H: students had had incorrect understanding of a proposition at the beginning and then they had kept it by the end. The agreement in their group is also incorrect, therefore, they did not have chance to change their understanding. This pattern accounts for over one out of ten propositions. Most of this case happened when no one had had the correct answer in the group. This suggests that it was hard for students to improve their understanding without the member who had had the correct understanding. Especially, the KB map in this study requires students to make an either-or decision. Therefore, this does not produce the diversities of opinion. The students often reached the agreement on the wrong answer from the beginning. It seems that the students did not discuss the propositions in this situation. A similar result was found in another study with the KB map.

These data supports the result of hypotheses testing in the last section. Although the above patterns are majority, there are some other patterns. In some patterns the students changed their answer from correct one to incorrect one in the post-map.

In Pattern-C and Pattern-F the students did not follow the decision in their collaborative-map and kept their own understanding that is different from the group's decision. In Pattern-D the students changed their correct understanding to incorrect one following the incorrect group decision. The number of Pattern-C is twice the number of Pattern-D. This may suggest that students prefer to keep their correct answer than to get groups' wrong answer.

Pattern-G and pattern-B mean that although a proposition in the pre-map was same to the propositions in the collaborative-map they change the proposition in the post map. They changed their thought after group work. However, there is no data to identify the reason in this study. In order to identify it, it is necessary to gather data about what they talked about in the group work other than data about KB map.

Regarding the propositions that changed from the pre-map to the post map, such as pattern-B, -D, -E and –G, nine out of ten propositions changed from incorrect to correct. Moreover, in this case, nine out of ten propositions changed to the propositions that were made collaboratively.

Table 4: the Patterns of Change of Understanding.

Pattern	Pre	Collaborative	Post	Rate (%)
A	Correct	Correct	Correct	53
B	Correct	Correct	Incorrect	1
C	Correct	Incorrect	Correct	4
D	Correct	Incorrect	Incorrect	2
E	Incorrect	Correct	Correct	23
F	Incorrect	Correct	Incorrect	6
G	Incorrect	Incorrect	Correct	2
H	Incorrect	Incorrect	Incorrect	9

5. Conclusions and Future Work

This study proposed learning activities with KB map in a group for knowledge construction and sharing prior to development of creativity, and implemented it to investigate the knowledge propagations among students. The result supports the three hypotheses, "Students share their understandings through discussion," "Students' understandings get close to the correct one" and "Teachers can teach based on the group map." Moreover, knowledge propagation in their learning is analyzed from the pre-maps to post- ones by way of collaborative- ones. Consequently, there were every possible patterns and the majority of the propositions changed to correct, however, some propositions changed to wrong.

From these results, students could acquire and correct their knowledge through the proposed group work. In knowledge propagation analyzed in Section 4.2 most of students keep their correct knowledge and correct it through discussion. This tells that the proposed method did not give negative effect on students in this case. KB map gives students with common parts to build a concept map. This might be effective in discussion. If the students built concept maps freely it would be difficult for them to organize their thought at a short time. This is highly controversial issue and must be investigated in the future.

The aim of the classes proposed in this study is to build common knowledge among students before forming the creative opinions in the next class. Therefore, what is important is to let learners be with a full understanding of basic knowledge required for forming creative opinions. To that end, learners must keep the correct knowledge and change the incorrect knowledge to correct. In the proposed classes, the students improved their understanding through group work with KB map after learning the subject. Also, with KB map, the teacher could teach about the propositions that the students did not understand well and conduct personal coaching based on the data from KB map the students made.

These classes were implemented in the regular classes and there is no data to improve the improvement of students' understanding. It is necessary to measure their understandings with other tests and to compare the effectiveness with learning methods. However, at least, the teacher conduct this classes said that he feels the reciprocal teaching by students have worked well and KB map is useful not only learners but also teachers to organize what to learn and recognize students' understanding.

Of course, the aim of this study is not general in the research area of collaborative learning. Most of studies focus on discovery learning in which students build their own knowledge other than knowledge given by teachers. However, Scardamalia distinguishes it into "guided discovery" and "knowledge building" (Scardamalia, 2007). The proposed lesson in this study can be considered as a kind of the former one. In this type of discovery learning learners try to discover a solution of a problem that have correct answers. Although, this does not require creativity in the true sense, they are required creative thinking. The authors think that the advantage of this style of discovery learning is the existence of correct answer and teachers can make clear assessment on the answers of learners. This is expected to use for training of creative thinking. In fact, KB map building in the lessons required the students to think about things other than they had learned. They needed to make a conjecture from the materials given in the lessons. Although this is closed in the KB map and the materials, thinking process is similar to inductive thinking in true discovery learning.

In addition to that, further research on this collaborative learning with KB map would clarify the knowledge propagation. A further direction of this study will be to support the classes for acquiring advanced knowledge by collaborative learning with KB map.

References

Dillenbourg P. (1999) What do you mean by collaborative learning?. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1-19). Oxford: Elsevier.

Griffin, P., Care, E. & McGaw, B. (2012). The changing role of education and schools. *Assessment and teaching of 21st century skills* (pp. 1-16). Dordrecht: Springer.

Herl, H. E., O'Neil, H. F. Jr., Chunga, G.K.W.K., & Schacter, J. (1999). Reliability and validity of a computer-based knowledge mapping system to measure content understanding. Computers in Human Behavior, 15, 315–333.

Jonassen, D. H. (1992). Evaluating constructivistic learning. *Constructivism and the technology of instruction: A conversation*, 137-148.

Maldonado, R. M., Kay, J., & Yacef, K. (2012). Analysing knowledge generation and acquisition from individual and face-to-face collaborative concept mapping. Proc. Of the Fifth Int. CMC, 17-24.

Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. Florida Institute for Human and Machine Cognition, 1.

Scardamalia, M., & Bereiter, C. (2007). Fostering communities of learners and knowledge building: An interrupted dialogue. In Campione, J. C., Metz, K. E. & Palincsar, A. S. (Eds.), *Children's learning in the laboratory and in the classroom: Essays in honor of Ann Brown* (pp.197-212). Mahwah, NJ: Erlbaum.

Scardamalia, M., Bransford, J., Kozma, B. & Quellmalz, E. (2012). New assessments and environments for knowledge building, *Assessment and teaching of 21st century skills* (pp.231-300). Dordrecht: Springer.

Sugihara, K., Nino, Y., Moriyama, S., Moriyama, R., Ishida, K., Osada, T., Mizuta, Y., Hirahima, T. & Funaoi, H. (2012). Interactive use of kit-build concept map with media tablets. In IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education, 325-327.

Yamasaki, K., Fukuda, H., Hirashima, T., & Funaoi, H. (2010). Kit-build concept map and its preliminary evaluation. In Proc. of ICCE, 290-294.

Yoshida, K., Sugihara, K., Nino, Y., Shida, M., & Hirashima, T. (2013). Practical use of kit-build comcept map system for formative assessment of learners' comprehension in a lecture, Proc. of ICCE2013, 892-901.

Liu, I. F. & Wei, C. W. (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Exploration of Improving Efficiency of Synchronous Discussion: e-Case Live Show

I-Fan LIU[a*], Chun-Wang WEI[b]
[a]*Department of Information Management, Oriental Institute of Technology, Taiwan*
[b]*Department of Management Information System, Far East University, Taiwan*
*ifliuG@gmail.com

Abstract: In traditional discussion style, the presentation and discussion of the case is usually led by the team leader, but there are still some members who do not actively participate in the discussion, causing the atmosphere of participation to become disjointed. In order to encourage students actively interactive with instructor and peers, the purpose of the study is to propose an advanced synchronous learning style, e-Case Live Show, which integrates live webcast technique with hosting by turns to support learners' case-based learning. 58 college students participated in this study. From the research results, we found that students from EG showed significantly higher average scores on Perceived Learning Autonomy and Perceived Interactivity than CG. The e-Case Live Show developed by this study to support case-based learning is indeed better than the traditional TV talk show discussion format for giving learners more autonomy and interactive learning.

Keywords: Case-based learning, synchronous discussion, learning autonomy, learning interactivity

1. Introduction

In the past ten years, Information and Communication Technology (ICT) has developed rapidly, reducing the distance of communication around the world. Learning no longer has to be fixed in time and place, but rather can be flexible. With the age of Web 2.0, social networking services, such as Facebook and YouTube make social interaction and sharing more and more frequent (Hughes, Rowe, Batey, & Lee, 2012). In Taiwan, many professors in business schools conduct the case-based learning in parts of their management courses in order to let students experience the conditions described in cases. Real business operating conditions can be even simulated through various software. One such program, for example, involves students taking a course on production management, in which they can discuss problems faced by enterprises within the context of a simulated supply chain system through case studies, and go through business operations, such as purchase of materials, sale of products, and inventory (Liu & Young, 2006). This software simulation practice will improve students' learning efficiency. The objective is to go beyond traditional teaching methods of teachers giving their courses to students from a podium.

According to Tian & Hong's (2003) research, students in Taiwan are less active in asking questions than those in United States and Europe. They also observe the following phenomena in Taiwan's college students in class of case-base discussion: (1) In traditional teaching method, teachers stand on the podiums to impart knowledge and students sit down to accept it. The problem is that students usually just sit there and do not ask questions or speak up; (2) In traditional discussion style, the presentation and discussion of the case is usually led by the team leader, but there are still some members who do not actively participate in the discussion, causing the atmosphere of participation to become disjointed; (3) Unless in group discussion, students prefer not to interact very often. They usually learn related knowledge individually and tend to become lone learners; (4) Students can learn and finish their works through group discussion and increase their interaction. However, it is likely that each student only cares about his or her work, integrating with the knowledge from others is difficult. In order to encourage students actively interactive with instructor and peers, the purpose of the study is to propose an advanced synchronous learning style, e-Case Live Show, which integrates live webcast technique with hosting by turns to support learners' case-based learning. Therefore, we also want to explore whether the efficiency of synchronous discussion style like e-Case Live Show had better than traditional discussion style.

2. The Call-in programs and learning

The TV talk shows originated in the United States in the 1930s from political talk shows and it also known as political call-in programs. The reason why the political call-in program is broadly popular to this day is mainly because it offers citizens and participants in the program with opportunity to engage in two-way communication and express personal opinions (McLeod, Dietram, & Moy, 1999). The call-in programs can raise the participation of discussions, leading to higher ratings. It is a kind of learning when people observe others' behavior or thought (Horowitz, 1993). Bandura (1986) has proposed observed learning and imitation in a social learning theory. That is, observed learning indicates that an individual, as an onlooker, observes others' behavior in a process of obtaining learning. Imitation, on the other hand, is the process of learning other individuals' or groups' behavior in observed learning (Chang, 1996). Bandura (1986) argues that the most important concept of social learning theory is vicarious learning, which means people learning and understanding what behaviors are shown in certain conditions by observing others' behaviors and outcomes.

3. The Case-based Learning Method

The case-based learning method originated from Harvard University in 1890, and has been practiced in the business courses at the Harvard Business School for a long time (Barnes, Christensen, & Hansen, 1994). This method is a unique type of teaching method, in which, as Shulman (1992) points out, educators use cases to familiarize students with the conditions under which cases happen. Through simulating the conditions and discussing possible decisions, students try to develop their capability to solve problems.

The case-based method, considered an effective way to combine theory and practice in business education (Harrington, 1995; Knirk, 1991), focuses on the interaction among peers and the solution from brainstorming (Levin, 1995). There are generally two arrangements in the case-based method, prearranged groups and *ad hoc* groups (Wassermann, 1994). The former involves a teacher gathering students with different characters in a group, such as talkers and non-talkers. The latter involves students grouping to work with others they are familiar with. Although the case-based method is a self-learning one in which students actively participate in discussion, learners still have to interact and cooperate with others in order to achieve better learning effects (Blumenthal, 1991).

4. e-Case Live Show

The synchronous discussion of *e-Case Live Show* is based on live webcast technique to conduct a videoconferencing. We adopt *i-Share* technique, a synchronous teaching module based on multimedia collaboration system developed by SUNNET Corporation (http://www.sun.net.tw). This module provides multi-player audio and video instant communication like a well-known synchronous discussion tool JoinNet (Chen, Ko, Kinshuk, & Lin, 2005; Chen & Ko, 2010) to offer online face-to-face teaching and learning for instructors and learners (as Figure 1).

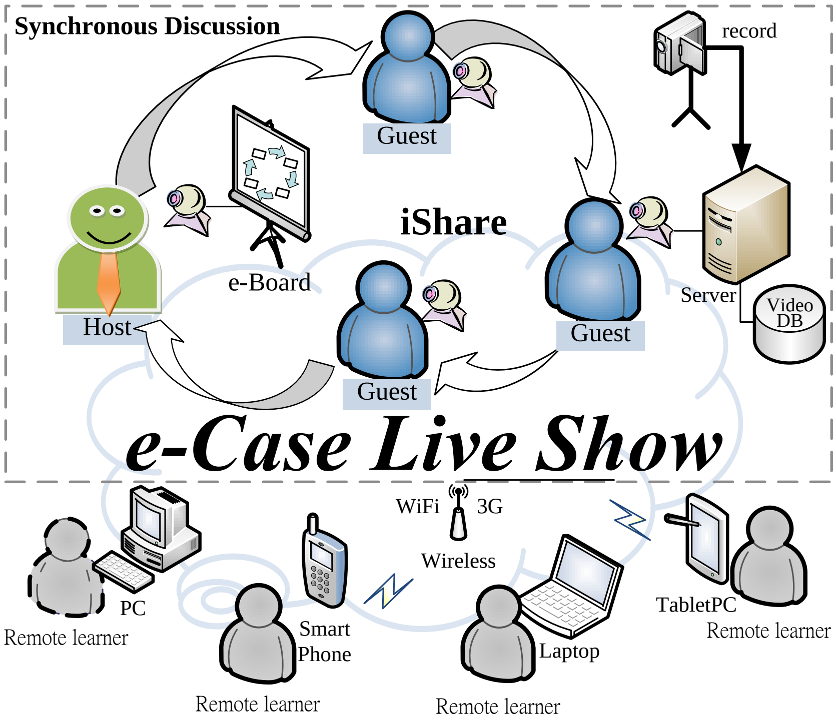


Figure 1. e-Case Live Show synchronous learning environment

5. Methodology

In order to evaluate students' interactive learning efficiency for the e-Case Live Show, the five-point Likert scale and independent sample *t-test* was adopted in this quasi-experiment. The participants were divided into an experiment group (EG) and a control group (CG). Thirty-six students volunteered to enter the EG group; 22 students were in the CG group. For the EG group, each team member had to play the role of leader by turns to host and discuss the content of the case study he/she was responsible for. For the CG group, the format was like a traditional TV talk show in which one leader hosted and discussed the content of the case study with guests. For both EG or CG groups, all of the students enrolled in the class called "Theory of Management and Case Study", and the professor carried out a blended form of instruction in the course for one semester. The section on management theory (2/3) was taught face-to-face by the professor; the case study (1/3) was carried out through the e-Case Live Show. This was a learner-centered and discussion-based course. Students had to look for partners, and each group consisted of 3 to 4 members. One group was responsible for one case study from the textbook. To prevent the situation where few team members speak or participate, each EG member took turns to be the host during the course of the one-hour online discussion; students in remote areas could also call in to participate in interactive discussions with the host.

6. Findings

From the research results, we found that students from the EG group showed significantly higher average scores on Perceived Learning Autonomy ($M_{EG} = 4.22 > M_{CG} = 3.64$, $t = 2.41$, $Sig = .019$ *) and Perceived Interactivity ($M_{EG} = 4.42 > M_{CG} = 3.86$, $t = 2.73$, $Sig = .008$ **) than the CG group. This suggests that the students from the EG group experienced much more autonomous and interactive learning in the whole process of case-based learning as compared to the CG group. In terms of Perceived Coordination ($M_{EG} = 3.69 < M_{CG} = 4.23$, $t = -2.01$, $Sig = .049$ *), we found that the EG students showed significantly lower average scores than CG ones. This indicates that because students from the EG group had to play a host role by turns, their coordination performance was not as efficient as the CG host. On the other hand, there was no significant difference on Ease of Use ($M_{EG} = 4.53 > M_{CG} = 4.18$, $t = 2.73$, $Sig = .147$) between the two groups.

7. Conclusions

The e-Case Live Show described in this study to support case-based learning clearly outperforms the traditional TV talk show discussion format for giving learners more autonomy and opportunities for interactive learning. Studies in the future could be focused on how to provide support for the students with poor oral communication skills when they are hosting the discussion of a case study, such as by encouraging assistance from their peers to increase their perceived coordination, so that these students can better express themselves.

Acknowledgements

This work was supported by the National Science Council of Taiwan under contract numbers NSC 101-2511-S-269-002-MY3.

References

Barnes, L., Christensen, R., & Hansen, A. (1994). *Teaching and the Case Method*. Harvard Business School Press, Boston.

Bandura, A. (1986). *Social Learning Theory*. New York: General Learning Press.

Blumenthal, J. (1991). Use of the case method in MBA education. *Performance Improvement Quarterly, 4*(1), 5-13.

Chang, C. H. (1996). *Contemporary Psychology*. Donghua Press, Taipei.

Chen, N. S., Ko, H. C., Kinshuk, & Lin, T. (2005). A model for synchronous learning using the Internet. *Innovations in Education and Teaching International, 42*(2), 181-194.

Chen, N. S., & Ko, L. (2010). An Online Synchronous Test for Professional Interpreters. *Educational Technology & Society,* 13(2), 153-165.

Harrington, H. L. (1995). Fostering reasoned decisions: Case-based pedagogy and the professional development of teachers. *Teaching and Teacher Education, 11*(3), 203-214.

Horowitz, E. (1993). Talk Show Politics: The Match that Rekindles American Democracy? Paper Presented to Association for Education in Journal and Mass Communication Convention.

Hughes, D. J., Rowe, M., Batey, M., & Lee, A. (2012). A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Computers in Human Behavior, 28(2), 561-569.*

Knirk, F. G. (1991). Case materials: Research and practice. *Performance Improvement Quarterly, 4*(1), 73-81.

Levin, B. B. (1995). Using the case method in teacher education: The role of discussion and experience in teachers' thinking about cases. *Teaching and Teacher Education*, *11*(1), 63-79.

Liu, I. F., & Young, S. S. C. (2006). A study of implementing Web-Based Learning Systems to enhance learning for the Supply Chain Management (SCM) course in higher education. *The 14th International Conference on Computers in Education (ICCE2006),* Beijing, China.

McLeod, J. M., Dietram, A. S., & Moy, P. (1999). Community, communication, and participation: The role of mass media and interpersonal discussion in local political participation. *Political Communication*, *16*(3), 315-336.

Shulman, J. H. (1992). *Case methods in teacher education*. New York: Teacher College Press.

Tian, N. C., & Hong, M. C. (1998). Computer-mediated communication and higher education teaching innovation. *The Journal of Taipei Teachers College*, 11, 295-312.

Wassermann, S. (1994). *Introduction to Case Method Teaching. A Guide to the Galaxy*. NY: Teachers College Press.

Game playing as a strategy to improve Team Cohesion, support for collaborative U-Learning

Pei-Yu Cheng[a], Wen-Yen Wang [b] & Yueh-Min Huang[a*]

[a] *Department of Engineering Science, National Cheng Kung University, Taiwan*

[b] *Department of Computer Science and Information Engineering, Kun Shan University*

* huang@mail.ncku.edu.tw

Abstract: This study investigated how a game playing strategy embedded in collaborative u-learning activities affect students' Team Cohesion and Learning attitude and their learning performance. Participants in this study were fifth grade students in elementary school (N=64); they were randomly assigned into the experiment group and the control group. The Experimental Group and Control Group will be assigned different English activity in the first stage, and in the second stage, two groups conducted the same collaborative u-learning activities. The results indicate that game playing strategy can greatly enhance students' Team Cohesion and Learning Attitude in our study.

Keywords: Game playing, Team Cohesion, U-Learning, Collaborative

1. Introduction

In Taiwan, English is the important second language. Consequently, enhancing student's English ability has become important educational policies. For learners, vocabulary knowledge and reading ability are the most important components of performance in second language learning(Huckin, 1995), Folse (2004) indicated that vocabulary is essential to English learning for second-language learners. Therefore, it's a vital issue to develop a sound approach by which to assist students in learning English vocabulary. In recent years, with the rapid evolution of computer technology and the prevalence of mobile devices, learning has changed transformed from traditional classroom learning to digital and mobile learning. For vocabulary learning, many studies have tried to explore, how to use mobile devices to support vocabulary learning (Chen & Chung, 2008; Hong, Hwang, Tai, & Chen, 2014; Y. M. Huang, Huang, & Lin, 2012).

According to related studies, ubiquitous learning is an effective teaching methods, because combining u-learning can effectively trigger learners' learning motivation(Chiou, Tseng, Hwang, & Heller, 2010; Jeng, Lu, & Lin, 2010; Ogata & Yano, 2004) and enhance their learning performance(El-Bishouty, Ogata, & Yano, 2007; Rogers et al., 2005). Liu and Chu (2010) indicated that incorporating ubiquitous into the English learning activities could achieve a better learning outcomes and motivation.

According to this viewpoint, many researchers have been interested in ubiquitous learning, and has been successfully applied to many subjects (Y.-M. Huang & Chiu, 2014; Y.-M. Huang, Huang, &

Wu, 2014). Researchers have pointed out that using mobile devices may enhance collaborative learning and promoted better interactions between students in the activities because students can use it to coordinate collaboration between them. Lai and Wu (2006) argued that using mobile devices can effectively enhance students' attitudes and performance in collaborative learning.

However, an earlier study show that there are many problems of online collaborative learning, such as difficulties in communication, the lack of shared, and the imbalance(Roberts & McInnerney, 2007; Tseng & Yeh, 2013). Therefore, the dynamic within the team is also an important consideration in building Team Cohesion(Kwon, Liu, & Johnson, 2014). Consequently, this study proposed a game playing strategy which is embedded in collaborative u-learning activities for helping students to building Team Cohesion.

2. Research Methods

2.1 Participants

This study investigated how a game playing strategy embedded in collaborative u-learning activities affect students' Team cohesion and their learning performance. Participants in this study were fifth grade students in elementary school (N=64); they were randomly assigned into the experiment group and the control group.

Figure 1 shows the experimental flow of this study. Before the experiment, this study distributed pretest to subjects to find if there is significant difference between two groups. After the subjects filled out the questionnaire, the researcher conducted experiment on two groups. In the first stage, the students in the experimental group conducted collaborative crossword game as Figure 2. While the students in the control group conducted ordinary learning activities. In second stage, the two groups conducted the same collaborative u-learning activities for learning English vocabulary as Figure 3.

Figure 1. Experimental procedure

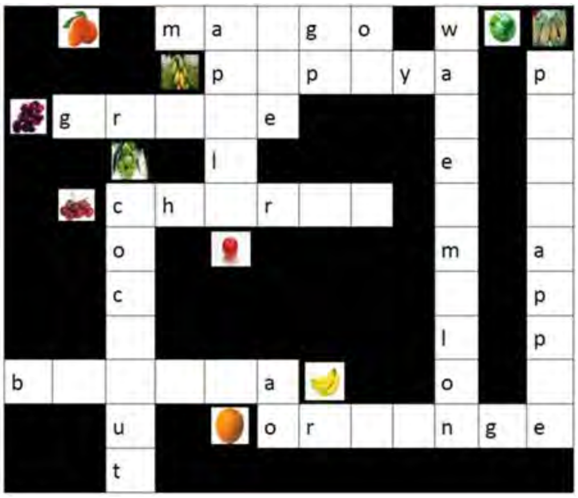


Figure 2. Collaborative crossword game

Figure 3. Collaborative u-learning activities for learning English

3. Result and Discussion

This study used the pre-test scores as covariate for one-way ANOVA to avoid any interaction effects from the pre-test on the students' learning outcomes. As listed in Table 1, the pre-test mean for the experimental group was 57.97 and 57.81 for the control group. The results did not reach a level of significance, f=0.005, p> .05. It suggests that homogeneity of two groups of variables is supported.

Table 1: The one-way ANOVA results for the pre-test scores.

	Group	N	mean	SD	f
Pre-test	Experimental	32	57.97	9.233	0.005
	Control	32	57.81	9.046	

According to Table 2, the post-test mean for the experimental group was 72.19, and 70.31 for the for the control group. Results of statistical analysis showed a no significant difference in learning performance between two groups, f=2.416, p>.05. This result suggests that learners in the first stage with different learning activities didn't produce a significant difference within learning performance.

Table 2: The one-way ANOVA results for the post-test scores

	Group	N	mean	SD	f
Post-test	Experimental	32	72.19	4.568	2.416
	Control	32	70.31	5.070	

*$p < .05$

As listed in Table 3, the experimental group students' Team Cohesion were significantly higher than the control group students, t=8.99, p<.001). That is, the students who conducted game playing strategy had higher Team Cohesion than those who conducted ordinary learning activities in the first stage.

Table 3: The one-way ANOVA results of Team cohesion

	Group	N	Mean	SD	f
Team Cohesion	Experimental	32	4.36	0.38	8.99**
	Control	32	4.12	0.23	

*p < .05

As listed in Table 4, the experimental group students' Learning Attitude were significantly higher than the control group students, t=-17.384, p<.001). That is, the students who conducted game playing strategy had higher Learning Attitude than those who conducted ordinary learning activities in the first stage.

Table 4: The one-way ANOVA results of Learning Attitude

	Group	N	Mean	SD	f
Learning Attitude	Experimental	32	4.33	0.23	17.384***
	Control	32	3.95	0.46	

*$p < .05$

4. Conclusion

In this study, we proposed a game playing strategy which is embedded in collaborative u-learning activities for helping students to building Team Cohesion. Based on the experimental results, we found that game playing strategy can greatly enhance students' Team Cohesion and Learning Attitude which is consistent with the findings of the past research(DeVries & Edwards, 1973; Huyen & Nga, 2003; Randel, Morris, Wetzel, & Whitehill, 1992; Roberts & McInnerney, 2007; Sharan & Sharan, 1976). Thus, we suggest that teachers can using game playing as a strategy to improve Team Cohesion, support for collaborative U-Learning. This study has certain limitations, such as manpower. The limitation of this study is too small of sample size in this experiment. In the future research, we will consider some experiments with a larger sample size of students and conduct more complete research investigate the relationships between Team Cohesion and Learning Styles.

Acknowledgements

This study is conducted under the "Cloud computing systems and software development project (3/3)" of the Institute for Information Industry which is subsidized by the Ministry of Economy Affairs of the Republic of China.

References

Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education, 51*(2), 624-645.

Chiou, C. K., Tseng, J. C. R., Hwang, G. J., & Heller, S. (2010). An adaptive navigation support system for conducting context-aware ubiquitous learning in museums. *Computers & Education, 55*(2), 834-845.

DeVries, D. L., & Edwards, K. J. (1973). Learning games and student teams: Their effects on classroom process. *American Educational Research Journal, 10*(4), 307-318.

El-Bishouty, M. M., Ogata, H., & Yano, Y. (2007). PERKAM: Personalized knowledge awareness map for computer supported ubiquitous learning. *JOURNAL OF EDUCATIONAL TECHNOLOGYAND SOCIETY, 10*(3), 122.

Folse, K. S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*.

Hong, J.-C., Hwang, M.-Y., Tai, K.-H., & Chen, Y.-L. (2014). Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety. *Computers & Education, 72*, 313-322.

Huang, Y.-M., & Chiu, P.-S. (2014). The effectiveness of a meaningful learning-based evaluation model for context-aware mobile learning. *British Journal of Educational Technology*, n/a-n/a. doi: 10.1111/bjet.12147

Huang, Y.-M., Huang, S.-H., & Wu, T.-T. (2014). Embedding diagnostic mechanisms in a digital game for learning mathematics. *Etr&D-Educational Technology Research and Development, 62*(2), 187-207. doi: 10.1007/s11423-013-9315-4

Huang, Y. M., Huang, S., & Lin, Y. T. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education, 58*(1), 273-282.

Huckin, T. (1995). Second Language Reading and Vocabulary Learning.

Huyen, N. T. T., & Nga, K. T. T. (2003). Learning vocabulary through games. *Asian EFL Journal, 5*(4).

Jeng, Y. C., Lu, S. C., & Lin, H. M. (2010). Implementing situated learning theory into e-Learning: Vocational special education students' learning outcomes. *Int J Digit Learn Technol, 2*, 100-119.

Kwon, K., Liu, Y.-H., & Johnson, L. P. (2014). Group Regulation and Social-emotional Interactions observed in

Computer Supported Collaborative Learning: Comparison between good vs. poor collaborators. *Computers & Education*.

Lai, C. Y., & Wu, C. C. (2006). Using handhelds in a Jigsaw cooperative learning environment. *Journal of Computer Assisted Learning, 22*(4), 284-297.

Liu, T.-Y., & Chu, Y.-L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*(2), 630-643.

Ogata, H., & Yano, Y. (2004). *Context-aware support for computer-supported ubiquitous learning*.

Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming, 23*(3), 261-276.

Roberts, T. S., & McInnerney, J. M. (2007). Seven problems of online group learning (and their solutions). *Educational Technology & Society, 10*(4), 257-268.

Rogers, Y., Price, S., Randell, C., Fraser, D. S., Weal, M., & Fitzpatrick, G. (2005). Ubi-learning integrates indoor and outdoor experiences. *Communications of the ACM, 48*(1), 55-59.

Sharan, S., & Sharan, Y. (1976). Small-group teaching.

Tseng, H. W., & Yeh, H.-T. (2013). Team members' perceptions of online teamwork learning experiences and building teamwork trust: A qualitative study. *Computers & Education, 63*, 1-9.

-

The relationship between parents addicted to mobile phone and adolescent addicted to the Internet

Ying ZHOU[a], Xiao ZHANG[a], Jyh-Chong LIANG[b], Chin-Chung TSAI[c]
[a] *School Educational Technology, Beijing Normal University, China*
[b]*Graduate Insitute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan*
[c]*Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan*
yzh@bnu.edu.cn

Abstract: Previous studies had revealed that the mobile phone addiction existed among adolescents. However, in China, along with the popular of smart phone, the situation of addicted to mobile phone changed and emerged some new features. Meanwhile, the influence between parents addicted to mobile phone and children addicted to the Internet has been realized gradually. The study evaluated the relation of both of them based on the factor analysis. The participants comprised one of parents of 311 middle school students in Beijing. The results revealed that they had significant positive correlation (r=0.32, p<0.001), and also Internet addiction of children was related to the parents' activities that parents participated by mobile phone, like Game & Video and Job & study.

Keywords: mobile phone addiction, Internet addiction, exploratory factor analysis, parents

1. Introduction

The popularity of mobile phone remains higher and higher, especially after emergence of smart phone. However, research has showed adolescents and adults are possibly addicted to the mobile phone(Lin 2010). In recent years, many scholars began to focus on mobile phone addiction.

A scale has developed to evaluate the mobile phone addiction for Korean adolescents. 20 items were divided into three factors: Factor 1 (withdrawal/ tolerance), Factor 2 (life dysfunction) and Factor 3 (compulsion/ persistence) explaining 55.45% of total variance. An item, for example, 'When I can't use a mobile phone, I am exasperated', belongs to Factor 1 (Koo 2009). A study on mobile phone addiction has done aiming at 269 Taiwanese female university students. The study presented two scales: one is the Mobile Phone Usage Behavior Scale (MPUB) surveyed the frequency of mobile phone use ('How many messages do you send pre day?'); the other is Mobile Phone Addiction Scale (MPA) by revising Young (1998) Internet Addiction Scale ('I neglect school assignment to spend more time using mobile phone'). And then there are 11 items classified three factors, named as (1) Time Management and its Problems (an example is While using mobile phone, you would think 'just give me some more minutes'). (2) Academic Problems in School and its Influence (an example is 'Because I spend too much time on mobile phone, my school work or my marks are influenced). (3) Reality Substitute (an example is 'Before having to do something I always check the mobile phone to see whether there are missed calls or text messages) explaining 65.95% of total variance. Results are that MPA can positively predict MPUB, and female university students with MPA will make more phone calls and send more text messages (Hong, Chiu et al. 2012). Also, With mobile phone addiction becoming serious, the sleep quality will be worse (Sahin, Ozdemir et al. 2013). Mobile phone addiction and academic-success has significant negative correlation， and addiction to mobile phone and the level of depression has a positive correlation (Cagan, Unsal et al. 2014).

As the use of smart phone more and more widely, application software developed for smart phone is more and more rich. It is fashionable for Users to surf the Internet by mobile phone and to use software downloaded from the Internet. Thereupon, the activities that users participate by phone are changing. People began to complain some new activities lend their friends and family to be more addicted to phone than phone calls and text messages. Moreover, accompanied by the activities, some new features of mobile phone addiction appeared. Particularly, it is probably harmful for students if parents addicted to mobile phone. Nevertheless the research is still a blank.

In this background, by absorbing and summarized the previous research results, the study aims at developing a scale to survey the addiction to mobile phone for adult. It is just for parents in this study. And further to explore the relationship between parents addicted to the mobile phone and students addicted to the Internet.

2. Method

2.1 Participants

The participants in this study comprised one of parents of 311 middle school students in Beijing. There were 104 fathers and 199 mothers. Based on what the parents filled in the questionnaire, there are 185 first grade students of junior middle school and 126 second grade students of senior high school. The child of 157 parents was schoolboy, and of 149 was schoolgirl. These parents were from a middle-level school. Among them, education background of 26.8% was high school or below, 30.7% was junior college, 30.1% was undergraduate, 7.8% was master and 4.6% was doctor.

2.2 Instrument

The instrument used in this study mainly included two scales, one was for mobile phone activities and addiction of parents, and other was for Internet addiction of their children.

Referring to previous research, the study developed a scale to measure parents' mobile phone activities and addiction (MPAA). The study considered some popular activities among smart phone adult users of china in the last two years. In addition, besides the items about withdrawal reaction which were contained by majority of Internet addiction scales' the study designed some items about smart phone addiction aiming at the addictive features of the new activities.

To measure Internet addiction of children, the Internet Addiction Test (IAT) (Tsimtsiou, Haidich et al. 2014) by Young was chosen in the study. IAT consisted of 20 items and a 6-ponits Likert scale, scoring as 0-5, measured each item. The total score ranged from 0 to 100, and represented the degree of an Internet user addicted to Internet. However, this study invited parents to evaluate the situation and degree of Internet addiction for their children, so I in the each item was revised as my child.

2.3 Data analysis and procedures

In analyzing the scale about the usage of mobile phone, exploratory factor analysis was used to reduce the items. The total score of IAT was calculated based on the sum of 20 Internet addiction items, and represented the Internet addiction degree of student. The Pearson correlation analysis of MPAA and IAT were used for explore the relation between parents addicted to mobile phone and students addicted to the Internet.

3. Results

3.1 Exploratory factor analysis for MPAA

To clarify the structure of the parents' MPAA, an exploratory factor analysis with a varimax rotation was performed. Table 1 presented the results, revealing seven factors: Social & information, Job & study, Game & video, Instant usage, Psychic gratification, Boring usage, Withdrawal reaction. The eigenvalues of the seven factors from the principle component analysis were all larger than 0.85. Items with a factor loading of less than 0.50 and with many cross-loadings were omitted. A total of 21items were retained in the final version of the MPAA, and the total variance explained was 71.96%. The Cronbach's alpha coefficients for the seven factors were 0.76, 0.81, 0.66, 0.72, 0.94, 0.82, and 0.61, and the overall alpha was 0.91, which revealed high reliability of these factors. Table 1 also presented the factor means and the standard deviations of MPAA. As shown in Table 1, the " Social & information" factor (mean=2.58) were scored highly by parents among the first three factors about activities that parents participated in by mobile phone, and "Instant usage" factor (mean=2.65) were scored highly by parents among the last four factors about parents addicted to mobile phone. Parent's score on "Withdrawal reaction" factor (mean=0.81) were lowly.

Table1: Rotated factor loading, Cronbach's alpha values, factor means, and standard deviations for the three factors of activities that parents participated in by mobile phone and the four factors of parents addicted to mobile phone in MPAA

		Factor 1	Factor 2	Factor 3
Factor 1: Social & information, alpha=0.76, mean=2.58, S.D.=1.26				
Social & information1	I access Social Networking Services by mobile phone.	0.73		
Social & information2	I search information on the Internet by mobile phone.	0.72		
Social & information3	I watch the Internet news by mobile phone.	0.70		
Factor 2:Job & study, alpha=0.81, mean=1.73, S.D.=1.23				
Job & study1	I assist my work by mobile phone.		0.81	
Job & study2	I study by mobile phone.		0.63	
Job & study3	I download application software by mobile phone.		0.63	
Job & study4	I use all sorts of small software on mobile phone.		0.59	
Factor 3: Game & video, alpha=0.66, mean=1.21, S.D.=1.08				
Game & video1	I play game by mobile phone.			0.83
Game & video2	I watch video by mobile phone.			0.61
Game & video3	I read novel by mobile phone.			0.57

		Fact-or4	Fact-or5	Fact-or6	Fact-or7
Factor4: Instant usage, alpha=0.72, mean=2.65, S.D.=1.34					
Instant usage1	When the phone rang, I will take it up immediately.	0.84			
Instant usage2	I take a look at the phone first after getting up.	0.70			
Instant usage3	I take a look at the phone before bed.	0.60			
Factor5: Psychic gratification, alpha=0.94, mean=1.50, S.D.=1.43					
Psychic gratification1	When using those applications on phone that I like, I feel relaxed.		0.91		
Psychic gratification2	When using those applications on the phone that I like, I am happy.		0.90		
Factor6: Boring usage, alpha=0.82, mean=1.84, S.D.=1.35					
Boring usage1	Encountering something I do not know how to do, I will pick up the phone to do something.			0.74	
Boring usage2	At bored time, I will find something to look at or get something to do from the phone.			0.73	
Boring usage3	when I have nothing to do, I will pick up the phone to do something.			0.56	

Factor7: Withdrawal reaction, alpha=0.61, mean=0.81, S.D.=0.98		
Withdrawal reaction1	In some cases not suitable for using a phone, I will look for opportunities to use the phone as much as possible.	0.72
Withdrawal reaction2	In front of my family, I will hide my thirst for the use of phone.	0.70
Withdrawal reaction3	When I cannot use the phone for a period of time, I feel restless.	0.53

Loadings less than 0.50 were omitted. Overall α=0.91, total variance explained=71.96%.

3.2 Relation between MPAA and IAT

To find out the relation between MPAA and IAT, a person correlation analysis was performed in this study, as shown in Table 2.

At first, the results showed that the activities that parents participated in by mobile phone were all significantly correlated with the features addicted to mobile phone (p<0.001). The correlation between boring usage and the three factors about the activities were all the highest than three other factors about the features.

And then, the results indicated that IAT of students had the highest correlation with Boring usage of parents (r=0.36,p<0.001), implying that higher degree of Internet addiction for students tended to higher boring usage for their parents. Next, parents' instant usage (r=0.21, p<0.001) and withdrawal reaction (r=0.22, p<0.001) also had higher correlation with IAT of students. Moreover the last four factors were all significantly correlated with IAT of students. Using the sum total of 11 items belongs to the last four factors represented the degree of parents addicted to mobile phone. The sum total was positively related to IAT of students (r=0.32, p<0.001), namely, the relation was established between parents addicted to mobile phone and students addicted to the Internet.

On the other hand, students' IAT and parents' Job & study (r=0.13, p<0.05) existed significant positive correlation, as well as to students' IAT and parents' Game & video (r=0.16, p<0.01). It means that what parents participated by mobile phone had the relation with students addicted to the Internet. In three factors about the activities, the factor Game & video, in other words, the activities like playing game, watching video or reading novel on mobile phone for parents existed an influence with Internet addiction of students. And the same influence existed between the factor Job & study of parents and Internet addiction of students.

Table 2: The correlations between MPAA and IAT

	Social & information	Job & study	Game & video	Instant usage	Psychic gratification	Boring usage	Withdrawal reaction
Social & information				0.43^{***}	0.40^{***}	0.54^{***}	0.35^{***}
Job & study				0.42^{***}	0.47^{***}	0.52^{***}	0.35^{***}
Game&video				0.28^{***}	0.39^{***}	0.50^{***}	0.35^{***}
IAT	0.09	0.13^{*}	0.16^{**}	0.21^{***}	0.16^{**}	0.36^{***}	0.22^{***}

*p<0.05, **p<0.01, ***p<0.001

4. Discussions and conclusions

The study developed the survey measuring the situation of parents using mobile phone, containing activities that the parents participated in by mobile phone, above all smart phone, and some new features of parents addicted to mobile phone, such as instant usage and boring usage of phone. It is different with previous scale for mobile phone addiction because pervious scales were generally developed based on referring to Internet addiction scales being composed of expending time, affecting social function and withdrawal reaction. However, there were some problems. Expending time during using mobile phone was often fragmentary. It led to be unconscious the passing of time for people. Secondly, the items about affecting social function in previous scales were mainly about adolescents, for instance, “I can’t do my homework or study because of cell phone use”(Koo 2009). Nevertheless, since adult’s identities were complex, their social functions were diverse. It is difficult to descript clearly by using several items. So the study redeveloped some items to measure the features about parents, namely adults. Relying on the exploratory factor analysis, it is proved that these items had good reliability and structural validity. In the present seven factors, the first three factors can describe what parents do on mobile phone; furthermore the last four factors can evaluate the mobile phone addiction of parents.

The correlation relation was shown between the seven factors and Internet addiction of students in this study. The results were consistent with our assumptions. It could be because, on the one hand, addicted to mobile phone of parents set children a bad example; on the other hand, due to addicting to mobile phone, parents ignored their children and got little time with them. These caused that some psychological needs of children cannot be satisfied, such as sense of safety, belonging, intimacy, etc. The missing of the psychological needs had the relation with the Internet addiction of adolescent (Yen, Yen et al. 2007, Yao, He et al. 2014).

The study suggested parents themselves to make some rules for using mobile phone. For example, to avoid the feature of instant usage of addiction to mobile phone, parents should strongly demand themselves not to touch the phone during 1-2 hours after getting up and before sleeping. At bored time, according to thinking deeply in advance, designedly to do some positive suitable activities so as not to take the phone to kill time. In addition, it is important to note, even though you are working or studying by phone, what your children notice is just that you immerse yourself in the phone. Further, the phone should be a tool. If you feel happy when using it and feel bothered when no using it, you have to be careful, because the phone may be grabbing your attention against your children. In a word, parents themselves need to notice the mobile phone usage, since it had the relation with addicting to the Internet of children.

Acknowledgements

This study is supported in part by the 2012 project of Education Science "the 12th five years" planning, Beijing under grant number AGA100209

References

Cagan, O., et al. (2014). Evaluation of College Students' the Level of Addiction to Cellular Phone and Investigation on the Relationship Between the Addiction and the Level of Depression. *4th World Conference on Psychology, Counseling and Guidance* (Wcpcg-2013), 114: 831-839.

Hong, F. Y., et al. (2012). A model of the relationship between psychological characteristics, mobile phone addiction and use of mobile phones by Taiwanese university female students. *Computers in Human Behavior, 28*(6): 2152-2159.

Koo, H. Y. (2009). Development of a Cell Phone Addiction Scale for Korean Adolescents. *Journal of Korean Academy of Nursing, 39*(6): 818-828.

Lin, J. C. (2010). Popularity, Funding for Health-Effect Research and Cell-Phone Addiction. *Ieee Antennas and Propagation Magazine, 52*(2): 164-166.

Sahin, S., et al. (2013). Evaluation of mobile phone addiction level and sleep quality in university students. *Pakistan Journal of Medical Sciences, 29*(4): 913-918.

Tsimtsiou, Z., et al. (2014). Greek Version of the Internet Addiction Test: A Validation Study. *Psychiatric Quarterly, 85*(2): 187-195.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Developing an instrument to assess teachers' belief, confidence and motivation about digital game-based learning

[a]Yu-Hsuan CHANG[*] & [a]Meng-Jung TSAI
[a]*Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan*
M10111017@mail.ntust.edu.tw

Abstract: The purpose of this study was to develop an instrument regarding teachers' belief, confidence and motivation of digital game-based learning pedagogies. A pool of 30 items was developed by revising items from Grove, Bourgonjon and Van Looy (2012)'s scale. A survey was used to examine and validate the instrument's structure and reliability. The participants include 134 in-service and pre-service teachers in Taiwan. An exploratory factor analysis with principal component method with varimax rotation was adopted to explore the factor structure of the instrument. Finally, a questionnaire was developed with a total of 13 items under three dimensions: confidence, belief and motivation. The internal reliabilities, alpha coefficients, were adequate for the overall scale (0.87) and for all the three subscales (0.83, 0.77 and 0.81, respectively). Applications of this instrument have been discussed at the end of this paper.

Keywords: Digital game-based learning, teacher education, confidence, motivation, belief

1. Introduction

Over the past few years, digital games have played an important role in education. Games not only bring entertainment to us, but also spark innovative thinking. Digital games offer teachers and students potentially power learning environments (Oblinger, 2004). Nowadays, digital game-based learning is widely used in education all over the world. There are several features about digital game-based learning: rules, goals, feedback, challenge and interaction, for instance (Prensky, 2001).

Because of these characteristics, digital games are used to try to raise students' motivation and promote students' performance. A study conducted by Papastergiou pointed out the gaming approach can enhance both students' knowledge of computer memory concepts and motivation. Educational computer games are used to improve learning environment, regardless of students' gender (Papastergiou, 2009). Besides, Divjak and Tomic´ (2011) revealed that adopting computer games can promote mathematics learning. To integrate computer games into learning may bring some good effects on learning, such as arising students' learning motivation and attitude; and further, it can make students have better learning outcomes than before.

From 2000 to 2010, there are more and more researches about digital game-based learning have been raised (Hwang & Wu, 2012). In recent years, the researches have indicated that digital game-based learning can bring others positive impacts for education. For instance, the experimental result showed that a game-based problem solving environment not only highly inspired users' concept learning, but greatly created entertainment (Cai, Bharathi, Klein, & Klein-Seetharaman, 2003).

As we can see, many studies have indicated that it is worthwhile adopting game-based learning to engage students. But in reality, there were not totally have positive responses. The prior research made a survey to survey 1048 in-service and 656 pre-service teachers about using video games in classroom. The result showed that there was less than half of the in-service teachers use gaming in their teaching (Ruggiero, 2013).

There are many innovative digital games, however, there are used in real classroom not very often. Hence, the purpose of this study is to develop a scale to assess teachers' belief, confidence and motivation about game-based learning pedagogies.

2. Method

2.1 The development of previous scales

A scale was developed by Frederik, Jeroen and Jan, to explore the teachers' adoption intention of digital games (Grove, Bourgonjon, and Van Looy, 2012). The scale focused on usefulness, ease of use and behavioral intention. The factors tended to investigate the technology acceptance of digital game-based learning.

Another scale is developed by Hsu, Su and Liang, they used a framework to measure the preschool teachers' Technological Pedagogical Content Knowledge-Games (TPACK-G) as well as their acceptance of digital game- based learning (Hsu, Su and Liang, 2012). The pervious scales all explore that teachers' can use digital games to represent the content or not.

2.2 The development of this scales

To develop this scale, a pool of 18 items was collected by mainly adapting items from developing by Frederik, Jeroen and Jan in 2012 and writing new items (Grove, Bourgonjon, and Van Looy, 2012). They proposed the following six following subscales for exploring teachers' adoption intention of digital games: usefulness, ease of use, experience, learning opportunities, curriculum relatedness, behavioral intention, including a total of 19 items. The rating range of the scale is from "strongly disagree" to "strongly agree" and is presented in a five-point Likert scale. The original reliability (Cronbach's alpha) coefficients are 0.86, 0.74, 0.89, 0.93, 0.69 and 0.92, respectively for usefulness, ease of use, experience, learning opportunities, curriculum relatedness, behavioral intention.

The items was developed in this study mainly basing upon these items, and this study change the term in the scale terms into "Confidence", "Belief", "Motivation". Besides, the author developed 8 additional items for the initial pool of items. The scale in this study hoped to develop a scale with low to high levels. As a result, the initial pool items in the scale included a total of 30 items. Each statement was measured on five-point Likert scale.

2.2.1 Participants

The participants in this study included 134 in-service and pre-service teachers in Taiwan. Among the participants, 71% are female, and 29% are male. Their age was from 21 to 51, and the mean age of all participants was 31.4 years (SD = 8.25).

3. Illustrations

3.1 Factor analysis

An exploratory factor analysis with principal component method with varimax rotation was adopted to explore the factor structure of the instrument. As seen in Table 1, the eigen values of the three factors were larger than one: 5.281, 1.735 and 1.146. Our three factors was retained in the final version of the scale, and they accounted for 62.8% of variance. Items were retained only when their loading was greater than 0.50 for the relevant factor and less than 0.50 for the non-relevant factor. The initial 30 items were reduced to 13 items. The internal reliability index, alpha coefficients, were adequate for the overall scale (0.87) and for the three subscales, 0.83, 0.77 and 0.81.

Table 1: Rotated factor loadings and Cronbach's alpha values for the three factors of the scale.

Item	Factor 1: Confidence	Factor 2: Belief	Factor 3: Motivation
Factor 1: Confidence α=0.83			
17	0.813		
23	0.741		
24	0.814		
27	0.742		
Factor 2: Belief α=0.77			
6		0.672	
19		0.609	
20		0.549	
26		0.766	
28		0.758	
Factor 3: Motivation α=0.81			
4			0.813
9			0.680
14			0.682
22			0.660
Eigen value	5.281	1.735	1.146
% of variance	40.62	13.35	8.82

3.2 Rerained items on the scale

The retained items and responding subscales was shown in Table2. A detailed description of the three subscales was presented below:

1. Confidence subscale: assessing teachers' experience and confidence of adopting digital games in teaching.
2. Belief subscale: measuring teachers' values and beliefs of digital game-based learning.
3. Motivation subscale: assessing teachers' potential and willingness to adopt digital game-based learning for teaching in the future.

Table 2: Retained items on the scale.

Item No.a	Subscale	Question
17	Confidence	It is difficult for me to integrate digital games into the instructions.[c]
23	Confidence	I have no idea how to integrate digital games into the curriculum.[c]
24	Confidence	I consider to use digital games in my courses.[c]
27	Confidence	I always feel upset when I use digital games in my classrooms.
6	Belief	Digital games offer opportunities to experiment with knowledge.[b]
19	Belief	Digital games increase my productivity in my job.[b]
20	Belief	Digital games offer opportunities to motivate students.[b]
26	Belief	Using digital games makes teaching easy.[b]
28	Belief	Digital games provide students opportunities to solve problems.
4	Motivation	It is easy for me to design a digital game- based curriculum.
9	Motivation	Digital games enhance my effectiveness in my job.[b]
14	Motivation	I'm planning to use digital games in my classrooms.
22	Motivation	When I prepare my teaching plans, I'll link the curriculum with digital games.

[a] The item number indicates the item order in the initial version of the scale (a total of 30 items).

[b] Not modified from Grove, Bourgonjon and Van Looy's (2012) scale.
[c] Scored in a reverse way.

3.3 Inter-correlation matrix of three factors

Table 3 further presented the inter-correlation matrix among three subscales. Because of the correlations reach significant level of 0.001, the three factors measure in a coherent way.

Table 3: Inter-correlation matrix matrix of three factors.

Three factors	Confidence	Belief	Motivation
Confidence	–		
Belief	0.390***		
Motivation	0.497***	0.599***	

$*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

4. Discussion

This study developed a scale regarding teachers' belief, confidence and motivation of digital game-based learning pedagogies. Over past few years, digital games have played an important role in education. There were some of scales have developed to assessing teacher' adoption intention and acceptance of digital games. Compare with these instruments in the following.

A previous framework was proposed by Frederik, Jeroen and Jan, to assess which factors influence the individual adoption intention of a teacher toward game-based learning (Grove et al., 2012). The instrument with the following six subscales: ease of use, usefulness, experience, behavioral intention and learning opportunities, constructing on the basis of previous research using the TAM (Bourgonjon et al., 2012). The study focued on measuring teacher' adoption intention toward adopt digital games.

CTPCK was developed Hsu, Su and Liang to examine the effects of the technology- and pedagogy-oriented course design on improving the in-service preschool teachers' Technological Pedagogical Content Knowledge- Games (TPACK-G). They assessed teachers' acceptance of digital game-based learning. Based on integrate ICT into classroom teaching and learning: which type of knowledge (e.g., TK, CK, or PK) should be instructed first during the course or not (Hsu et al., 2012).

This study was developed the scale about exploring the teachers' beliefs of teaching toward digital game-based learning. Because of beliefs might influenced teachers' teaching method. Hence, the study was developed the scale with the following three subscales: confidence, belief and motivation. Firstly, assess teachers' experience with design a digital game-based curriculum. Then, measuring whether they feel that adopting digital games into classrooms were usefulness or not. Finally, measuring teachers' confidence and they will adopt digital games in the future or not. Through this instrument to realize teachers' opinion toward integrate digital games in teaching.

Hope can use this instrument through from inside mental to outside behavior's way to realize teachers' opinion toward integrate digital games in teaching; and further, to improve and promote the implement of digital game-based learning.

References

Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Computers & Education*, *54*(4), 1145-1156.

Cai, Y., Snel, I., Bharathi, B.S., Klein, C. & Klein-Seetharaman, J. (2003). Towards biomedical problem solving in a game environment. *Lecture Notes of Computer Science*, *2659*, 1005-1014.

Divjak, B., & Tomić, D. (2011). The impact of game-based learning on the achievement of learning goals and motivation for learning mathematics-literature review. *Journal of Information and Organizational Sciences*, *35*(1), 15-30.

Grove, De., Bourgonjon, J., & Van Looy, J. (2012). Digital games in the classroom? Acontextual approach to teachers' adoption intention of digital games in formaleducation. Computers in Human Behaviour, 28, 2023–2033.

Hsu, C. Y., Su, Y. C., & Liang, J. C. (2012). Examining the effects of integrating technological pedagogical content knowledge into preschool teachers' professional development regarding science teaching: using digital game-based learning as an example. ICCE, Bali, Indonesia.

Hwang, G. J., & Wu, P. H. (2012). Advancements and trends in digital game-based learning research: A review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*, *43*(l), E6-10.

Oblinger, D. (2004). The Next Generation of Educational Engagement. *Journal of Interactive Media in Education*. From : http://jime.open.ac.uk/jime/article/viewArticle/2004-8-oblinger/198

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: impact on educational effectiveness and student motivation. *Computers & Education*, *52*, 1–12.

Prensky, M. (2001). Digital game-based learning. New York: McGraw-Hill.

Ruggiero, D. (2013). Video games in the classroom: The teacher point of view. The Foundations of Digital Games conference, Chania, Greece.

Eye-tracking analyses of text-and-graphic design effects on E-book reading process and performance: "Spanish color vocabulary " as an example

Tse-Wen PAN[a*] & Meng-Jung TSAI[a]
[a]*Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan, ROC*
*M10111008@mail.ntust.edu.tw

Abstract: This study aimed to explore how graphical design effects on learner's E-book control and visual behaviors by an eye-tracking experiment. The participants were twelve university students with novice experience of Spanish who were asked to read and learn basic Spanish vocabularies in E-book display on iPad. Participants were randomly assigned into two groups of reading materials with high-related graphics and low-related graphics. An ASL MobileEye eye tracker was used to track and record the gaze data of learners. Each participant's action of controls and visual attention during the experiment were observed and coded. Wilcoxon test, Mann–Whitney U tests and Pearson's correlation analyses were used to analyze the data. The results of this study suggest that low-related graphical design may reduce the effectiveness of learning. Future study is suggested to explore the effects of E-book cue design on students' visual attention and learning performance.

Keywords: E-book; eye tracking; visual attention; multimedia, reading

1. Introduction

1.1 E-books

Recently, since tablet computers are equipped with sundry features, such as touch screen, wireless networking capabilities and the storage of memory (Lim, Lee, and Ahn, 2013), it becomes popular in education field; therefore there has been much research about how it affect student's learning performance. E-book is a popular format to represent multimedia as the learning material; it includes text, text-speech, music, sound and animation (Korat, and Shamir, 2004). E-book also includes some features, such as hidden button, quick view, keyword searching and bookmark (Vassiliou, and Rowley, 2008). In terms of the content, E-book on mobile device is more flexible and interesting than the content on the printed material (Woody, Daniel, and Baker, 2010). Previous study compared the effect of read text in print and read on screen, the result showed that students who read E-books compared to printed texts had higher affective learning and psychomotor learning (Rockinson-Szapkiw, Courduff, Carter, and Bennett, 2013). However, another result of study found that students didn't enjoy in E-book reading and had very low scores on comprehension tests (Lam, Lam, Lam, and McNaught, 2009). Those inconsistent findings suggests researcher to explore deeply in this field.

1.2 Visual attention and reading behavior

Eye-fixation is the instinctive reaction of stimuli, therefore, visual attention analysis result may be the stronger evidence of mental processes of reading and other visual activity of learner (Rayner, 1998).

Visual attention is based on eye movement, and can be trace back to "eye-mind" assumption (Just, and Carpenter, 1976) which suggests that tracking one's eye movements could help people to realize his visual attention. Recently, eye tracking technology was started applying on cognitive process of

learning research to disclose how human process information while learning. Previous researchers divided different eye movement measures into three categories (Lai, et al.,2013):
•Temporal: to answer the “when” and “how long” questions related to cognitive processing.
•Spatial: to answer the “where” and “how” questions related to cognitive processing.
•Count: to show the importance of the visual content.

1.3 Cognitive theory of multimedia

According to dual-coding theory of Paivio, compare with single media, previous researcher found that multimedia may help learner effectively and enhance learning performance. Besides, in the past findings, researcher found that non-related material may distracts learner’s attention and waste learner’s time on non-related graphic, reducing information process in working memory and increase cognitive overload (Mayer, 2009).

2. Purpose

The aim of this study is to exam the effects of multimedia design on students’ visual attention; researcher used an eye tracking examination in this study. Specifically, this study disclosed how two different design of graphic in the E-book (i.e., high-related graphic of the Spanish vocabulary and low-related graphic of the vocabulary) effect learning performance (i.e., posttests for Spanish vocabulary) and visual attention (i.e., number of times zone observed, duration before first fixation arrival, percent time fixated related to total fixation, pupil size, total time in zone, fixation count, total fixation duration). This study further explored the relation between learners’ background and reading behavior (i.e. total click number of pronunciation button). The research questions included in following:
RQ1: What is the rule of eye movement of different groups’ learner’s between text and graphic while E-book reading?
RQ2: Is there any significantly difference in students' E-book reading behaviors between high-related and low-related graphic groups?
RQ3: Is there any significant relationship between learners’ visual attention and posttests?

3. Methods

3.1 Participants

Twelve university students form north Taiwan were novice Spanish learner were randomly assigned into high-related graphic group or low-related graphic group.

3.2 Learning material

Two versions E-book of Spanish vocabulary presented in iPad was used in this study, both version include ten pages with ten color vocabulary in each page. The high-related material as shown in Figure1, since apple is red, the graphic is high-related to the vocabulary “rojo (red)”. On the other hand, in the

Figure1. Material for the high-related group.

Figure2. Material for the low-related group.

low-related graphic material (Figure2), the graphic of "rojo (red)" is a man doing gymnastics which is low-related to the vocabulary.

3.3 Eye-Tracking system

In this study, ASL Mobile Eye-XG with a sampling rate of 30 Hz was used to track participants' eye-movements through the experiment and Participants was free to move their head. During the experiment, participants' E-book control and eye movement was recorded separately as the same time.

3.4 Background Questionnaire

To realize participants' language learning experience, preferences, attitude and personal background, the background questionnaire was assessed before the reading task.

3.5 Posttest

To observe the effect of reading task, an immediately posttest (i.e., posttest1), a three day delayed posttest (i.e., posttest2) and a one week delayed posttest (i.e., posttest3) was assessed after the reading task. The better scores the participants had indicated the more learning content they remembered.

3.6 Reading behavior Analyses

According to prior literature (Lai, et al., 2013), E-book control behavior and several eye movement measures were observed, which are defined as follows:

- Total clicking number: the number of time that participant clicked the hidden bottom to listen the pronunciation of the vocabulary.
- Number of times zone observed: sum of visited fixations count and saccade count.
- Duration before first fixation arrival (seconds): time spent on the first fixation.
- Percent time fixated related to total fixation: total time spent on fixations in a look zone related to all look zone.
- Average pupil size (pixel2): the average size of pupil while looking in specific look zone.
- Total time in zone (seconds): sum of fixations saccade time in a look zone.
- Fixation count: total number of fixation counted in a look zone.
- Total fixation duration (seconds): total time spent on fixations

3.7 Statistical

Mann–Whitney U test was used to analysis the difference in students' E-book reading behaviors between high-related and low-related graphic groups. Besides, correlation analyses was used to analyses the relationship between learners' backgrounds and reading behaviors, as well as the relationship between learners' visual attention and posttests.

3.8 Statistical

In this study, Mann–Whitney U test was used to analysis the differences of three posttests and reading behavior between two groups. Besides, Wilcoxon test was used to analysis the retention of each posttest. Third, Pearson's correlation was use to analysis participants' background and reading behavior.

4. Results

4.1 Resuls of Wilcoxon test on text and graphic

Though the Wilcoxon test, the findings showed that all eye movement measures of text and graphic were significant different with each other in low-related graphic group. However, in high-related graphic group, there was no significant different between text DBFFA (duration before first fixation arrival) and graphic DBFFA ($p = 0.24$). Meanwhile, in the high-related graphic group, the pupil size of text had no significant different to the pupil size of graphic ($p = 0.46$).

4.2 Results of Mann–Whitney U test result on visual attention

Though the Mann-Whitney U test, we found that low-related graphic group (mean = 13.16) had significant higher number of times zone observed ($p = 0.02$) compared to the high-related graphic group (mean = 7.10). It meant learner in low-related graphic group spent significant more time of observing the graphic zone than the high-related group.

4.3 Correlation between posttests and eye movement measures

The results of correlation between posttests and reading behavior showed that in high-related graphic group, the average pupil size was significantly negative related to the posttest1 ($r = -.832$, $p = 0.040$) and the posttest2 ($r = -.882$, $p = 0.020$).

In the low-related graphic group, the total fixation duration was significantly negative related to posttest1 ($r = -0.820$, $p = 0.045$), posttest2 ($r = -0.895$, $p = 0.016$) and posttest3 ($r = -0.817$, $p = 0.047$). The percent time fixated related to total fixation was significantly negative to posttest1 ($r = -0.882$, $p = 0.02$) and posttest2 ($r = -0.852$, $p = 0.31$). Total time in zone was significant negative to posttest2 ($r = -0.860$, $p = 0.028$). Fixation count was significantly negative to posttest2 ($r = -0.867$, $p = 0.025$).

These result may indicated that low-related graphic was unhelpful to the performance of learner.

5. Discussion and conclusion

This study was aimed to investigate the effect of graphic design on E-book reading behavior and learning performance. First, we found that learners who were in low-related graphic group had significant longer DBFFA on text zone than graphic zone, which may indicate learner who read low-related graphic thought the graphic is not helpful for them to memorize the content. Secondly, low-related graphic group spent more time observing the graphic zone, which is inconsistent with previous findings (Tsai, Hou, Lai, Liu, and Yang, 2012) (Balslev, et al., 2012) that learners fixated relevant information longer within text-based environment. Third, in the high-related graphic group, the smaller the students' pupil size is, the higher scores the students had. According to previous study (Seeber, and Kerzel, 2011), this result may indicated that when learners read high-related graphic, the cognitive load may be lower and could result in better learning efficiency. On the contrary, in the low-related graphic group, the more effort they spent on the graphic zone, the lower learning outcome they had. It may be indicate that low-related graphics created cognitive load to learners.

In sum, according to these findings and previous study (Mayer, 2009), relevant graphic and text is suggested to arrange in pairs while design vocabulary learning material. In other words, vocabulary learning material designers should be careful to selecting appropriate graphics.

Future study was suggested to explore the relation between mental effort and pupil size of learner through interview method. In addition, more complex cognitive test was suggested to assess learner's learning performance. Moreover, future study is suggested to explore the effects of E-book cue design on students' visual attention and learning performance

References

Balslev, T., Jarodzka, H., Holmqvist, K., de Grave, W., Muijtjens, A. M., Eika, B., ... and Scherpbier, A. J. (2012). Visual expertise in paediatric neurology. *European journal of paediatric neurology*, *16*(2), 161-166.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4),

441-480.
Korat, O., & Shamir, A. (2004). Do Hebrew electronic books differ from Dutch electronic books? A replication of a Dutch content analysis. *Journal of Computer Assisted Learning*, *20*(4), 257-268.
Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., ... & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90-115.
Lam, P., Lam, S. L., Lam, J., & McNaught, C. (2009). Usability and usefulness of eBooks on PPCs: How students' opinions vary over time. *Australasian journal of educational technology*, *25*(1), 30-44.
Lim, S. H., Lee, S., & Ahn, W. H. (2013). Applications IO profiling and analysis for smart devices. *Journal of Systems Architecture*, *59*(9), 740-747.
Mayer, R. E. (2009). *Multimedia learning*. NY: Cambridge University.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.
Rockinson-Szapkiw, A. J., Courduff, J., Carter, K., & Bennett, D. (2013). Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning. *Computers & Education*, *63*, 259-266.
Seeber, K. G., & Kerzel, D. (2011). Cognitive load in simultaneous interpreting: model meets data. *International Journal of Bilingualism*, 1367006911402982.
Tsai, M. J., Hou, H. T., Lai, M. L., Liu, W. Y., & Yang, F. Y. (2012). Visual attention for solving multiple-choice science problem: An eye-tracking analysis. *Computers & Education*, *58*(1), 375-385.
Vassiliou, M., & Rowley, J. (2008). Progressing the definition of "e-book". *Library Hi Tech*, *26*(3), 355-368.
Woody, W. D., Daniel, D. B., & Baker, C. A. (2010). E-books or textbooks: Students prefer textbooks. *Computers & Education*, *55*(3), 945-948.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Weblog as Learning Community for Supporting Astronomy Teaching in Thailand

Sakanan ANANTASOOK and Chokchai YUENYONG*

Science Education Program, Faculty of Education, Khon Kaen University, Thailand
* ychok@kku.ac.th

Abstract: Previous studies indicated that there were some difficulties of astronomy teaching in Thailand because of teachers' lacking of content and pedagogical knowledge. Building up some spaces for Thai astronomy teacher learning community may allow them to gain their knowledge and pedagogy about astronomy. A weblog (blog) as a web-based technology may allow Thai astronomy teachers' community to improve pedagogical knowledge and astronomy concepts through expressing and exchanging ideas interaction and collaboration, social networking and group work. This paper share ideas of provide weblog for learning community of Thai astronomy teachers. The knowledge of sharing in weblog may have implications for professional development of Thai astronomy teachers.

Keywords: Weblog, learning community, astronomy, pedagogy

1. Introduction

Astronomy is one of the key areas for citizens in the 21st century. It can play a unique role in facilitating education and capacity building of citizens in society; and in furthering sustainable development throughout the world (Percy, 1996; Plummer, 2006). Knowledge of astronomy can increase public awareness, understanding and appreciation of science and technology which are important in all countries, both developed and developing (Percy, 2005). Astronomy is a science that transcends cultures, has been prominent in the news in recent years, and can generate excitement in young minds as no other science can. Astronomy is useful for understanding other science and mathematics concepts, and for developing problem solving skills, which are important in our technological world (Lebofsky, Canizo, & Lebofsky, 1996).

Teaching and learning astronomy in schools is highly justifiable. It not only prepare individuals for becoming astronomers in the future but also prepares citizens as a new generation who have deeper understanding of science, the ability to think critically and be able to explain astronomical phenomena that occur in everyday life and are relevant to their culture and society. However, it seems that there are some difficulties of astronomy teaching in Thailand. Astronomy is just obviously appearing in both the 2001 and later the 2008 Thailand school science curriculum (IPST, 2008). Not many Thai school science teachers graduated in astronomy area. Some previous studies found that the cause of Thai students' misconception about astronomy was teachers' low capability in teaching astronomy. Teachers could not provide constructivist learning models or instructional media for astronomy (Khongpugdee, Sukonthachat, & Phonphok; 2009; Khongpugdee, 2010; Dahsah et al., 2012).

Building up some spaces for Thai astronomy teacher learning community may allow them to gain their knowledge and pedagogy about astronomy. Teachers are embracing technology as a way to increase instructional effectiveness and reach the 21st century learner. The Internet is currently applied in a wide range of e-learning settings, and various innovative web-based learning systems have been developed over the last few years. Numerous information technologies have emerged to support this type of educational environment, by facilitating communication and collaboration among online learners. Blogs, wikis, Facebook, and microblogs are all technologies in educational contexts in many studies (Cole, 2009; Huang, 2011; Robertson, 2011).

A weblog (blog) is a web-based technology that allows people to quickly share their thoughts and comments with the entire web population. Persons who are not necessarily familiar with web design codes (HTML, CSS) are able to successfully post an article with multimedia materials. This advantage has increased recent attention to pedagogical roles of blogs in the e-learning literature. Compared with other popular social software applications (wikis, online forums, facebook, and microblogs), blogs can be more applied more broadly, allowing simplified web pages, links, and resource collections (Huang, 2011). There were many studies using the blog for diverse learning groups, ranging from primary (Davis 2006, Tse et al. 2010) and secondary education (Angelaina & Jimoyiannis 2009) to higher education (Blau et al. 2009, Kerawalla et al. 2009, Tan et al. 2010) and teachers' professional development as well (Makri & Kinigos 2007).

Since 2010, The Technology for Teaching and Learning Center, under the Thai Ministry of Education every years training Thai teachers to create blog content through "wordpress" free blog service (URL: http://www.wordpress.com). Many teachers who attend in the workshop created blogs and used as effective strategies for teaching and learning in their class. Some studies show that students more satisfied and they was higher achievement after learning by using a weblog that applying with social media (Facebook, Google Docs, Slideshare, and Youtube) (Sudprakone, 2012; Anantasook, 2014). In addition, other teachers can use content from various educational blogs which some teachers designed, in their classroom. Because of their organizational and pedagogical features, blog for Thai astronomy teachers' community can offer enhanced opportunities to teachers, not only to improve pedagogical skills, but to construct new knowledge through expressing and exchanging ideas, critical and reflective thinking, interaction and collaboration, social networking and group work.

2. Weblog as Learning Community for Thai astronomy teachers

The weblog which concern in this study is astronomy education's blog [www.astroeducation.com], was started blogging in January 2013. The objective's blog setting was provided for teaching and learning on astronomy, astronomy education, in the Thai's basic educational levels (Grade 1-12). This blog consist of seven main menus; (1) author's astronomical activities, (2) news and announce, (3) astronomy lesson (4) astronomy knowledge, (5) learning medias, (6) astronomy curriculum, and (7) CERN physics teacher projects. These all menus and some articles are show in figure 1.

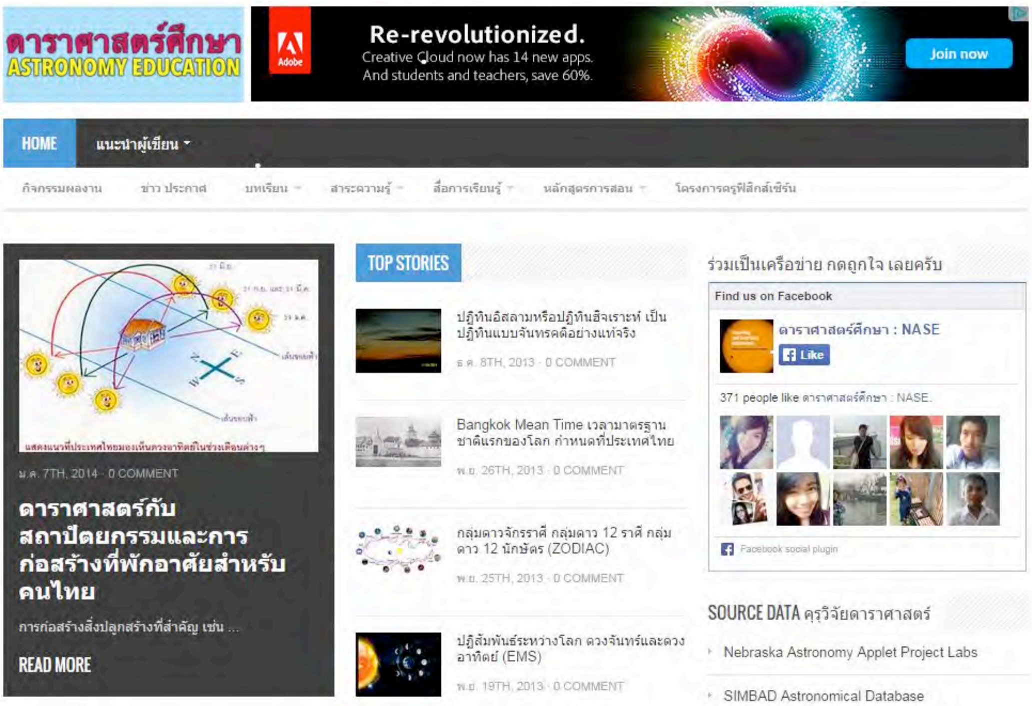


figure 1. The first page of astronomy education's blog [www.astroeducation.com].

The first author's astronomical activities menu includes the author's news or activities when he attended in the astronomy workshop and conference. For example, the author attend in the Eratosthenes workshop; measure circumference of the Earth on 25-27 March 2013, and the Institute for the Promotion of Teaching Science and Technology (IPST)'s workshop for astronomy teacher on 11-15 March 2013. The author present his research study in the title "The Poster for Archaeoastronomy and Geology of Prasats in Surin Province: the Learning Material" in the International Conference of Educational Research (ICER) 2011 on 9 November 2011.

The second news and announce menu that include the news of astronomy workshop or projects which many organizations such as the Institute for the Promotion of Teaching Science and Technology (IPST), the Thai Astronomy Society, The National Astronomical Research Institute of Thailand (NARIT) and The Learning centre for Earth Science and Astronomy (LESA), create activities for teachers who teaching astronomy.

The third astronomy lesson menu consists of two archaeoastronomy lessons. The first lesson was designed for enhancing learners' understanding on how the direction of 30 Surin province prasats' (Hindu temple) structure was provided in order to face to the East at the position of sunrise on Equinox day. The lesson provided the picture and detail of all prasats in term of age, purpose, material and direction of construction which were checked and survey by the author before created the lesson. The learning requires students to take action according to the guidelines for two hours in weblog class and one hour outside of the class. They were enthusiastic and interested in the group activity and knew more about Hindu temple that relevant to the movement of the Sun. The second lesson is the operation of measure circumference of the Earth on Vernal Equinox day. The author used the Eratosthenes method, and a difference technique for measure circumference of the Earth. Because of the Sun move on the Equator in the Vernal Equinox day. The author applied the Google Earth free online program for measured the distance from the participants' location to the Equator. In this part, the author created many articles while he stayed in New Zealand. Thai and Lao teachers and students (N=67) in difference locations, who submitted in this operation, learn the content from this part and successfully to measure circumference of the Earth on 23-24 March 2014.

The fourth astronomy knowledge menu that include several articles which wrote and post on the blog by the author. It can categories in two sub-menus; (1) the concept of time and calendar, (2) the concept of astronomy in daily life. This part benefits for person who interests on astronomy in Thai societal and cultural context.

The fifth learning media menu that involve astronomy media such as video clip, astronomy textbook, and free astronomical program that were shared on the internet. The author selected and categorized them which benefit for students and teachers in the basic educational level. This main menu includes three sub-menus; (1) video clips content for students, (2) video clips training for astronomy teachers, (3) astronomy textbooks and astronomical program.

The sixth astronomy curriculum menu that involve three sub-menu which link to the astronomy level in Thai science curriculum that include primary level, lower secondary level and upper secondary level. However, only astronomy lessons and learning material on the upper secondary level were stipulated in the first year. This menu can support both teachers and students for teaching and learning astronomy in their classroom.

The final main menu is CERN physics teacher projects menu. This part include particle physics knowledge from the European Organization for Nuclear Research (CERN)-Physics High School Teacher Programme which the author interest individually. This menu benefits for physics teachers who interests in particle physics laboratory and planning to attend in the CERN-Physics High School Teacher Programme.

3. Community Building of Thai Astronomy Teacher Weblog: First Year

The first year, it is good start for community building of Thai astronomy teacher weblog. Number of teacher access to the blog and teachers who become member in astronomy educations' fanpage [https://www.facebook.com/astroeducation] are increasing in everyday. The blogger normally post many articles in each menu but focus on the sixth menu. The author provides the astronomy learning material such as learning activities, learning packets, and lesson plans for teachers who teach astronomy content at the upper secondary level [http://www.astroeducation.com/category/curriculum/highlevel/]. All documents contain necessary information for teacher looking to start using it in their astronomy classroom. The blog is currently not many articles then it was visited around 200-250 visitors per day.

The community seems to consider in pedagogical knowledge. They usually come to the sixth menu, astronomy curriculum menu, through google search for using learning material for teaching. Then, they download all files by themselves or e-mail to the astronomy education blogger for send files them. Their communication was related to need of pedagogical knowledge as follow:

> *"I don't have any learning material, could you please send it to me."* Ketsaraporn [02/07/2013]
> *"Now, I don't have lesson plan, the plan for teaching astronomy, could you please send it to me. Thank you very much"* Pensriri kaewnongsong [22/11/2013]
> *"I am biology teacher, but I was assigned to teach astronomy in this semester, I Requesting assistance the astronomy learning plan from you."* Phacharanapat Boonseela [07/01/2014]
> *"This is the first time for me to teach astronomy, could you please send lesson plan and files of astronomy content to me."* Sririyakorn Jullee [19/04/2014], Rewadee chansamut [06/05/2014]
> *I" am a new teacher, I am lake the teaching technique. Could I receive your files."* Panida [30/04/2014]

It indicated that astronomy teaching had difficulty for them. They asked for learning materials or even lesson plan because most of them were first time to teach astronomy. Even some of them are not new science teachers, they are new astronomy teachers. This similar to previous studies found the problem of astronomy teaching in Thailand (Khongpugdee, Sukonthachat, & Phonphok; 2009; Khongpugdee, 2010; Dahsah et al., 2012). It could be mentioned that the contents in this astronomy education blog can support teachers for teaching and learning astronomy in Thailand.

4. Conclusion

The astronomy education blog could build Thai astronomy teachers' community up. First year of sharing in the blog indicated that teachers required not only pedagogical knowledge but also astronomy concepts. It is probably because most of astronomy teachers have no background of astronomy areas. The weblog may increase of providing media to enhance constructing astronomy concept. More interaction between experts and naïve astronomy teachers should be provided in order to set up the atmosphere of professional development through the blog.

References

Anantasook, R. (2014). *The effect of Weblog instruction applying social media entitle "Basic C Program language" upon achievement learning and satisfaction of Grade 9 students.* Classroom research. Surin : Rattanaburi school.

Angelaina, S. & Jimoyiannis, A. (2009). The educational blog as a tool for social construction of knowledge: Analysis of students' cognitive presence. *Proceedings of the 6th Panhellenic Conference "Teaching Sciences and New Technologies in Education"* (eds P. Kariotoglou, A. Spyrtou & A. Zoupidis), pp. 137-145. Florina (in Greek).

Blau, I., Mor, N., & Neuthal, T. (2009). Open the windows of communication: Promoting interpersonal and group interactions using blogs in higher education, *Interdisciplinary Journal of E-Leaning and Learning Objects*, 5, 233-246.

Cole, M. (2009). "Using Wiki Technology to Support Student Engagement: Lessons from the Trenches," *Computers & Education*, 52, 141-146.

Dahsah, C., Phonphok, N., Pruekpramool, C., Sangpradit, T. & Sukonthachat, J. (2012). Students' Conception on Sizes and Distances of the Earth-Moon-Sun Models. *European Journal of Social Sciences*, *32* (4): 583-597.

Huang, T. C. (2011). Creating a Knowledge Development Model for Blog-Based Learning, *International Journal of Information and Education Technology*, *1*(3), 261 – 267.

Institute for the Promotion of Teaching Science and Technology [IPST]. (2008). *The Basic Education Core Curriculum B.E. 2551 (A.D. 2008) Science*. Bangkok : Kurusapha Ladpraw.

Kerawalla, L., Minocha, S., Kirkup & Conole, G. (2009). An empirically grounded framework to guide blogging in higher education, *Journal of Computer Assisted Learning*, 25, 31-42.

Khongpugdee, S. (2010). *Effect of teaching astronomy based on inquiry method using Earth Moon Sun system model (EMS-Model) innovation*. Dissertation, Ph.D (Science Education). Bangkok: Graduate School, Srinakharinwirot University.

Khongpugdee, S., Sukonthachat, J. & Phonphok, N. (2009). A study of conceptual understanding in basic astronomy of grade 9 and 10 students in rural areas of Thailand, *Thai Journal of Physics*, 4: 124-126.

Lebofsky, L. A., Canizo, T. A., & Lebofsky R. N. (1996). "Project Artist and Project Access! Integrating astronomy and planetary sciences into the elementary and middle school curriculum." In J. R. Percy. (ed.). *Astronomy Education: Current Developments, Future Coordination*. San Francisco: BookCrafters, Inc., 235-236.

Makri, K. & Kynigos, C. (2007). The role of blogs in studying the discourse and social practices of mathematics teachers, *Educational Technology & Society*, *10*(1), 73-84.

Namwar, Y., & Rastgoo, A. (2008). Weblog as a learning tool in higher education. *Turkish Online Journal of Distance Education*, *9*(3). Retrieved August 8, 2014, from http://tojde.anadolu.edu.tr/tojde31/articles/article_15.htm

Percy, J. R. (1996). "Astronomy education: An international perspective." In J. R. Percy. (ed.). *Astronomy Education: Current Developments, Future Coordination*. San Francisco: Book Crafters, Inc., 1-8.

Percy, J. R. (2005). "Why astronomy is useful and should be included in the science curriculum." In J.M. Pasachoff, and J.R. Percy. (eds.) *Teaching and Learning Astronomy: Effective Strategies for Educators Worldwide*. New York: Cambridge University Press, 10-13.

Plummer, J. (2006). *Students' Development of Astronomy Concepts Across Time*. Doctoral Dissertation, Ph.D (Astronomy and Education). USA: The University of Michigan.

Robertson, J. (2011). "The educational affordances of blogs for self-directed learning," *Computers & Education*, *57*(2): 1628-1644.

Sudprakone, S. (2012). *The effect of Web-base instruction applying social media entitle "Basic Computer" upon achievement learning of mattayousuksa 1*. Master of Education Thesis, Graduate School, Khon Kean University.

Tan S. M., Ladyshewsky, R. K. & Gardner, P. (2010). Using blogging to promote clinical reasoning and metacognition in undergraduate physiotherapy fieldwork programs. *Australasian Journal of Educational Technology*, *26*(3), 355-368.

Tse, S. K., Yuen, A. H. K., Loh, E. K. Y., Lam, J. W. I. & Ng, R. H. W. (2010). The impact of blogging on Hong Kong primary school students' bilingual reading literacy, *Australasian Journal of Educational Technology*, *26*(2), 164-179.

The Perceived and Expected User experiences of AR Book Reading: the Perspective of Parents

Kun-Hung CHENG[a*] & Chin-Chung TSAI[b]
[a]*Department of Communication and Technology, National Chiao Tung University, Taiwan*
[b]*Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan*
*kuhu@mail.nctu.edu.tw

Abstract: Studies regarding augmented reality (AR) books generally indicated its advantages on reading performance. Positive attitude toward the usage of AR books and acceptance of usability with AR books was also responded by users including samples of students and parents. AR books create a new experience of reading; and the further exploration of user experience beyond usability when using AR applications can be considered. The present study hence invited 29 pairs of parents and their children coming from different regions in Taiwan to join a shared AR book reading activity. Each parent was interviewed for understanding their perceived user experiences and expected user experiences of the selected AR book in this study; and the qualitative data was analyzed through phenomenographic method. The results showed that the parents perceived using the AR book in both emotional senses (e.g., joy and surprise) and cognitive senses (e.g., increasing attention). However, some of them perceived negative experiences such as interrupting traditional book reading by AR. Also, AR books are expected to provide parents with rich interaction and vivid animation for experiencing the augmented information. The diverse user experiences of the AR book reading perceived by the parents could be elaborated to further explore their relationships with intention to use and usage behaviors. The characteristics of perceived and expected user experiences responded by the parents found in this study can also be the basis for the practical design of AR books in the future.

Keywords: augmented reality, book reading, user experience, parents, phenomenographic method

1. Introduction

An augmented reality (AR) book, combining conventional paper book and real-time synthetic information for its simultaneous presentation through a screen device, has been considered its advantages on learning recently (Cheng & Tsai, 2014; Ibáñez et al., 2014). In addition to the experiences of paper book reading, AR books offer more virtual information such as 3D objects, video, or audio relevant to the book pages for users to read and thus originate a new reading experience. For the effects of AR books on reading performance, past studies have indicated their benefits (Abas & Zaman, 2011; Vate-U-Lan, 2012). AR book systems could enhance students' motivation to learn (Di Serio et al., 2013). Acceptance of usability (e.g., effectiveness, efficiency, satisfactory) with AR books was also responded by students (Chang et al., 2011). In terms of parents' attitude toward their children's use of AR technology in preschool education, they were inclined to agree with that AR could motivate their children to learn (Cascales et al., 2013). However, these findings only initially investigated the usability issues regarding users' attitude toward AR. Since AR books involve a new experience of reading, the further exploration of user experience beyond usability when using AR applications can be considered (Cheng & Tsai, 2013).

According to ISO, the definition of user experience is "a person's perceptions and responses that result from the use or anticipated use of a product, system or service." The definition may elicit the ideas that while a user's perceptions of "use" of a system stands for the concept of *perceived user experience*, a user's perceptions of "anticipated use" of a system indicates the concept of *expected user experience*. Moreover, researchers mostly agree with the notion that user experience moves beyond usability toward more emotional concern about the interaction between users and products or systems (Desmet et al., 2007). Nevertheless, user experience involves a consequence of a user's internal state (e.g., motivation or mood), the characteristics of the designed system (e.g., usability or functionality), and the context within which the interaction occurs (Hassenzahl & Tractinsky, 2006). The findings imply that user experience is a broader scope including usability issues that designers are suggested to concern.

With regard to the research on the user experience of AR applications, recently, a study interviewed 28 participants in a shopping mall for exploring the expected user experience of mobile AR services for business purposes (Olsson et al., 2013). They found several characteristics of user experience (e.g., feelings of being connected with other people using the services) and requirements (e.g., requirements of privacy protection in interacting with the services) that affect and facilitate the mobile AR experiences. However, related research on user experience studies with AR applications is still limited, particularly on reading AR books along with parents' perspectives. Based on the two facets of *perceived user experience* and *expected user experience* discussed previously, this study therefore explored parents' experiences when reading an AR book with their children and qualitatively categorized the characteristics of user experience. Specifically, the research questions are as follows:

(1) What is the parents' *perceived user experiences* of the AR book reading?
(2) What is the parents' *expected user experiences* of the AR book reading?

2. Method

The participants in this study included 29 pairs of parents and their children coming from different regions in Taiwan (i.e., 10 pairs from northern region 9 pairs from central region, and 10 pairs from southern region). While the ages of the children are between 4 and 9 (mean=6.48, SD=1.21), the parents are between 24 and 60 years old (mean=37.72, SD=7.79). Most of the parents were familiar with using smartphones or tablet PCs; however, they generally did not experience the demonstration of AR.

An AR picture book with artistic introduction, namely "*The adventures of Yuyu: Yuyu Yang artistic journey (published by National Chiao Tung University Press in Taiwan),*" was utilized as the learning material in this study. In a pair setting, the parents and their children were required to freely share reading with an iPad. A trained research assistant introduced the reading process and the usage of the book before the activity began. When the participants finished the AR book reading, all the parents were interviewed for understanding their *perceived user experiences* and *expected user experiences* of AR book. The guiding interview questions include:

(1) Please describe the feelings when you read the AR book.
(2) What do you experience in the process of the AR book reading?
(3) What do you expect to experience when reading an AR book?
(4) What content of an AR book and the format it present do you expect to see?

The interviews were undertaken in Chinese and audio-recorded, as well as fully transcribed. To reveal the user experiences of the parents in the learning activity, this study conducted a phenomenographic method, which was used to analyze and further categorize students' conceptions of learning in previous studies (e.g., Tsai et al., 2011; Lee et al., 2013), to examine the verbatim transcripts of the parent interviews. Specifically, for each parent's interview transcripts, the researcher firstly marked the most important sentences that could represent their main idea of experiencing the AR book reading. The content-specific consistencies and differences across the interviewed parents' responses were then explored and summarized by comparing the selected sentences. Following previous steps, the qualitatively different categories of the user experiences of AR book reading perceived and expected by the parents can be constructed.

3. Results and discussion

3.1. Perceived user experiences

With an initial attempt to probe parents' user experiences of an AR book reading, the present study conducted phenomenographic analysis and preliminarily found that the parents perceived both positive and negative user experiences. In terms of the positive user experiences, some parents can feel the reality of objects or environments via the AR technology; and they emotionally expressed their joy, amusement and playfulness that arise from a brand-new experience of reading. Feelings about the traditional book reading being surpassed by the experiences of AR book reading, particularly for the 3D elements or virtual information overlapping from the paper book, also surprised them. In addition to these emotional senses, they perceived the advantages of AR technology on reading in a cognitive sense, for example enhancing awareness, memory, and knowledge about the content of the book.

However, some of them perceived negative experiences that reading book with AR technology likely interrupted traditional book reading, as well as operating the AR book alone with problems of usability. Some parents felt an antipathy toward technology use and tended to forbid their children to use electronic devices. In the consideration of this study, the diverse user experiences of the AR book reading perceived by the parents could be elaborated to further explore their relationships with intention to use and usage behaviors.

3.2. Expected user experiences

Regarding the parents' expectations of reading AR books, the present study found that they expected AR books to include more interactive design for users to interact with the virtual elements or information (e.g., tap the screen to control the augmented 3D objects). Also, they expected AR books to include more vivid animation regarding the book story to fulfill the content of the virtual information within the AR book. In other words, AR books are anticipated to provide parents with rich interaction and vivid animation for experiencing the augmented information with their children.

With regard to the help of learning, some parents expected more suitable content of AR books to draw on their children's interests or needs (e.g., interesting topics, life-relevant materials, or abstract knowledge).Also, AR books with value-added materials integrated into the paper books (e.g., a craft of cutting paper for increasing the children's impression of the artistic work) for their children were anticipated by the parents to be developed. Similar to the perceived user experiences found in this study, the parents also expected AR books to foster learning in a cognitive sense. Finally, some of them anticipated to read AR books with more ease of control, and even imagined to experience actual feelings of touching objects. An AR glass (e.g., Google glass) in the developing process may be up to the expectations of portability.

4. Conclusion

In sum, the characteristics of perceived and expected user experiences responded by the parents found in this study can be the basis for the practical design of AR books in the future. With more considerations in parents' user experience, the acceptance of adopting AR books for their children to learn may be increased.

References

Abas, H., & Zaman, H. B. (2011). Visual learning through augmented reality storybook for remedial student. *Proceedings of the Second international conference on Visual informatics: sustaining research and innovations - Volume Part II*, 157-167.

Cascales, A., Pérez-López, D., & Contero, M. (2013). Studies on parent's acceptance of the augmented reality use for preschool education. *Proceedings of 2013 International Conference on Virtual and Augmented Reality in Education*, 420-427.

Chang, Y. J., Chen, C. H., Huang, W. T., & Huang, W. S. (2011). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of English learning using augmented reality. *In: Proceedings of IEEE International Conference on Multimedia and Expo*.

Cheng, K. H & Tsai, C. C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology, 22*(4), 449-462.

Cheng, K. H., & Tsai, C. C. (2014). Children and parents' reading of an augmented reality picture book: Analyses of behavioral patterns and cognitive attainment. *Computers & Education, 72,* 302-312.
Desmet P. M. A., Porcelijn R, van Dijk M. B. (2007). Emotional design; Application of a research-based design approach. *Knowledge Technology & Policy*, *20*(3), 141-155.
Di Serio, A., Ibáñez, M. B., & Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education, 68,* 586-596.
Hassenzahl, M., & Tractinsky, N. (2006). User Experience - a Research Agenda. *Behaviour & Information Technology, 25*(2), 91-97.
Ibáñez, M. B., Di Serio, A., Villarána, D., & Kloos, C. D. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education, 71*, 1-13.
Olsson, T., Lagerstam, E., Kärkkäinen, T., & Väänänen-Vainio-Mattila, K. (2013). Expected user experience of mobile augmented reality services: a user study in the context of shopping centres. *Personal and Ubiquitous Computing*, *17*(2), 287-304.
Vate-U-Lan, P. (2012). An augmented reality 3D pop-up book: the development of a multimedia project for English language teaching. *Proceedings 2012 IEEE International Conference on Multimedia and Expo*, 890-895.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Design of MOOC for In-service Teacher Professional Development: A Case of Teachers' Refresher Training Course in Hong Kong

Silu LI[a*], Eric T. H. LUK[b] & Morris S. Y. JONG[c]
[abc]*Centre for the Advancement of Information Technology in Education, The Chinese University of Hong Kong, Hong Kong.*
[c]*Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong*
*siluli@cuhk.edu.hk

Abstract: This paper describes the design of a video-based massive open online course (MOOC) with a scalable activity structure for in-service teachers' continuing professional development (CPD). The course has been developed and delivered in the open learning platform. We design this MOOC to aid teachers to learn anytime and anywhere, and exploit teacher professional training with the infusion of collaborative learning and self-directed learning.

Keywords: Massive open online courses (MOOCs), in-service teacher professional development, self-directed learning, computer supported collaborative learning

1. Introduction

The Internet has influenced the way we learn, work, live and socialize. The widespread use of computers and the Internet have made distance learning easier and faster. Massive Open Online Courses (MOOCs) are a recent development in distance education which began to emerge in 2012 (Pappano, 2012; Lewin, 2013). It has been anticipated that MOOCs will gain wide spread as an educational form. The rapid growth of information and communication technologies (ICT) and rising computer knowledge of people make possible appearance of this new educational form. Wikipedia (2014a) noted that there are two key features to a MOOC that contrast it with established university course delivery: (1) Open access-anyone can participate in an online course for free; (2) Scalability-courses are designed to support an indefinite number of participants.

In recent years, in-service teachers' continuing professional development (CPD) in ICT is a major priority in K-12 education in Hong Kong. In-service teachers are demanded to learn continuously so as to enhance their professional capacity. Through engaging in continuing professional learning and reflection, teachers can acquire the professional knowledge and skills in ICT for assuming the role and responsibilities of a teacher (Advisory Committee on Teacher Education and Qualifications, 2009). Education Bureau, the Government of the Hong Kong SAR has organized a lot of refresher training courses (RTCs) in information technology in education professional development programmes for teachers.

However, we found that the attendance rates of these courses were unsatisfactory. RTCs generally must take place after school, in the weekends or in the summer holiday, thus imposing on teachers' personal time, which cuts into time needed for other tasks. At the same time, we found that professional development experienced by substantial proportions of teachers lacks key pedagogical qualities that make it effective, including time to think about, collaborative learning, follow-up activities, and sharing with teachers from other schools (Smylie et al., 2001). So the sources and means of which RTCs are delivered to teachers should be carefully assessed.

To improve teacher professional development and foster their collaborative learning and self-directed learning, we design a MOOC, which has been developed to aid primary and secondary school teachers to implement effective teacher professional development training.

In the paper, we elaborate the reasons, principles and framework of the design of the above mentioned MOOC and conduct effective teacher professional development training to facilitate teachers' self-directed learning and computer-support collaborative learning.

2. Background

2.1 Massive open online courses (MOOCs)

MOOCs are online courses aimed at unlimited participation and open access via web (Wikipedia, 2014a) and are receiving increasing attention and interest from several communities involved in online distance education. The first MOOC emerged from the open educational resources (OER) movement, named "Connectivism and Connective Knowledge (CCK08)", was a unique event in 2008, which was led by George Siemens of Athabasca University and Stephen Downes of the National Research Council. In 2012, starting with the widely-publicized online courses at Stanford University, several universities are engaged in offering online versions of regular courses, through companies such as Coursera, Udacity and edX. Most in-service teachers cannot afford time to participate in conventional face-to-face professional development training courses. Thus MOOC provides a promising solution to the problem.

2.2 In-service Teachers Continuing Professional Development (CPD) in Hong Kong

Advisory Committee on Teacher Education and Qualifications (ACTEQ) in Hong Kong released "CPD Document 2003" in November 2003 and proposed a teachers' CPD framework. A "soft" target of 150 CPD hours in a three-year cycle is set, within which teachers can deliberate on the direction and content. Teachers' CPD refers to all kinds of learning opportunities that help them strengthen their professional practices. Today, ICT can facilitate not only delivery of instruction, but also learning process itself (Jung, 2005). Thus the government wishes to provide a variety of refresher training courses (RTCs) to teachers in Hong Kong to improve teachers' ability to promote student learning and development and to update teachers' knowledge and skills in teaching and learning.

Generally, the course duration of each training event is 6 hours (in two 3-hour sessions on two different days to give ample time for participants to complete assignments between the two sessions). Event must take place in the evening of workdays, in the weekend or summer. The maximum number of participants for each event is 25.

2.3 WebQuest

A WebQuest is an inquiry-oriented format in which all the information that learners work with comes from the web (Dodge, 1995). It can foster collaborative learning through collaborative activities with a group-based project, encourage independent thinking and motivate students. The use of WebQuest can serve as a powerful and an efficient tool for teacher professional development (Johnson, 2004). It is also a learner-centric project-based learning approach for facilitating K-12 students to pursue collaborative inquiry learning on the Internet, and is becoming an integral part of education. For effective integration of WebQuest into school education, trainings should be provided to teachers to equip them with the pedagogical knowledge and skills required, which will facilitate teachers' self-directed learning and collaborative learning. So we choose WebQuest as a production case to design this MOOC.

3. Design

3.1 Principles of the MOOC design

3.1.1 Self-directed learning (SDL)

Self-directed learning (SDL) is learning in which the conceptualization, design, conduct and evaluation of a learning project are directed by the learner (Brookfield, 2009). In the context of a self-directed learning environment, all decisions about how and what to learn, and how or whether to consult external resources, are decided by the learners. Some educational institutions are finding ways to support self-directed study through open-learning programs, non-traditional course offerings and other innovative programs (Knowles, 1975). Rita Kop mentioned that SDL on open online networks is now a possibility as communication and resources can be combined to create learning environments. Downes proposed that teaching strategies in his MOOC, named *Connectivism & Connective Knowledge 08,* allowed the educator to have the role of facilitator. Nowadays most of scholars and educators believed that MOOCs can facilitate participants' self-directed learning.

In a traditional classroom of RTCs, knowledge transfer from educators to learners. However, in this study, teachers learn WebQuest through MOOC can manage their time, find resources and choose the subject they want to learn about or the activities they want to engage in.

3.1.2 Computer Supported Collaborative learning (CSCL)

Computer Supported Collaborative learning (CSCL) regards learning as a social process, where each individual participant, or learner, is responsible for creating his or her own knowledge through social interaction with other human beings by interacting with physical objects (Miyake, 2007). CSCL refers to an instruction method in which learners at various performance levels work together in small groups toward a common goal (Gokhale, 1995; Johnson & Johnson, 1986). The shared learning gives learners an opportunity to engage in discussion, take responsibility for their own learning, and thus become critical thinkers (Sills, 1991).

According to Vygotsky, knowledge is social in nature and is constructed through a process of collaboration, interaction and communication among learners in social settings. The collaborative learning through MOOC provides learners with opportunities to analyze, synthesize, and evaluate ideas cooperatively. Because in traditional classroom, teachers barely discuss and just listen what the educators taught. They barely have chance to share and receive constructive feedback. The massive open online platform can facilitate discussion and interaction among teachers.

In this MOOC, we set up a discussion forum for teachers to ask questions and discuss with peers. During the learning process, this WebQuest MOOC can facilitate teachers' collaborative learning and then they can use WebQuest to teach their students to improve students' high-level thinking and inquiry-based learning skills.

3.2 MOOC Design

3.2.1 Platform

Moodle is an open-source Course Management System (CMS) that universities, communities, colleges, K-12 schools, and even individual instructors use to add web technology to their courses (Cole & Foster, 2007). It presents one of the most widely used open-source e-learning platforms, enabling the creation of a course website and ensuring their access only to enrolled participants (Dougiamas & Taylor, 2003). Moodle allows the exchange of information among users geographically dispersed, through mechanisms of synchronous (chats) and asynchronous communication (discussion forums). It also has easily configurable features, allowing the creation of participants assessment processes (quizzes, online tests and surveys), as well as managing their tasks with their timetable (Mahmoud, 2008; Costa, 2012). According to the advantages of Moodle platform to enrich the process of teaching and learning, in this study we design and establish all the course materials using learning modules of Moodle platform. Figure 1 shows the modules of MOOC platform. Teachers who would like to participant in this course can login the platform anytime and anywhere. It solves the limitation of teachers' time and geographical locations to a great extent.

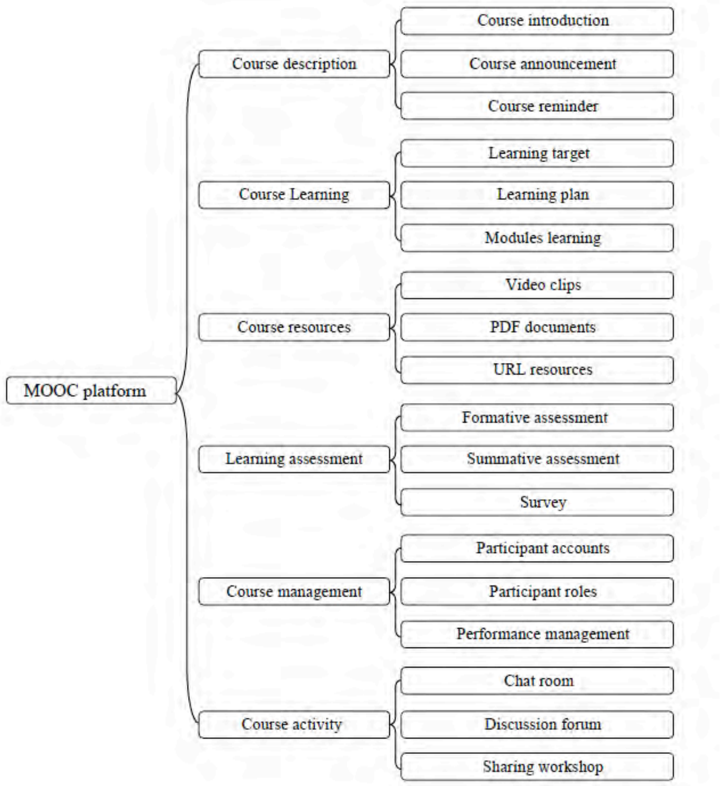


Figure 1. Modules of the MOOC platform

.

3.2.2 Basic requirements

This MOOC is mainly composed of 12 video-based lecture modules, formative assessment quizzes (each articulated to the end of a module), discussion forums for teacher community building (facilitated by an online tutor), and summative assessment tests. Teachers can select their own learning track by taking different lecture modules in accordance with their own grade of teaching (primary or secondary) and their own subject of teaching (four Key Learning Areas (KLAs): English Language, Chinese Language, Mathematics, and Humanities). After passing all formative assessment quizzes and completing all community-building and summative assessment tests, participants will be awarded an e-Certificate, and six hours of CPD by EDB.

3.2.3 Quality

The design and content of this course captures some trends that have emerged in recent years in the field of MOOCs learning. To ensure teaching and learning effectiveness, course designers, instructors and online tutors involved in this course have provided teacher coaching support or teacher professional development in designing and implementing WebQuest in schools over five years.

3.2.4 Module design

The delivery structure for MOOC is based around a self-enrollment, self-paced completion of the learning activities that are presented over 2 months. In the system, teachers from different disciplines will decide when they would like to login the system to learn, where they want to set as a connection to assess to the course. What's more, they could select their own learning track by taking different lecture modules in accordance with their own disciplines. The instructional content is delivered via a series of short video clips, which range from 10 to 20 minutes in length. Total duration of these video clips is 6 hours. The video clips were high-quality and developed with a professional television crew in a recording studio setting.

We design this MOOC with 12 modules. Each module contains a video clip. Except Module 4, to tailor-made the needs of teacher participants teaching in different KLAs in primary and secondary schools, we prepare 8 versions of video clips correspond to 8 types of KLAs. Figure 2 shows the modules design of this WebQuest MOOC.

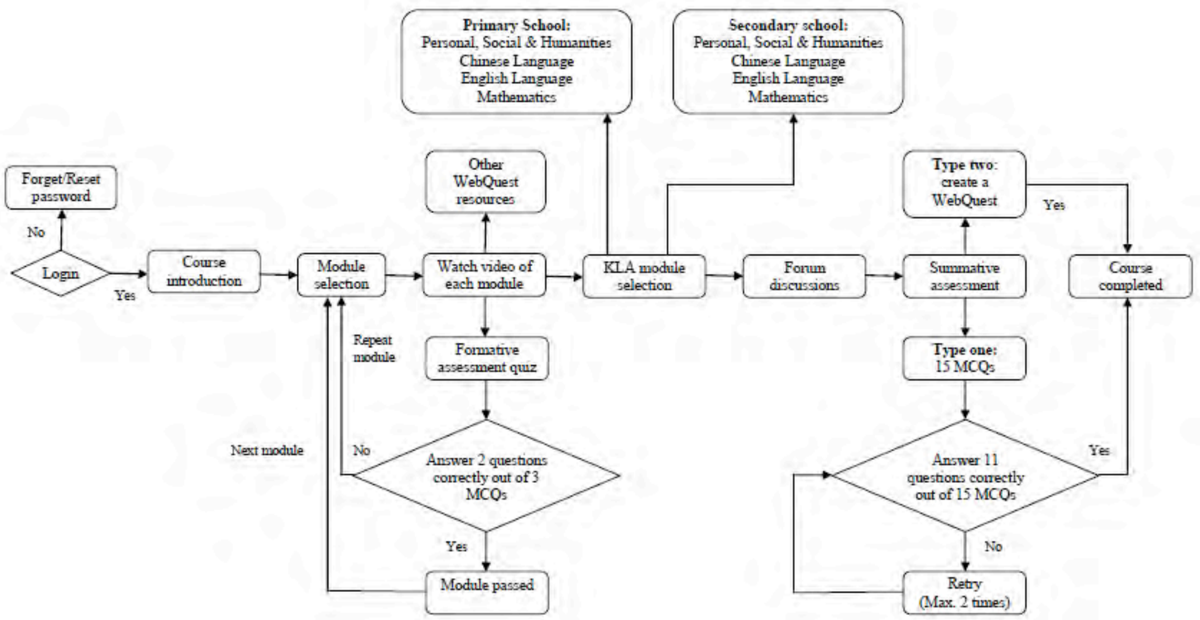


Figure 2. Modules design of this WebQuest MOOC

3.2.5 Assessment

Since teachers may take different modules in different time, to give them larger flexibility to join this course, we arranged formative assessment quizzes with 3 multiple-choice questions (MCQs) at the end of each module. The MCQs were presented to teachers with the goal of allowing teachers to gauge their understanding of the material presented in each video clips. Teachers have to get a pass with at least 2 correct answers in a quiz before proceeding to the next one. Participants can download the PowerPoint file of course material in each module as a reference when taking the quizzes. Any queries after joining the assessment, participants can raise questions in the online discussion forum. Online tutors will give teachers guidance to acquire the knowledge.

For the sake of modifying teachers' thinking or behavior to improve their learning, we design formative assessment quizzes at the end of each module. And the feedbacks should be supportive, timely, and specific. The feedbacks come in a variety of types, such as verification of response accuracy, explanation of the correct answer, hints, and worked examples. If teachers didn't pass the quiz, they can repeat watching the correspondent video and search for supporting resources from the web to improve their learning.

At the end of the course, we designed two types of summative assessment quizzes for teacher participants. Type one is to finish a quiz with 15 MCQs. Teacher must get a pass with at least 11 correct answers. Teachers will have 3 chances to attempt. The platform will give the prompt feedback to teachers. Type two is to create a WebQuest originally according to the specific KLA and level (primary and secondary) the teachers teach. Teachers who will choose this type of summative assessment are recommended to form a group to discuss how to create, evaluate, adopt and improve a WebQuest. This process achieves the CSCL that teacher will engage in the discussion to form their own construction of learning.

3.2.6 Activities

Instead of merely watching video clips online passively, this MOOC also want to facilitate teachers' SDL and CSCL and let their online learning actively. We established a wide variety of learning activities, like hands-on practice and online discussion forum responding reflective and collaborative questions.

In addition, inter-school subject-specific teacher communities are built for users to share good teaching practices.

4. Conclusion

In this WebQuest MOOC, with computer and Internet, teachers have the opportunity to study anywhere regardless of their busy schedule. The principles and flows of design of this MOOC provide educational designers with insights on designing MOOCs for in-service teachers' professional development training course. Next step, we will evaluate the acceptance and learning experience of teachers who have completed this MOOC. And we will explore if our work can apply to other courses of teacher professional development.

References

Advisory Committee on Teacher Education and Qualifications (ACTEQ) (2009). *Towards a Learning Profession: Third Report on Teachers' Continuing Professional Development,* Government Printer, Hong Kong.

Brookfield, S. D. (2009). Self-directed learning. In *International handbook of education for the changing world of work* (pp. 2615-2627). Springer Netherlands.

Cole, J., & Foster, H. (Eds.). (2007). *Using Moodle: Teaching with the popular open source course management system.* O.R. Media: United Sates of America.

Costa, C., Alvelos, H., & Teixeira, L. (2012). The use of Moodle e-learning platform: a study in a Portuguese university. *Procedia Technology, 5,* 334-343.

Dodge, B. (1995). Some thoughts about WebQuests. *The distance educator, 1*(3), 12-15.

Dougiamas, M., & Taylor, P. (2003). Moodle: Using learning communities to create an open source course management system. *World conference on educational multimedia, hypermedia and telecommunications, 2003*(1), 171-178.

Downes, S. (2008). Places to go: Connectivism & connective knowledge. *Innovate: Journal of Online Education, 5*(1).

Gokhale, A. A. (1995). Collaborative learning enhances critical thinking.

Johnson, R. T., & Johnson, D. W. (1986). Cooperative learning in the science classroom. *Science and children, 24*, 31-32.

Johnson, D., & Zufall, L. (2004). Web watch not just for kids anymore: WebQuests for professional development. *Reading Online, 2*, 26-30.

Jung, I. (2005). ICT-Pedagogy Integration in Teacher Training: Application Cases Worldwide. *Educational Technology & Society*, *8*(2), 94-101.

Knowles, M. S. (1975). *Self-directed learning*. New York: association Press.

Kop, R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *The International Review of Research in Open and Distance Learning, 12*(3), 19-38.

Lewin, T. (2013). Universities abroad join partnerships on the web. *The New York Times, 21*.

Mahmoud, S. S. (2008). A proposed model for distributing e-courses content through mobile technology architectures. *World Academy of Science, Engineering and Technology, 37*(48), 274-289.

Miyake, N. (2007). Computer supported collaborative learning. *The SAGE handbook of e-learning research,* 248-265.

Pappano, L. (2012). The Year of the MOOC. *The New York Times, 2*(12), 2012.

Sills, T., Digby, A., & Russ, P. (1991). *Cooperative learning: A guide to research*. New York: Garland.

Smylie, M. A., Allensworth, E., Greenberg, R. C., Harris, R., & Luppescu, S. (2001). *Teacher professional development in Chicago: Supporting effective practice. Chicago,* IL: Consortium on Chicago School Research.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Wikipedia (2014a). Massive open online course. http://en.wikipedia.org/wiki/Massive_open_online_course

Learning to create Technological Pedagogical Content Knowledge through distributed leadership: A Case Study of a Singapore Future School

Ching Sing Chai[a*], Benjamin Wang[a] and Chun Ming Tan[b]

[a]*National Institute of Education, Nanyang Technological University, Singapore*
[b]*Nan Chiau Primary School, Singapore*
* chingsing.chai@nie.edu.sg

Abstract: This case study analyzed the forms of distributed leaderships practice in a future school in Singapore. From the technological Pedagogical Content Knowledge (TPACK) framework, the study attests that the various levels of leadership in the school are focused on different aspect of the TPACK and the collective effort in engaging in school reform for future learning can be viewed from the TPACK perspective.

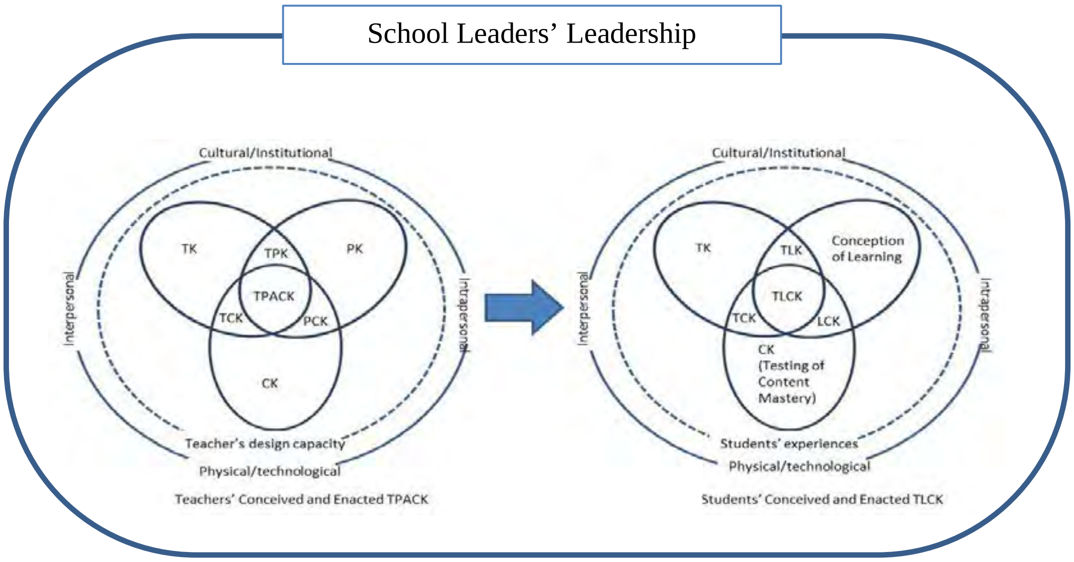
1. Introduction

Current development of Technological Pedagogical Content Knowledge (TPACK) research has documented a variety of methods of developing preservice and inservice teachers' TPACK (Kramarski & Michalsky, 2010). Most intervention studies that engage teachers and educators in collaborative designing of TPACK report positive outcomes (see Chai, Koh & Tsai, 2013). While substantial research has been carried out under the framework of TPACK among teachers, Chai, Koh, Lim and Tsai (2014) have pointed out the possibility of extending the TPACK research to school leaders. They argue that teachers' creation of TPACK is enabled through or constrained by higher level instructional decisions made by school leaders. These decisions include the technological environment the school leaders envisage, the pedagogical directions they set and the content areas for lesson design. These decisions shape subsequent TPACK creation among teachers and consequently students learning experiences. How school leaders enact their leadership and thus form the technological pedagogical content environment has however not been research empirically. In addition, Figure 1 below (adapted from Chai et al., 2013) depicts the possible design space where school leaders have to exercise their leadership to create the conditions and directions for TPACK to emerge.

2. Method

The study adopts a case study design in its approach to answer the research question. The research questions addressed in this study is how do school leaders enact their leadership as viewed from a TPACK perspective? The case school is a future school that has devoted itself in several pedagogical initiatives including the building of MyCLOUD (Wong, Chai, Zhang & King, in press) for Chinese language learning; the Idea Garden (Tsai, Chai & Hoe, 2014) to facilitate knowledge creation among student communities for social studies and MyDesk (Tan, 2013) for seamless science learning. The principal, vice-principal, head of department (HOD) for Information and Communication Technologies (ICT), and the social studies subject coordinator from a future school (primary) were interviewed for this study in the context of innovating the social studies through Idea Garden. In addition, records of meeting from 2013-2014 was collected. The data was analyzed using figure 1 as a guide. The data coded situate the loci of instructional decisions and actions that the school leaders have been devoting their energy for the past two years. The themes formulated below were sent to the participants for members checking to ensure that the researchers have represented their views fairly.

Figure 1: TPACK Design Space for School Leaders



3. Findings

With regards to the leadership practice, distributed leadership emerged naturally as the school leaders were engaged in actualizing the future school effort. The different roles undertaken by the leaders are shaped by institutional norms. The principal assumed leadership in contextualizing school directions towards cultivating 21st century competencies and ensuring learning of good foundational disciplinary-based knowledge. Based on that, he also enabled staff by structuring professional developments activities and sourcing for additional funding. His role is therefore more on interfacing with the external environment and people. The principal articulated his work as follow:

> At the beginning of the future school projects, my role is to lead my key personals, including the vice principals and the HODs (head of department) to contextualize the MOE (ministry of education) 21st century competencies framework for our school. This involves looking at what we have in terms of technology capacity, manpower capital and the students' general learning and family profiles. We went through rounds of intensive discussion and we agree to focus on values and key 21 competencies students need to have without compromising the foundational knowledge. Then we operationalize the framework and put up a proposal to obtain research and development fund for the school. I am also the person to source for potential collaborators from the IHL (institute of higher learning) to provide TPACK training for the KPs (key personals) and later the whole school. I set up collaborations with international researchers, industry partners and researchers from local university to bring in the necessary expertise to work with our teachers.

With the strategic directions set, the vice principal supported the school transformation by working with technology service providers, looking into the infrastructure of the school and ensuring that financial procedures are adhered to. For example, he drafted the documents to call for interested vendors to bid for the services and he reviewed the specification for learning analytics needed for Idea Garden. The vice principal's is therefore more focused on laying the necessary technological conditions within the school context. He acquired his technological expertise at the infrastructure level through his experience of working as education technology officer in the ministry of education. As for the HOD for ICT, he described his roles as ensuring alignment of school goals with the efforts devoted to ICT integration. He understands his work as guiding curriculum redesign and integrating the efforts for the various subject matters and providing support for the professional development activities. In his words,

> I see the need to take a step back and view KB and other projects from the system's perspective and see how, these projects collectively can enhance the learning experience of the student. The experience with KB reaffirms again the potential ICT has being integrated with the curriculum. Working closely with the co-PI and his team of researchers, project drivers, teachers, curriculum developers and even allied educators provide multiple insights to redesigning the existing curriculum and executing it in lessons. Research serves as an affirmation to delivering the planned redesigned curriculum. Anchored by a sound pedagogy with meaningful integration of the Idea Garden platform, we are experiencing first hand a new way of learning in our students, one that is important for the 21st century.
>
> The PD plan looks into the teacher's capacity building in three areas i.e. curriculum, pedagogy and technical aspects. In delivering the PD to the teachers, we work closely with Prof C and the curriculum developer and the ICT support staff team.

For the subject coordinator, Mr E, his main role is to ensure that the ministry curriculum goals are not compromised. He views his inputs and leadership for the project as follow:

> In creating new practices and procedures in my teaching, there is a rather drastic shift towards student-centered learning as part of 21CC skills. It takes a lot of getting used to especially relinquishing control of the students' everyday learning. A KB teacher is no longer a 'vessel' of knowledge to the students and requires a unique skills set of classroom management in order to have a successful KB classroom. In every lesson, the students are required to use various KB principles in their research and discussions.

In summary, he is responsible in ensuring adherence to the syllabus and at the same time creating the new practices needed for the new technologically supported pedagogy. He is also consolidating his experiences to pass on to his colleagues for future professional development "to equip teachers before they embark in KB".

4. Discussion

From the perspective of TPACK, the school leaders each occupy mainly a sub-domain of the TPACK factors. The principal established the general pedagogical push towards 21st century learning while the vice principals work mainly on the technological dimension in providing the necessary infrastructure. Building on the pedagogical direction sets the directions for reform. The ICT HOD works are situated in the sphere of technological pedagogical dimension while the expertise of the subject coordinator lies in content knowledge and pedagogical content knowledge. The distributed expertise was brought to bear on collectively on the creation of the platform and pedagogical practices encapsulated in the form of lesson plans and materials, and also professional development materials for scaling up purposes. The creation of platform, practices and materials can be viewed as knowledge creation efforts in the school context. This becomes the main form of professional development for the teachers to draw upon their existing knowledge and ideas and refine them in collaborative discussion. In other words, the learning processes is in essence a knowledge creation processes (Chai et al., 2014).

School improvement and transformation is complex in nature, especially given the rapid advancement of technology. Multiple levels of leadership have to work synergistically in a coordinated manner. Heck and Hallinger (2014) termed this as “leadership for learning”, which encompasses both instructional leadership and transformational leadership. While the principals in this case study are providing transformative leadership and developing knowledge for the technology and pedagogical dimensions, the HOD and subject head are working on instructional leadership in the technological pedagogical dimensions and the pedagogical content knowledge dimension.

Acknowledgements

This study was funded by the research grant NRF2011-EDU002-EL005.

References

Chai, C. S., Koh, J. H. L., & Tsai, C.-C. (2013). A review of Technological Pedagogical Content Knowledge. *Education Technology and Society, 16*(2), 31-51.

Chai, C. S., Koh E., Lim, C. P. & Tsai, C.-C. (2014). Deepening ICT integration through multilevel design of Technological Pedagogical Content Knowledge. *Journal of Computers in Education, 1*(1), 1-17.

Heck, R. H. & Hallinger, P. (2014). Modeling the longitudinal effects of school leadership on teaching and learning. *Journal of Educational Administration, 52*(5), 653-681.

Kramarski, B., & Michalsky, T. (2010). Preparing Preservice Teachers for Self-Regulated Learning in the Context of Technological Pedagogical Content Knowledge. *Learning and Instruction, 20*(5), 434-447.

Tsai, P.-S., Chai, C. S., & Hoe, K. E. (2014). Harnessing emerging technologies to build the next generation knowledge creation platform for school students. In Tan. S. C., So, H.-J., & Yeo, J. (eds.). *Knowledge creation in education,* (pp.75-96). Singapore: Springer.

Tan, C. M. (2013). Building the 21st century classroom with 3G smartphones in Sengkang New Town. *Edtech Digest.* Retrieved 25th August from http://edtechdigest.wordpress.com/2013/05/13/smart-in-singapore/

Wong, L.-H., Chai, C. S., Zhang, X., & King, R. B. (In-press). Employing the TPACK framework for researcher-teacher co-design of a mobile-assisted seamless language learning environment. *IEEE Transactions on Learning Technologies.*

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Development and validation of an instrument for exploring Taiwanese undergraduates' approaches to Internet-based learning

Yu-Chih TSAO*, Chi-Ling WU, & Min-Hsien LEE
Institute of Education, National Sun Yat-sen University, Taiwan
*watsao32@gmail.com

Abstract: Few studies have developed questionnaires that attempt to assess approaches to learning in the context of Internet-based learning. To obtain a better understanding of undergraduates' approaches to Internet-based learning, this study aimed to develop and validate the Approaches to Internet-based Learning (AIL) instrument which was created by referring to the qualitative results of Ellis' (2011) research and the structure of Lee et al.'s (2008) Approaches to Learning Science questionnaire. The AIL consists of the six factors of "deep motive," "criticism and evaluation", "reflection and integration," "surface motive," "collecting and summarizing," and "replicating information." In addition, the former three factors could be categorized as "Deep approaches to Internet-based learning" while the latter three could be grouped into "Surface approaches to Internet-based learning." To establish the reliability and validity of the AIL and to confirm its second-order structure as hypothesized in this study, exploratory factor analysis and second-order confirmatory factor analysis were conducted. A total of 598 undergraduates from seven Taiwanese universities participated in this study. The results support our hypothesized second-order structure of the AIL and indicate that the instrument items have good reliability and validity. We also found that the Taiwanese undergraduates tended to adopt deep approaches to Internet-based learning. The newly developed AIL could provide educators with a valid instrument to examine students' approaches to Internet-based learning.

Keywords: Approaches to Internet-based learning

1. Introduction

Following the advances in technology and associated innovations, Internet-based learning or online learning has been recognized as an effective way to enhance students' learning. In one study, 71% of students said that the Internet tended to be the major source of information for their school learning (Lenhart et al., 2001). Additionally, recent research has indicated that enrollment in online courses is growing at a rate approximately ten times that of traditional classroom-based instruction in higher education (Shea & Bidjerano, 2009). Due to the rapid development of Internet technology in education, numerous studies have investigated how students' characteristics contribute to learning in an Internet-based context. Some studies have attempted to investigate students' approaches to Internet-based learning. Ellis et al. (2011) investigated students' experiences of learning through research on the Internet, and identified four categories: "Critical focus and Evaluation," "Reflection and Integration," "Collecting and Summarizing," and "Replicating information." The four categories seem to reflect the various forms of students' strategies in the context of Internet-based learning. In addition, Marton and Säljö (1976) found that undergraduates' approaches to learning could be classified as 'surface approaches to learning' and 'deep approaches to learning.' In Ellis et al.'s (2011) study, the first two categories could be categorized as deep approaches while the latter two could be categorized as surface approaches.

In the area of learning approaches, several studies have used questionnaires to explore students' approaches to learning such as Kember et al. (2004) and Lee, Johanson and Tsai (2008). Kember et al.'s (2004) and Lee et al.'s (2008) questionnaires consisted of the four main factors of deep motive for learning, deep strategies for learning, surface motive for learning, and surface strategies for learning, which could also be grouped as deep approaches and surface approaches. However, few studies have developed questionnaires that attempt to assess approaches to learning in the context of Internet-based learning.

A larger scale quantitative survey might obtain a better understanding of students' approaches to Internet-based learning. Moreover, second-order confirmatory factor analysis can be used to examine whether the motive for learning and strategies for learning can be framed by higher-order categorization such as surface and deep approaches. By referring to the four categories of Ellis et al. (2011) and the structure of Lee et al.'s (2008) instrument, the aim of this study was to develop and validate an instrument, namely the Approaches to Internet-based Learning (AIL) instrument, for exploring Taiwanese undergraduates' approaches to Internet-based learning.

2. Method

2.1 Participants

The participants consisted of a total of 598 undergraduates (261 male) from seven universities in Taiwan. All participants responded to the AIL instrument. The participants were then split into two subsets for the exploratory factor analysis (EFA) (n = 445) and the second-order confirmatory factor analysis (CFA) (n = 153).

2.2 Instrument assessing undergraduates' approaches to Internet-based learning (AIL)

The AIL instrument was created with reference to the qualitative results of Ellis' (2011) research and the structure of Lee et al.'s (2008) ALS (Approaches to Learning Science) questionnaire. The AIL consists of the six factors of "deep motive," "deep strategy A: criticism and evaluation", "deep strategy B: reflection and integration," "surface motive," "surface strategy A: collecting and summarizing," and "surface strategy B: replicating information." In addition, the former three factors could be categorized as "Deep approaches to Internet-based learning" and the latter three could be grouped into "Surface approaches to Internet-based learning."

Through consulting with four experts in this field for the content validity, this study constructed 5 items for each of the six factors, giving a total of 30 items presented in a five-point Likert mode, ranging from "strongly agree" to "strongly disagree." A detailed description of the six factors, with a sample item for each, is presented below:

- Deep motive (DM): The student has a deep motive (e.g., intrinsic interest) for Internet-based learning, e.g., When I learn in the context of Internet-based learning, I feel happy and contented.
- Criticism and Evaluation (CE): The student uses critical thinking and information evaluation in the context of Internet-based learning, e.g., When I am learning in the context of the Internet, I check different websites at the same time to judge information.
- Reflection and Integration (RI): Knowledge reflection and integration are used in the context of Internet-based learning, e.g., When I use the Internet for learning, I like to create a theory to help me put the fragmented content together.
- Surface motive (SM): The student uses Internet-based learning just to pass exams or meet the requirements of the course, e.g., I use the Internet for learning in order to get a good grade.
- Collecting and Summarizing (CS): Collecting and summarizing information are the main strategies adopted in the context of Internet-based learning, e.g., When I use the Internet for learning, irrelevant contents do not make sense to me.
- Replicating information (Rep): Replicating information is the main strategy used in Internet-based learning, e.g., When I use the Internet for learning, I think the best way to get a good grade is to memorize the answers to related questions.

2.3 Data analysis and procedure

The purpose of this study was to develop and validate an instrument, namely the Approaches to Internet-based Learning (AIL) instrument, for exploring Taiwanese undergraduates' approaches to Internet-based learning. In order to establish the reliability and validity of the AIL, both exploratory and second-order confirmatory factor analyses were performed. The reduction of items of the AIL was based on two sets of evidence: an exploratory factor analysis, followed by the use of reliability statistics. In an EFA, only those items with a factor loading of at least 0.40 within their own factor should be retained (Stevenson, 1996). EFA was employed first and then the second-order confirmatory factor analysis (CFA) was performed to analyze the construct validity and structure of the AIL.

Accordingly, the validity and reliability of the AIL were evaluated. In addition, the participants were split into two subsets for the exploratory factor analysis (EFA) (n = 445) and for the second-order confirmatory factor analysis (CFA) (n = 153).

3. Results and Discussion

3.1 Exploratory factor analysis for the approaches to Internet-based learning (AIL) (n=445)

An exploratory factor analysis with a varimax rotation was performed to clarify the structure of the AIL. As a result, the 445 students' responses were grouped into the following six factors: Deep motive (DM), Criticism and Evaluation (CE), Reflection and Integration (RI), Surface Motive (SM), Collecting and Summarizing (CS), and Replicating information (Rep). The eigenvalues of the six factors from the principle component analysis were all larger than one, while six items with a factor loading of less than 0.40 were omitted from the instrument. As a result, a total of 24 items were retained in the final version of the AIL (as shown in Table 1), and the total variance explained is 65.62%.

Table 1: The exploratory factor analysis for the AIL factors (n = 445)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Factor 1: Deep Motive (Mean = 3.48, SD = .65, α = .84)						
DM 1	.72					
DM 2	.81					
DM 3	.75					
DM 4	.71					
DM 5	.63					
Factor 2: Criticism and Evaluation (Mean = 3.84, SD = .65, α = .83)						
CE 7		.74				
CE 8		.68				
CE 10		.83				
CE 11		.64				
Factor 3: Reflection and Integration (Mean = 3.67, SD = .65, α = .81)						
RI 12			.53			
RI 13			.80			
RI 14			.77			
RI 15			.68			
RI 16			.67			
Factor 4: Surface Motive (Mean = 3.14, SD = .79, α = .68)						
SM 18				.68		
SM 20				.61		
SM 21				.68		
Factor 5: Collecting and Summarizing (Mean = 2.92, SD = .80, α = .81)						
CS 22					.80	
CS 23					.81	
CS 24					.81	
CS 25					.63	
Factor 6: Replicating information (Mean = 3.00, SD = .85, α =.79)						
Rep 28						.82
Rep 29						.78
Rep 30						.77

Total variance = 65.62%, overall α = 0.87

In addition, the reliability (Cronbach's alpha) coefficients respectively for these factors were 0.84, 0.83, 0.81, 0.68, 0.81, and 0.79, and the overall alpha was 0.87, suggesting that these factors had sufficient reliability in assessing the students' approaches to Internet-based learning.

3.2 Second-order confirmatory factor analysis for the AIL (n=153)

The CFA further confirmed the construct validity and the second-order structure of the 24 item version of the AIL through the 153 students' responses. In addition, the second-order factor analysis model of the AIL was hypothesized in this study. To examine whether the second-order structure of the AIL can be established, the first-order factors which converged to the second-order constructs were tested. Each factor of the AIL was a first-order construct (i.e., DM, CE, RI, SM, CS, and Rep). The deep and surface approaches to Internet-based learning served as the second-order constructs of the AIL. As shown in Table 2A, the factor loading values for the six factors are significant and larger than 0.4. The ratio of chi-square per degree of freedom = 1.46, RMSEA = 0.055, CFI = 0.95, GFI = 0.84, AGFI = 0.80. Moreover, the results shown in Table 2B support that the first-order factors converge to the second-order constructs. These results suggest an acceptable model fit which supports our hypothesized second-order structure of the AIL and indicates that the instrument items have good convergent and construct validity in this model.

Table 2: The Second-order Confirmatory factor analysis (CFA) for the AIL factors (n = 153)

A. Standardized CFA first-order loading

Factors	Number of item	Factor loading	*t*- value	CR
Deep motive (DM)	5	0.69 - 0.86	8.97^{*} - 11.82^{*}	0.88
Criticism and Evaluation (CE)	4	0.70 - 0.75	7.38^{*}- 7.84^{*}	0.81
Reflection and Integration (RI)	5	0.45 - 0.69	4.51^{*}- 4.80^{*}	0.74
Surface Motive (SM)	3	0.54 - 0.71	4.61^{*}- 4.60^{*}	0.66
Collecting and Summarizing (CS)	4	0.62 - 0.70	6.02^{*}- 6.35^{*}	0.76
Replicating information (Rep)	3	0.48 - 0.91	5.12^{*}- 6.12^{*}	0.75

B. Standardized CFA second-order loading

Second-order factor model	Loading value	*t*- value
Deep approaches		
DM	0.71	7.21^{*}
CE	0.79	6.81^{*}
RI	0.88	4.83^{*}
Surface approaches		
SM	0.61	3.54^{*}
CS	0.51	3.71^{*}
Rep	0.68	4.67^{*}

* $p < 0.05$; CR: Composite Reliability

3.3 Paired-t-test for the AIL (n = 598)

In order to understand students' learning motive and strategy in the context of Internet-based learning, this study calculated the mean values for deep strategy as combining CE and RI, and for surface strategy as combining CS and Rep. Then, a paired t-test was conducted to examine whether the students tended to use deep or surface approaches to learn in the context of Internet-based learning. As shown in Table 3, significant differences were found. The results seem to indicate that, rather than surface approaches, these Taiwanese undergraduates tended to have deep motives and adopt deep strategies to learn in the context of Internet-based learning.

Table 3: Differences between Deep AIL and Surface AIL (n = 598)

Deep AIL	Surface AIL (M, SD)	*t* value
Deep Motive (M, SD)	Surface Motive	
3.52 (0.67)	3.20 (0.79)	9.25***
Deep Strategy (M, SD)	Surface Strategy (M, SD)	
3.75(0.57)	2.97(0.69)	21.96***

***p <.001.

In conclusion, the abovementioned results suggest that the newly developed AIL instrument has sufficient reliability and validity, and could serve as a valid instrument for evaluating undergraduates' approaches to Internet-based learning.

References

Ellis, R. A., Goodyear, P., Bliuc, A. M., & Ellis, M. (2011). High school students' experiences of learning through research on the Internet. *Journal of Computer Assisted Learning, 27*, 503-515.

Kember, D., Biggs, J., Leung, D. Y. P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, 74, 261-280.

Lee, M.-H., Johanson, R. E., & Tsai, C.-C. (2008). Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. *Science Education, 92*, 191-220.

Lenhart, A., Rainie, L., & Lewis, O. (2001). *Teenage life online: The rise of the instant- message generation and the internet's impact on friendships and family relationships.* Washington, DC: Pew Internet and American Life Project.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—outcome and process. *British journal of educational psychology, 46*, 4-11.

Shea, P., & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster "epistemic engagement" and "cognitive presence" in online education. *Computers & Education, 52*, 543 – 553.

Stevenson, J. (1996). *Applied multivariate statistics for the social science* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Investigating Chinese University Students' Perceptions about Blackboard Platform to Support Their Online Learning

Weisheng LI[a], Lizhu CHEN[a], Ge QU[a] & Yan DONG[a*]
[a]*Faculty of Education, Beijing Normal University, China*
*dongyan98@126.com

Abstract: In this paper, we investigated the current situation of BB (Blackboard) platform using and the Chinese university students' learning attitudes towards BB platform. Data were collected from 164 students (51 male and 113 female) in Beijing via self-reported questionnaires: Questionnaire of Perceptions about Blackboard platform (PBPQ), including three scales (Ease of Use, Usefulness, Attitude). Results showed that the PBPQ can be used for the Chinese context with good validity and reliability. Further analysis results indicated that different frequencies of using BB platform were related the most to all the three scales except gender, major and grade. The results also showed that students' learning attitudes can be improved by ameliorating the Ease of use and Usefulness of BB platform, and teachers should encourage students to use BB more frequently if they want to use BB platform for teaching. These results give some advice for future course designs on BB platform.

Keywords: Blackboard platform, perception, current situation, learning attitude

1. Introduction

The BB platform is one kind of network platforms that was developed by the U.S. Blackboard company. In October 2003, it was introduced to China by the Beijing CERNET Blackboard company. BB platform is the largest network teaching platform in the world, which can support millions of users. In the market of network teaching platform, BB platform has occupied 85% of the American market and more than 50% of the global market (Song, 2009). Previous studies (Athena & Maker, 2004; Zhang, 2012; Meng, 2011) mainly paid attention to user experience of BB platform and how to design a course with BB platform. For example, Athena and Maker (2004) introduced many teachers' use experience and course management results with BB platform. Zhang (2012) did a research on network collaborative learning. Meng (2011) did the research about how to design an interactive mode based on Blackboard platform. However, there was less research on the university students' learning attitudes towards BB platform. In order to explore about the Chinese college students' perceptions about BB platform, this study aims to take the PBPQ questionnaire, developed by Tsai, Tsai and Hwang (2012) to make a deep understanding about Chinese students.

In sum, this study included the following research questions:

1) Can this questionnaire be used for Chinese mainland college students to measure their perceptions about BB platform?
2) What is the current using situation about BB platform in China?
3) Is there any difference between different major, grade and gender students' perceptions about BB platform?
4) Is there any difference among different frequencies of using BB students' perceptions about BB platform?

2. Methods

2.1 Participants

The participants in this study included 164 college students in Beijing, China. The study took part in the convenience sampling. There were 51 male and 113 female students. They were from different grade (G1: G2: G3: G4 =54:44:42:24) and different major (Liberal arts: Science: Others=69:89:8) in a famous university in Beijing, China. All participants were volunteers to complete the questionnaires in one sitting to explore the current using situation of BB platform and their learning attitudes towards BB platform.

2.2 Instruments

This present study used the instrument developed by Tsai (2012). It was validated for Taiwan students, similar to the Chinese context. This instrument employed a multi-dimensional framework to show student' perceptions about Blackboard platform.

This instrument included the following three scales, with a sample item for each scales:

a) Ease of Use: This scale measures perceptions of the extent to which students prefer that BB platform are easy to use. The example item is 'The search results which are displayed in BB platform are clear and easy to read'.
b) Usefulness: This scale measures the usefulness of BB platform in different students' eyes. The example item is 'The use of BB platform would enhance students' ability to search for information when problem-solving'.
c) Attitude: This scale measures students' learning attitudes towards BB platform. The example item is 'I have more willingness to use BB platform than before'.

2.3 Data Collection and Analysis

Before filling the questionnaires, an instructor trained by the research team informed the responding students of the purpose of the study. The process first involved a survey of each participant's demographic data such as gender and grade. Then, the participant was asked to answer the PBPQ. It took about 15 min to complete this whole survey.
The PBPQ was utilized in this study. Exploratory factor analysis was used to clarify the questionnaire structure. Accordingly, the validity and reliability of the questionnaire was also evaluated. Moreover, an independent t-test was conducted to explore whether there were differences between the two major groups of students. In addition, Pearson's correlation analysis was performed on the students' responses to PBPQ for the students.

3. Results

3.1 Factor Analyses

To ensure a well factor structure and reliability of the factors, it is necessary to make a reexamination. The following Table 1 respectively reveals the consequences of factor analyses for the aforementioned three factors: Ease of use, Usefulness, Attitudes. To get the final instrument, items whose factor loading weighs more than 0.5 were retained. Therefore, there are 10 items showed in Table 1 in total, and the total variance explained is 77.121%. For every scale, alpha coefficient was from 0.819 to 0.897, indicating favorable interaction consistency to do the statistical analysis.

Table 1: Rotated factor matrix and the alpha coefficient for the Ease of use and Usefulness of BB platform (N=164).

	Factor 1	Factor 2	Factor 3
	Ease of use	Usefulness	Attitudes
Factor1 : Ease of use, α=0.897, mean=2.53, SD= 0.41			
Ease of use 1	0.83	—	—
Ease of use 2	0.85	—	—
Ease of use 3	0.78	—	—
Ease of use 4	0.80	—	—

Factor2 : Usefulness, α=0.853, mean=2.97 , SD= 0.45			
Usefulness 1	—	0.86	—
Usefulness 2	—	0.83	—
Usefulness 3	—	0.70	—
Factor3 : Attitudes, α=0.819, mean=2.99 , SD= 0.34			
Attitudes 1	—	—	0.81
Attitudes 2	—	—	0.82
Attitudes 3	—	—	0.70

3.2 The correlation between the frequency of Using BB and the Ease of use, Usefulness, the Attitudes

According to the Table 2, Ease of use was related to both student's age and frequency of using BB. Usefulness was related to their frequency of using BB while Attitudes were linked to both major and frequency.

Table 2: The correlation between the frequency of Using BB and the Ease of use, Usefulness, the Attitudes.

	Ease of use	Usefulness	Attitudes
Age	0.190^{*}	0.063	0.126
Major	0.049	0.100	0.175^{*}
frequency of using BB	0.318^{**}	0.309^{**}	0.358^{**}

** p<0.05; **p<0.01*

3.3 Step-wise Regression Analysis

After analyzing the correlation between the factors, the study attempts to predict student's attitudes towards the BB. As following(Table 3), in the regression equation, the coefficients of Ease of use, Usefulness, Frequency of Using BB respectively were 0.403, 0.284, 0.154 while the constant equaled to 0.648.

Table 3: Step-wise Regression Analysis for predicting Students attitudes towards the BB.

		B	B	T	R
Attitudes	Ease of use	0.403		6.140	0.632
	Usefulness	0.284		4.358	0.689
	Frequency of using BB	0.154	0.181	2.173	0.700

3.4 Major difference

The study compared the possible difference between Liberal arts majors and Science majors. As shown in Table 4, Science major students had higher scores than students who majored in liberal arts, especially on attitudes (p<0.05). It means that Liberal arts major students are more likely to accept this platform easily, to get more convenient learning method and better sources.

Table 4: Major comparisons of the scores for all the scale.

scale	major(n)	Mean	SD	T	sig
Ease of use	Liberal arts majors(51)	2.48	0.97	0.14	0.704
	Science major(105)	2.54	1.00	0.10	
Usefulness	Liberal arts majors(51)	2.80	0.98	0.14	0.107
	Science major(105)	3.06	0.96	0.09	
Attitudes	Liberal arts majors(51)	2.72	0.77	0.11	0.014*
	Science major(105)	3.11	0.99	0.09	

3.5 Grade difference

This study also made a comparison between junior grade (including freshman and sophomore students) and senior grade (the third and fourth grade) to look for possible differences. Similarly, from the result of t-test, several significant differences were showed in Table 5. It revealed that senior grade students had better scores on Ease of use ($p<0.05$) than junior grade students. That is, senior grade students thought it easier to use than junior grade students.

Table 5: Grade comparisons of the scores for all the scale.

scale	grade(n)	Mean	SD	T	Sig
Ease of use	Junior grade(73)	2.73	1.00	0.12	0.020*
	Senior grade(91)	2.37	0.94	0.10	
Usefulness	Junior grade(73)	3.06	0.93	0.11	0.312
	Senior grade(91)	2.90	1.00	0.11	
Attitudes	Junior grade(73)	3.05	0.96	0.11	0.455
	Senior grade(91)	2.94	0.90	0.09	

3.6 Frequency of using BB difference

The comparison between high- frequency (frequently or often) and low-frequency (occasionally or sometimes) of using BB also revealed some differences by t-test. What showed in Table 6 is, Students who used BB in high-frequency had more identity on BB than the rest part students, they thought the BB platform were not only easy to use ($p<0.01$), but also useful ($p<0.01$), in addition, their attitudes ($p<0.01$) towards BB were more desired.

Table 6: BB using frequency comparisons of the scores for all the scales.

scale	frequency(n)	mean	SD	T	Sig
Ease of use	high- frequency(59)	2.26	0.84	0.11	0.008**
	low-frequency(105)	2.68	1.02	0.10	
Usefulness	high- frequency(59)	2.67	0.95	0.12	0.002**
	low-frequency(105)	3.15	0.94	0.09	
Attitudes	high- frequency(59)	2.71	0.80	0.10	0.003**
	low-frequency(105)	3.15	0.95	0.09	

4. Discussion and Conclusion

This study reports the validation of Questionnaire of Perceptions about Blackboard platform (PBPQ). A sample of students joined in reporting their perceptions of all prepared items. The instrument was certified reliable and valid based on three scales for Chinese mainland university students.

In this study, Table 2 indicated that Frequency of using BB was related the most to all the three scales while major, gender and grade had little relation. Table 3 predicted that students' learning attitudes can be improved by ameliorating the BB platform's Ease of use, Usefulness and encouraging students to use BB more frequently.

Table 4 and Table 5 uncovered that there were noticeable difference between different majors and grades students in the perceptions of BB platform. It showed that Liberal arts major students were likely to learn by BB more than Science major students were. The probably reason was that BB had stronger function on sharing literal files instead of Science instruments or so. As for the Ease of using BB, the senior grade students considered it easier, and this might be related to their using times or frequency compared to the junior grade students. It was also confirmed in Table 6 that higher frequency of using BB led to better experience. Its reason may be that students found more functions convenient to learn subjects to get a better impression of BB platform.

There were some advises from participants for BB platform such as associating with mobile applications, making friendlier interface, and setting modules of adding online friends and so on. All these results can give those courses designers of BB platform more valuable content.

Acknowledgements

We would like to thank for the funding project from School of Educational Technology, Beijing Normal University in China (No 705105602).

References

Tsai, P.-S., Hwang, G.-J., Tsai, C.-C., Hung C.M., & Huang, I. (2012). An electronic library-based learning environment for supporting web-based problem-solving activities. *Educational Technology & Society*, 15(4), 252-264.

Koevoets Rosanne, Allaart Cornelia F, van der Heijde Désireé M F M, Huizinga Tom W J. (2010). Disease activity monitoring in rheumatoid arthritis in daily practice: experiences with METEOR, a free online tool. [J]. *Journal of Rheumatology*, 2010, 3712.

Skhal Kathryn J, Thureson Catherine. (2007). ICONic training: use of a course management system to provide continual reference student education. [J]. *Medical Reference Services Quarterly*, 2007, 262.

Song Chengying. (2009). Instructional design problems and countermeasures based on Blackboard platform web. [J]. *Guangzhou City Polytechnic College*, Jan. 2009: 80-83.

Engineering Graduate Students' Literature Searching Behaviors

Ying-Hsueh CHENG[a*] & Chin-Chung Tsai[b]
[a]*National Pingtung University of Science and Technology, Taiwan*
[b]*National Taiwan University of Science and Technology, Taiwan*
*sherrycheng85@gmail.com

Abstract: In this paper, we drew on Kuiper, Volman and Terwel's (2008) three aspects of web literacy skills (searchig, reading, and evaluating) and investigated engineering graduate students' literature searching strategies. Through in-depth interviews and think-aloud protocols, we compared strategies employed by twenty-two engineering graduate students in Taiwan. The findings showed that the students' literature searching behaviors included searching, reading, and evaluating sources for their research purposes. Implications for enhancing information literacy of engineering students are suggested.

Keywords: Engineering, graduate students, literature searching

1. Introduction

During the past decades, studies have been conducted to examine graduate students' literature searching behaviors in several disciplines: education (Moselen, 2011), humanities (Bronstein, 2007), physics and astronomy (Jamali & Nicholas, 2010), and basic and medical sciences (Hemminger, Lu, Vaughan, & Adams, 2007). However, few have examined engineering graduate students' seeking behaviors. How engineering students actually search literature and use information for research purposes has been little studied. Thus, this study aims to explore Taiwanese engineering graduate students' literature searching behaviors in three aspects: (1) how do they seek and obtain online information for research? (2) how do they read and interpret sources? and (3) how do they assess and evaluate those sources for their research tasks?

2. Literature Review

2.1 Graduate Students' Literature Searching Behaviors

During the past decades, studies have been conducted to understand graduate students' literature searching behaviors. Some have conducted large-scale studies with the purpose to categorize students' search behaviors across diverse disciplines (Du & Evans, 2011; Rempel, 2010). However, such approach has been criticized by those who value a discipline-based approach (Grafstein, 2002; Jamali & Nicolas, 2010). These researchers argue that it is equally important to examine skills required for acquiring knowledge or conducting research in a specific subject area (Talja, Vakkari, Fry, & Wouters, 2007). For instance, Barrett (2005) explored humanities graduate students' seeking practices and found that they used *Google* to find general information on a topic, and other techniques such as citation chasing, identifying primary sources to validate their theories and hypotheses, having interpersonal contact for guidance, and constantly reading in a subject area.

Thus far, to the best of our knowledge, only one study conducted by Ismail and Kareem (2011) is related. Studying information seeking by Malaysian master's students in a computer science and information technology program, Ismail and Kareem found that these students had difficulties obtaining relevant information for their research. Unfamilar with specific scholalry tools, the students relied on *Google* or *Google Scholar* for searching but were overloaded with massive information.

Ismail and Kareem's research demonstrates struggle and strategies novice engineering researchers are likely to encounter when obtaining the required information and it indicates the importance of facilitating them to enhance their searches for academic literature.

2.2 Searching, Reading, and Evaluating Online Information

Several frameworks have been proposed to explore students' searching practices. Related to this study, Kuiper, Volman and Terwel's (2008) three aspects of skills: searching, reading, and evaluating, are considered. *Searching skills* refer to the ability to define appropriate keywords and use search engines or library databases. *Reading skills* include the ability to sort through massive information and decide what to use and what to neglect. *Evaluating skills* consist of the ability to assess the authority, relevance, and reliability of Web information. According to Kuiper et al. (2008), these three skills are mutually connected. Considering this study, we adopted Kuipert et al.'s framework and aimed to investigate how graduate students search, read, and evaluate information sources for their research purposes.

3. Methods

3.1 Participants

For the purpose of the study, fifteen master's students and seven doctoral students from two national universities in Taiwan were recruited. The students' research-related searching experiences ranged from one to seven years. Their specialized areas varied from material science and engineering, automation and control, to applied science and technology. They conducted academic literature searching for a variety of writing purposes (course assignments, grant projects, conference papers, journal articles, and theses/dissertations) and sometimes for experimental purposes (details of methodology in other studies).

3.2 Data Collection

Data sets included two parts: a semi-structured retrospective interview (30-40 minutes) and a concurrent think-aloud protocol (30-40 minutes). Audio recordings were collected from the interview and screen recordings were gathered from the think-aloud protocol.

3.3 Data Analysis

To answer the three research questions, the analyses were guided by Kuiper et al.'s (2008) three skills. We focused on how the engineering graduate students were engaged in searching, reading, and evaluating research-related information.

4. Findings and Discussion

The findings correspond with the three aspects of web literacy skills raised by Kuiper et al. (2008): searching, reading, and evaluating. In the following, we offer responses to the research questions and discussions of the findings in relation to the literature.

4.1 Searching

The participants reported four strategies they used for seeking and obtaining online information. These strategies included: using search engines and library databases, and networking for retrieving articles or materials.

In this study, seven students reported that they turned to *Google* or *Google Scholar* before the library databases because they included major science databases and thus provided sufficient sources related to their research. Frequent use of *Google* or *Google Scholar* by graduate students has been reported in the literature. Purposes indicated include conducting broad searches (Rempel, 2010) and using them as a starting point (Du & Evans, 2011) or to acquire new knowledge (Vibert, Rouet, Ros, Ramond, & Deshoullieres, 2007). The finding implies that it is beneficial to begin research by using search engines to browse and construct background knowledge and then use library databases for specific search purposes.

Moreover, networking for retrieving articles or materials was frequently reported by the engineering graduate students in this study. Aligned with previous studies, the finding revealed that engineering literature searching behaviors involve dynamic source-seeker connections for retrieving information efficiently (Xu, Tan, & Yang, 2006). It is suggested that graduate students initiate contacts with colleagues, librarians, or experts inside or outside their programs to request information.

4.2 Reading

The participants reported that they selected main ideas for reading, including browsing the title first and then the abstract, and referring to specific sections in the article (e.g., methods or conclusion). These involve quick decisions to read online or to download an article for in-depth reading. Such practice corresponds with M. J. Tsai and C. C. Tsai's (2003) finding of selecting main idea strategy regarding how students grasp or summarize the main information provided in each Web page. The engineering students' practices for online and printed resources are more intertwined and interactive than those observed in Ellis, Cox, & Hall's study (1993).

The current study also observed that the engineering graduate students assessed the usefulness or relevance of information based on the graphs or images in the methods section. This is similar to Aurisicchio, Bracewell, & Wallace's (2010) finding regarding aerospace designers' use of drawings.

Moreover, it was found that to develop a comprehensive understanding of a research topic, the participants were aware that they had to read broadly in a subject area. Such finding is similar to humanities graduate students' practices (Barrett, 2005) and it indicates the process of building domain knowledge through constant searching and reading.

4.3 Evaluating

Four evaluation criteria were discussed by the participants regarding the usefulness of online sources. These included: *relevance*, *recency*, *credibility,* and *authority*. Different from the processes of searching and reading, these criteria are considered higher-order thinking skills, facilitating the users to differentiate and sort out online information critically. Without this process, literature searching would not be successful and thus it plays a crucial role in information seeking (Biddix, Chung, & Park, 2011; Currie, Devlin, Emde, & Graves, 2010; Head & Eisenberg, 2009).

5. Limitations and Implications

This study has some limitations and thus draws some implications for future research. In this study, a sample of Taiwanese engineering graduate students enrolled in the master's and doctoral programs was recruited. It is suggested that the theoretical frameworks and the conclusions of the study can be further explored in other student populations.

Interviews with those (e.g., librarians, peers, colleagues, or advisors) who facilitate the participants to conduct their literature searching may be needed in order to triangulate varied perspectives of their searching processes.

Finally, it is suggested that adopting mixed methods including surveys and interviews could perhaps gather a broad picture of users' searching perceptions and behaviors.

Acknowledgements

We would like to thank the 22 students for their participation in this study. Without their time and help, this study would not have been possible.

References

Aurisicchio, M., Bracewell, R., & Wallace, K. (2010). Understanding how the information requests of aerospace engineering designers influence information-seeking behaviour. *Journal of Engineering Design*, *21*(6), 707-730.

Barrett, A. (2005). The information-seeking habits of graduate student researchers in the humanities. *The Journal of Academic Librarianship*, *31*(4), 324-331.

Biddix, J. P., Chung, J. C., & Park, H. W. (2011). Convenience or credibility? A study of college student online research behaviors. *The Internet and Higher Education,* 14(3), 175-182.

Bronstein, J. (2007). The role of the research phase in information seeking behaviour of Jewish scholars: A modification of Ellis's behavioural characteristics. *Information Research*, *12*(3), paper 318.

Currie, L., Devlin, F., Emde, J., & Graves, K. (2010). Undergraduate search strategies and evaluation criteria: Searching for credible sources. *New Library World*, *111*(3/4), 113-124.

Du, J. T., & Evans, N. (2011). Academic users' information searching on research topics: Characteristics of research tasks and search strategies. *The Journal of Academic Librarianship*, *37*(4), 299-306.

Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, *49*(4), 356-369.

Grafstein, A. (2002). A discipline-based approach to information literacy. *The Journal of Academic Librarianship*, *28*(4), 197-204.

Head, A. J., & Eisenberg, M. B. (2009, February 4). Finding context: What today's college students say about conducting research in the digital age. *Project information literacy progress report*. The Information School, University of Washington. http://projectinfolit.org/pdfs/PIL_ProgressReport_2_2009.pdf. Accessed on 25 November, 2012.

Hemminger, B. M., Lu, D., Vaughan, K. T. L., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology, 58*(14), 2205-2225.

Ismail, M. A., & Kareem, S. A. (2011). Identifying how novice reseurachers search, locate, choose and use web resources at the early stage of research. *Malaysian Journal of Library & Information Science, 16*(3), 67-85.

Jamali, H. R., & Nicholas, D. (2010). Interdisciplinarity and the information-seeking behavior of scientists. *Information Processing & Management*, *46*(2), 233-243.

Kuiper, E., Volman, M., & Terwel, J. (2008). Integrating critical Web skills and content knowledge: Development and evaluation of a 5th grade educational program. *Computers in Human Behavior*, 24(3), 666-692.

Moselen, C. (2011). *Is it all just Google? The research behaviour of education masters students*. Unpublished master's thesis, University of Auckland.

Rempel, H. G. (2010). A longitudinal assessment of graduate student research behavior and the impact of attending a library literature review workshop. *College & Research Libraries*, *71*(6), 532-547.

Talja, S., Vakkari, P., Fry, J., & Wouters, P. (2007). Impact of research cultures on the use of digital library resources. *Journal of the American Society for Information Science and Technology*, *58*(11), 1674-1685.

Tsai, M. J., & Tsai, C. C. (2003). Information searching strategies in web-based science learning: The role of Internet self-efficacy. *Innovations in Education and Teaching International*, *40*(1), 43-50.

Vibert, N., Rouet, J. F., Ros, C., Ramond, M., & Deshoullieres, B. (2007). The use of online electronic information resources in scientific research: The case of neuroscience. *Library & Information Science Research*, *29*(4), 508-532.

Xu, Y., Tan, C. Y., & Yang, L. (2006). Who will you ask? An empirical study of interpersonal task information seeking. *Journal of the American Society for Information Science and Technology, 57*(12), 1666-1677.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Developing an Online Formative Assessment System for a Chinese EFL Course

Chunping ZHENG*, You SU & Jingjing LIAN
School of Humanities, Beijing University of Posts and Telecommunications, China
*zhengchunping@bupt.edu.cn

Abstract: This study aims to develop an Online Formative Assessment System for a Chinese EFL course to support conducting formative assessment activities and further improving learners' English language skills. Two questionnaire surveys were conducted on 653 participants at the end of the course to explore their attitude toward the course (ATTC) and perceptions of the system (POTS) respectively. The results showed that students held favorable attitude toward the course as it improved their language skills. Students also expressed their preference for the system in terms of its usefulness, ease of use, and their willingness for future use. Besides, stepwise regression analyses indicated consistency between students' ATTC and their POTS.

Keywords: Online formative assessment, language tasks, EFL, higher education

1. Introduction

Assessment plays a crucial role in the process of education for evaluating and further improving the teaching and learning. With the popularization of online and blended learning in higher education since the 21st century, effective integration of formative assessment in online learning environments was claimed to have the high potential for sustained meaningful interactions among learners and the teacher (Sorensen & Takle, 2005).

Formative assessment is increasingly being implemented through policy initiatives in the Chinese higher education setting (Chen, Kettle, Klenowski, & May, 2013). For instance, Wen (2011) and Zheng (2014) applied formative assessment to Chinese EFL courses. Though there is a refocused emphasis on online formative assessment in order to create learner and assessment centered learning environments (Pachler, Daly, Mor, & Mellar, 2010; Wang, Wang, & Huang, 2008), online formative assessment studies are still less established in Chinese EFL courses.

In the present study, an online formative assessment system was developed for supporting web-based formative assessment activities in an EFL course at a key university in North China. The system provides an innovative approach for guiding teachers to systematically and efficiently assign language learning tasks and supporting students to carry out online formative assessment activities. At the end of the course, two survey questionnaires have been conducted among 653 students who have experienced the course and the system to investigate their attitude toward the course and the system.

2. Development of the online formative assessment system

The online formative assessment system was developed in year 2009 and has been applied in the EFL course at the university for five years. As shown in Figure 1, the structure of the system consists of two major components: an "Online Course Materials Module" and an "Online Formative Assessment Module".

The "Online Course Materials Module" was designed by teachers, including three sections, namely, "Resources Center for Language Learning Tasks", "Rubrics for Online Assessment" and "Students' Demo Work". "Resources Center for Language Learning Tasks" is a database for 23 categories of language tasks. In total, 280 language tasks were designed meeting the learning objectives of 28 units in four EFL textbooks at different proficiency levels. "Rubrics for Online Assessment" are specific assessment criteria adapted from College English Curriculum Requirements (2007) to assess learners' language performance when completing the tasks and "Students' Demo Work" is for collecting and sharing all the sample works from the students.

The "Online Formative Assessment Module" enables the teacher to assign learning tasks, conduct teacher assessment and recommend students' demo work. It also provides students with functions for completing the tasks, uploading outcomes of tasks and conducting self and peer assessment.

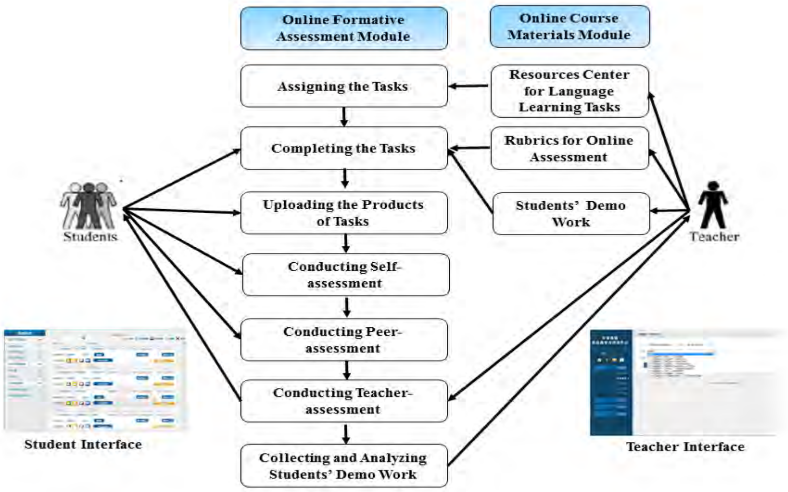


Figure 1. The Structure of Online Formative Assessment System

3. Results and Discussion

3.1 Students' attitude toward the EFL course and related course activities

After the EFL course, a questionnaire survey concerning the learners' attitude toward the course (ATTC) was conducted. In the questionnaire, students received 27 statements, each of which they rated on a 5-point scale the degree to which they agreed or disagreed (1= most strongly disagree and 5= most strongly agree). The questionnaire was modified based on the perception questionnaire designed by Fang (2010). A detailed description of the six factors is presented below:

- Leaners' attitude toward the EFL course (attitude): assessing students' general view of the course.
- Perceived usefulness for improving skills (improving skills): exploring effectiveness of the course for improving students' listening, speaking, reading, writing skills.
- Task completeness: measuring the degree to which students worked to complete the tasks.
- Extra effort: delineating students' after-class work for improving their language learning.
- Learners' attitude toward teacher assessment (teacher assessment): disclosing students' perceptions of teacher assessment activities in class.
- Learner' attitude toward peers' demo work (demo work): revealing the degree to which the peers' demon work is welcomed.

The reliability (Cronbach's alpha) for each factor is high (i.e., 0.89, 0.77, 0.75, 0.67, 0.84 and 0.75 respectively; the overall alpha coefficient is 0.91), indicating that these factors had sufficient reliability for measuring students' perceptions of the EFL course and the related student and teacher activities in class.

Students gave the highest ratings (Mean=4.56 on a 1-5 Likert scale) on the "learners' attitude toward teacher assessment" scale, implying the high importance of timely teacher feedback and assessment on learners' task performance. In addition, students have positive ratings of "task completeness" (Mean=4.13), "improving skills" (Mean=4.11), and "demonstration" (Mean=4.06), indicating their hard work for completing tasks and the effectiveness of the EFL course and peers' demo work for improving learners language skills. On the other hand, students gave relatively lower score to "attitude" and "extra effort" (Mean<4), which implied that this EFL course still had a room for improvement and students should be encouraged to make more extra efforts for completing tasks and improving overall performance.

3.2 Students' perceptions of using the online formative assessment system

The other questionnaire survey was conducted to evaluate the students' perceptions of the system (POTS). It consists of five factors, presented with strongly agree/disagree statements on a five-point Likert scale. The questionnaire was adapted from the questionnaire by Davis (1989), and Li, Dong, and Huang (2011), based on the TAM (technology acceptance model). Moreover, the researchers invited two experts in the field of online formative assessment to comment on the items of the questionnaire for face validity, and two EFL professionals to clarify the wording of all survey items. The following information is the detailed description of the five factors of the questionnaire.

- Perceived usefulness（usefulness）: assessing students' perceptions of the degree to which using the system will enhance their learning performance.
- Perceived ease of use (ease of use): exploring perceptions of the degree to which students expect the system to be free of effort.
- Learner satisfaction (satisfaction): measuring perceptions of the degree to which students will be willing to use the system and recommend it to peers.
- Effectiveness of online formative assessment (assessment effectiveness): delineating students' perceptions of the system for improving their second language learning skills.
- Willingness for future use scale (willingness): disclosing students' continuous efforts to use the system in the future.

The reliability (Cronbach's alpha) for each factor is high (i.e., 0.83, 0.78, 0.82, 0.88 and 0.71 respectively; the overall alpha coefficient is 0.92), indicating that these factors had sufficient reliability for measuring students' views toward the system. Students gave the highest ratings (Mean=4.23 on a 1-5 Likert scale) on the "perceived ease of use" factor, implying that the system has high potential supporting online formative assessment with proper user interface design. In addition, students have positive perspectives of the system and gave high ratings on the factors of "learner satisfaction", "willingness for future use" and "perceived usefulness" (all scores above 4), implying that the system has high potential for assisting students in completing online formative assessment activities, hence, students are willing to use it, recommend it to the peers and will continue to use it in the future. However, students gave the comparatively lowest score to the "effectiveness of online formative assessment" factor (Mean=3.88), indicating the need to improve the effectiveness of online formative assessment activities for improving their second language learning skills. In sum, this study can conclude that the effectiveness of system is accepted by most of the students but further efforts should be made to improve the effectiveness of online formative assessment activities.

3.3 The relations between students' attitude toward the course and their perceptions of using the system

In order to find the relationships between the above two surveys, the Pearson's correlation was conducted. As shown in Table 1, the results of the two questionnaires are highly related.

Table 1: The correlations among the factors of ATTC and POTS

	Usefulness	Ease of use	Satisfaction	Assessment effectiveness	Willingness
Attitude	0.59***	0.44***	0.42***	0.46***	0.52***
Improving skills	0.66***	0.39***	0.43***	0.49***	0.42***
Task completeness	0.49***	0.48***	0.50***	0.44***	0.41***
Extra effort	0.40***	0.26***	0.27***	0.47***	0.32***
Teacher assessment	0.35***	0.37***	0.42***	0.28***	0.33***
Demonstration	0.32***	0.25***	0.32***	0.39***	0.34***

Notes: ***$p<.001$

The regression analysis was performed to evaluate the predictive effects of the POTS factors on each factor of ATTC. As in Table 2, all the POTS factors are positive predictors to their corresponding ATTC factors. In other words, students with higher level of POTS tend to possess higher-level of ATTCs. Among all the POTS factors, "usefulness" could make positive prediction for all factors of students' attitude toward the EFL course and related course activities (t = 9.72, 14.34, 6.04, 4.53, 3.33, 2.42 respectively). In addition, the POTS factor "willingness" could also make significant prediction for all ATTC factors except "task completeness" (t = 6.96, 3.10, 2.20, 2.07, 4.32 respectively). Hence, we consider POTS factors "usefulness" and "willingness" as two important predictors to ATTC factors.

Table 2 Stepwise regression model of predicting ATTC (n=653)

ATTC	Predictors	B	S.E.		T	R^2
Attitude	Usefulness	0.41	0.04	0.36	9.72***	0.45
	Willingness	0.23	0.03	0.25	6.96***	
	Ease of use	0.15	0.04	0.13	3.54***	
	Assessment effectiveness	0.12	0.04	0.11	3.12**	
	Constant	0.11	0.18		0.63	
Improving skills	Usefulness	0.48	0.03	0.51	14.34***	0.49
	Assessment effectiveness	0.13	0.03	0.15	4.34***	
	Willingness	0.08	0.03	0.11	3.10**	
	Ease of use	0.07	0.03	0.07	2.11*	
	Constant	1.02	0.14		7.11***	
Task completeness	Satisfaction	0.20	0.04	0.23	5.36***	0.37
	Usefulness	0.21	0.04	0.23	6.04***	
	Ease of use	0.19	0.04	0.19	4.89***	
	Assessment effectiveness	0.10	0.03	0.12	3.16**	
	Constant	1.26	0.15		8.30***	
Extra effort	Assessment effectiveness	0.41	0.05	0.34	8.33***	0.27
	Usefulness	0.24	0.05	0.19	4.53***	
	Willingness	0.09	0.04	0.09	2.20*	
	Constant	0.26	0.21		1.26	
Teacher assessment	Satisfaction	0.17	0.04	0.23	4.78***	0.22
	Usefulness	0.11	0.03	0.14	3.33**	
	Ease of use	0.13	0.04	0.15	3.33**	
	Willingness	0.06	0.03	0.09	2.07*	
	Constant	2.64	0.15		18.01***	
Demonstration	Assessment effectiveness	0.27	0.04	0.27	6.20***	0.20
	Willingness	0.16	0.04	0.18	4.32***	
	Usefulness	0.12	0.05	0.11	2.42*	
	Constant	1.90	0.18		10.49***	

Notes: *p<.05, **p<.01, ***p<.001

4. Conclusion

This paper reports on the development of an Online Formative Assessment System in an EFL course and examines the effectiveness of the course and the system. The research results indicate that the EFL course adopting formative assessment in online learning environment is well received by students. It is also evident that most students find it rewarding to use the system. Besides, students also expressed their preference to the system in terms of its usefulness, ease of use, and their willingness for future use. The regression results reveal that students' perceptions of the system could make significant prediction for their responses of attitude towards the course. In particular, POTS factors of "willingness" and "usefulness" were significant variables for predicting students' attitude toward the course. However, it should be noted that the present study mainly employed quantitative measures. More in-depth qualitative studies are recommended to explore the issue further.

Acknowledgements

This paper is supported by the Beijing Higher Education Young Elite Teacher Project (YETP0463) and the Fundamental Research Funds for the Central Universities (2013RC0702), (2013RC0708).

References

Chen, Q. X., Kettle, M., Klenowski, V., & May, L. (2013). Interpretations of formative assessment in the teaching of English at two Chinese universities: a sociocultural perspective. *Assessment & Evaluation in Higher Education*, 38(7), 831-846.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *Mis Quarterly, 13*(3), 319-340.

Fang, Y. C. (2010). Perceptions of the Computer-Assisted Writing Program among EFL College Learners. *Educational Technology & Society, 13*(3), 246-256.

Li, Y. Y., Dong, M. K., & Huang, R. H. (2011). Designing Collaborative E-Learning Environments based upon Semantic Wiki: From Design Models to Application Scenarios. *Educational Technology & Society, 14*(4), 49-63.

Ministry of Education, (2007). College English Curriculum Requirements. Beijing: Foreign Lanugage Teaching and Research Press.

Pachler, N., Daly, C., Mor, Y., & Mellar, H. (2010). Formative e-assessment: Practitioner cases. *Computers & Education, 54*(3), 715-721.

Sorensen, E. K., & Takle, E. S. (2005). Investigating knowledge building dialogues in networked communities of practice: A collaborative learning endeavor across cultures. *Interactive Educational Multimedia*, 10, 50–60.

Wang, T. H., Wang, K. H., & Huang, S. C. (2008). Designing a Web-based assessment environment for improving pre-service teacher assessment literacy. *Computers & Education*, *51*(1), 448-462.

Wen, Q. F. (2011). Formative assessment in the course "reading and evaluating research papers": theories and practice. *Foreign Language Testing and Teaching*, 3, 39-49.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Role-Play in Computer-Supported Collaborative Learning-An Explorative Study

Yu-Chen HSU[a] & Yen-Lin CHIU[b*]
[a]*School of Public Administration, Sichuan University, China*
[b]*Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan*
*E-mail: erin0825@ms58.hinet.net

Abstract: The role-play has been regarded as an important function which may facilitate effective computer-supported collaborative learning (CSCL). In this paper, we explored undergraduates' experiences in undertaking team-based task in CSCL environment. The main purpose of this study was to examine the impacts of role-play on CSCL. 90 participants grouped into 18 teams were surveyed. For the team-based analysis level, the results indicated that roles of team members in CSCL teams have relations with their appraisal for CSCL. This finding suggested that teachers and students should be aware of the roles played in CSCL environment and govern the roles to stimulate effective online collaborative learning.

Keywords: collaborative learning, CSCL, performance, role-play, satisfaction

1. Introduction

Nowadays collaboration has been extremely concerned and adopted in learning activities. Collaborative learning is regarded as students' interaction while they are making efforts to solve problems and accomplish tasks together in learning process (Dewiyanti et al., 2007). In recent years computer supported collaborative learning (CSCL) is viewed as a good way to facilitate knowledge acquisition and to improve learning in online environments (Noroozi et al., 2013; Wecker et al., 2014).

However, it was indicated that putting students together does not necessarily result in effective outcomes of collaborative learning (Weinberger et al. 2005). On the one hand, it was pointed out that assigning roles to students in collaborative learning process has positive effects on learning outcomes (De Wever et al., 2008). Roles can be defined as learners' responsibilities which may guide individual behavior and govern group interaction to achieve group's goal (Strijbos & De Laat, 2010). The role concept has attracted increasing attention and become a promising construct for facilitating CSCL (Strijbos & De Laat, 2010; Pozzi, 2011). However, the impact of role-play on CSCL is still an emerging topic which has not been completely studied. This study aimed to explore team members' roles in CSCL teams and its impacts on team members' appraisal for CSCL.

2. Methods

2.1 Participants

All participants were undergraduates of Sichuan University in China. The average age of them was 21.52 with a range from 20 to 24. The ratios of gender type were almost equal. Most of the participants majored in public administration and public policy. They were grouped into 18 teams to collaboratively accomplish a team-based task which was part of their assignments related to their study topic. The number of team members ranged from 3 to 7 with an average of 5 members. The roles they played while undertaking team-based task in CSCL environment were conductor (16.7%), information provider (15.6%), active actor (54.4) and general actor (13.3%).

Table 1: Participants' demographics.

Demographics		
Variables	Frequency	%
Gender		
Female	48	53.3
Male	42	46.7
Major		
Public administration and Public Policy	34	37.8
Land Resource and Real Estate Management	9	10.0
Social Security and Insurance	13	14.4
Philosophy	3	3.3
Secretary and Archive Science	9	10.0
Information management technology	15	16.7
Information Resource Management	7	7.8
Country level of hometown		
Level 1	5	5.6
Level 2	17	18.9
Level 3	23	25.6
Village	45	50.0
Role in CSCL team		
Conductor	15	16.7
Information provider	14	15.6
Active actor	49	54.4
General actor	12	13.3

2.2 Instruments

In addition to the demographic variables, a self-report instrument for appraising participants' experiences of CSCL including attitude toward CSCL, satisfaction with CSCL method and satisfaction with CSCL outcome were evaluated. Moreover, the Internet Self-efficacy Scale was also utilized and measured.

2.3 Analysis Procedure

The descriptive statistics were analyzed to explore the demographics of the participants. The exploratory factor analysis was conducted to validate the instruments including CSCL experience survey and Internet self-efficacy scale. Finally, the hypotheses were tested by executing t-test and regression analysis.

3. Results

3.1 Results of exploratory factor analysis

For evaluating the validation of instruments the exploratory factor analyses were conducted. The Kaiser-Meyer-Olkin (KMO) measure and the Bartlett's test of sphericity were examined to determine whether the sample was appropriate for executing the EFA.

For the Internet Self-Efficacy Scale (ISES), it was reported that the KMO measure had a value of 0.89 with a significant Bartlett's test (chi-square = 624.30, $p < 0.001$) showing the EFA was appropriate. As a result, the items were grouped into 2 factors, namely Basic Internet Self-Efficacy (BISE) and Advanced Internet Self-Efficacy (AISE). The Cronbach's alpha for two factors were 0.77and 0.85, suggesting that two factors had high reliability. Both BISE and AISE contained 5 items and the total variance explained reached 74.17%, implying the ISES was appropriate for assessing the participant' Internet-based self-efficacy. With respect to the CSCL experience, the EFA results revealed that the KMO measure (0.84) and Bartlett's test (chi-square = 116.25, $p < 0.001$) were well examined. 65.87 % of total variances were explained by two factors namely overall appraisal for method (3 items) and overall appraisal for outcome (3 items) with alpha values of 0.69 and 0.68, respectively.

3.2 Comparisons of gender difference

Table 1 shows the differences between genders. It reveals that males have higher frequency in computer usage and Internet usage than females does. However, there are no differences of appraisal for CSCL method and appraisal for CSCL outcome between males and females.

Table 1: Differences between genders.

Variables	Male		Female		t -test	p value
	Mean	S.D.	Mean	S.D.		
age	21.71	0.97	21.35	0.86	-1.86	0.066
computer usage (hr/week)	35.83	25.23	23.00	12.81	-3.06^{**}	0.003
Internet usage (hr/week)	28.93	25.05	16.79	11.52	-2.98^{**}	0.004
Satisfaction with performance	3.36	1.12	2.92	1.22	-1.78	0.079
Satisfaction with communication	4.40	0.66	4.10	0.83	-1.88	0.064
Attitude toward CSCL	3.71	0.83	3.48	0.92	-1.26	0.211
Attitude toward cooperation	4.90	0.45	4.85	0.54	-0.46	0.649
Satisfaction with cooperation	4.76	0.69	4.66	1.03	-0.54	0.588
Attitude toward online discussion	4.24	0.98	4.07	1.01	-0.78	0.440
Basic Internet self-efficacy	5.64	0.46	5.59	0.48	0.47	0.643
Advanced Internet self-efficacy	5.51	0.72	5.02	0.99	2.64^{*}	0.010
Overall appraisal for CSCL method	4.33	0.56	4.21	0.69	-0.84	0.406
Overall appraisal for CSCL outcome	4.17	0.58	3.89	0.87	-1.78	0.079

Note: * $p < 0.05$; ** $p < 0.01$

3.3 Correlations among variables for individuals

As shown in table 2, males have higher computer usage and Internet usage than females; however, males possess higher Advanced Internet self-efficacy than their counterparts. Moreover, it reveals that computer usage has positive relations with attitude toward CSCL, suggesting that raising learners' computer usage may increase their positive attitude toward CSCL.

3.4 Correlations among variables for CSCL teams

Table 3 shows the correlations among variables for team level. It should be noted that all variables were calculated for team level; for instance, the roles (conductor, information provider, active actor and general actor) were counted as ratios in a CSCL team.

As revealed in table 3, CSCL teams have higher satisfaction with communication while they have higher ratio of information provider ($r = 0.57$, $p < 0.05$), suggesting that arrange more information provider in a CSCL team may increase team members' satisfaction with their communication. However, the teams reveal negative attitude toward CSCL ($r = -0.51$, $p < 0.05$) if the CSCL teams have more general actors in their team; moreover, the teams with higher ratio of general actor tend to possess lower satisfaction with communication ($r = -0.50$, $p < 0.05$), online discussion ($r = -0.48$, $p < 0.05$) as well as overall appraisal for CSCL outcome ($r = -0.50$, $p < 0.05$). Moreover, the teams have more positive attitude toward CSCL ($r = 0.49$, $p < 0.05$), online communication ($r = 0.71$, $p < 0.01$) and appraisal for CSCL method ($r = 0.62$, $p < 0.01$) while these teams have higher average computer usage. Finally, it is indicated that Basic Internet self-efficacy has positive relationships between attitude toward CSCL ($r = 0.47$, $p < 0.05$) and online discussion ($r = 0.54$, $p < 0.05$), showing that improving teams' basic Internet self-efficacy may increase their attitude toward CSCL.

Table 2: Correlations among variables for individuals.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Gender (male: 1; female: 0)	------												
2. Age	0.19												
3. Computer usage (hr/week)	0.31^{**}	0.12											
4. Internet usage (hr/week)	0.31^{**}	0.14	0.94^{***}										
5. Satisfaction with performance	0.19	0.18	0.09	0.06									
6. Satisfaction with communication	0.20	-0.04	0.05	0.04	0.48^{***}								
7. Attitude toward CSCL	0.13	-0.04	0.22^{*}	0.20	0.44^{***}	0.51^{***}							
8. Attitude toward cooperation	0.05	-0.19	-0.11	-0.16	0.24^{*}	0.38^{**}	0.35^{**}						
9. Satisfaction with cooperation	0.06	-0.05	-0.18	-0.20	0.38^{***}	0.47^{***}	0.35^{**}	0.49^{***}					
10. Attitude toward online discussion	0.09	-0.16	0.12	0.03	0.33^{**}	0.37^{**}	0.58^{***}	0.32^{**}	0.20				
11. Overall appraisal for CSCL method#1	0.10	-0.07	0.19	0.14	0.39^{**}	0.56^{***}	0.86^{***}	0.61^{***}	0.48^{***}	0.86^{***}			
12. Overall appraisal for CSCL outcome#2	0.19	0.07	0.00	-0.03	0.84^{***}	0.79^{***}	0.55^{***}	0.45^{***}	0.75^{***}	0.38^{***}	0.58^{***}		
13. Basic Internet self-efficacy	0.09	-0.21^{*}	-0.00	-0.02	0.08	0.08	0.16	0.20	0.18	0.14	0.13	0.14	
14. Advanced Internet self-efficacy	0.26^{*}	-0.15	0.12	0.11	0.11	0.14	0.17	0.18	0.10	0.04	0.05	0.14	0.77^{***}

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$
#1 Overall appraisal for CSCL method = (attitude toward CSCL + attitude toward cooperation + attitude toward online discussion)/3
#2 Overall appraisal for CSCL outcome = (Satisfaction with performance+ Satisfaction with communication+ Satisfaction with cooperation)/3

Table3: Correlations among variables for CSCL teams

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Gender (male: 1; female: 0)	----																
2. Age	0.28																
3. Conductor[#1]	0.14	0.13															
4. Information provider[#2]	0.16	-0.15	-0.11														
5. Active actor[#3]	-0.11	-0.04	-0.17	-0.67^{**}													
6. General actor[#4]	-0.14	0.04	-0.32	0.02	-0.59^{*}												
7. Computer usage (hr/week)	0.13	0.04	-0.17	0.47^{*}	-0.38	0.16											
8. Internet usage (hr/week)	0.21	0.07	-0.27	0.47^{*}	-0.21	0.00	0.94^{***}										
9. Satisfaction with performance	0.39	-0.07	0.00	0.24	0.06	-0.37	0.26	0.28									
10. Satisfaction with communication	0.45	-0.17	0.15	0.57^{*}	-0.14	-0.50^{*}	0.24	0.31	0.73^{**}								
11. Attitude toward CSCL	0.40	-0.34	0.04	0.33	0.11	-0.51^{*}	0.49^{*}	0.55^{*}	0.55^{*}	0.71^{**}							
12. Attitude toward cooperation	0.08	-0.42	0.03	0.21	0.09	-0.37	0.27	0.20	0.41	0.51^{*}	0.70^{**}						
13. Satisfaction with cooperation	0.22	-0.36	-0.19	0.26	0.25	-0.48^{*}	-0.10	0.01	0.43	0.57^{*}	0.54^{*}	0.55^{*}					
14. Attitude toward online discussion	0.38	-0.22	0.08	0.38	-0.18	-0.20	0.71^{**}	0.70^{**}	0.58^{*}	0.65^{**}	0.79^{**}	0.40	0.18				
15. Basic Internet self-efficacy	0.04	-0.61^{**}	-0.19	-0.02	0.16	-0.09	0.07	0.07	0.30	0.36	0.54^{*}	0.46	0.31	0.47^{*}			
16. Advanced Internet self-efficacy	0.13	-0.68^{**}	-0.12	0.18	-0.17	0.13	0.10	0.05	0.29	0.42	0.46	0.43	0.28	0.45	0.91^{***}		
17. Overall appraisal for CSCL method	0.38	-0.27	0.08	0.31	0.01	-0.39	0.62^{**}	0.62^{**}	0.47^{*}	0.63^{**}	0.93^{**}	0.65^{**}	0.39	0.90^{***}	-0.47	-0.38	
18. Overall appraisal for CSCL outcome	0.43	-0.19	-0.01	0.39	0.06	-0.50^{*}	0.20	0.27	0.92^{***}	0.89^{***}	0.70^{**}	0.55^{*}	0.70^{**}	0.59^{*}	-0.38	-0.38	0.58^{*}

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

[#1] Ratio of conductor in CSCL team; [#2] Ratio of information provider in CSCL team; [#3] Ratio of active actor in CSCL team; [#4] Ratio of general actor in CSCL team

4. Conclusions

This study indicated that team members' roles in a CSCL team have relations with their appraisal for CSCL. This finding suggested that teachers and students should be aware of the roles played in CSCL environment and govern the roles to stimulate effective online collaborative learning.

.

Acknowledgements

This study was supported by the Ministry of Science and Technology, Taiwan, under grant number MOST 103-2511-S-011 -008 -MY3.

References

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2009). Structuring asynchronous discussion groups: the impact of role assignment and self-assessment on students' levels of knowledge construction through social negotiation. *Journal of Computer Assisted Learning, 25*, 177-188.

Dewiyanti, S., Brand-Gruwel, S., Jochems, W., Broers, N. J. (2007). Students' experiences with collaborative learning in asynchronous computer-supported collaborative learning environments. *Computers in Human Behavior, 23*, 496-514.

Noroozi, O., Teasley, S.D., Biemans, H. J. A., Weinberger, A., & Mulder, M. (2013). Facilitating learning in multidisciplinary groups with transactive CSCL scripts. *International Journal of Computer-supported Collaborative Learning, 8*(2), 189-223.

Pozzi, F. (2011). The impact of scripted roles on online collaborative learning processes. *Computer-Supported Collaborative Learning, 6*, 471-484.

Strijbos, J. W., & De Laat, M. F. (2010). Developing the role concept for computer-supported collaborative learning: An explorative synthesis. *Computers in Human Behavior, 26*, 495-505.

Wecker, C., & Fischer, F. (2014). Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning. *Computers & Education, 75*, 218-228.

Weinberger A., Reiserer M., Ertl B., Fischer F., & Mandl H.(2005). Facilitating collaborative knowledge construction in computer-mediated learning environments with cooperation scripts. In R. Bromme, F.W. Hesse, & H. Spada (Eds), *Barriers and Biases in Computer-Mediated Knowledge Communication* (pp. 15-38). Kluwer, Boston.

Exploring the Interactive Use of Video Cases in Scaffolding Prospective Teachers in Learning Clinical Interview Method

Yu-Ling Hsu*

[a] *Graduate Institute of Digital Learning and Education,*

National Taiwan University of Science and Technology, Taiwan

*Yuling.Hsu.NY@gmail.com

Abstract: Numerous studies have shown that the clinical interview method has played an important role in helping educators and researchers to gain deep insight into children's mathematical thinking, and thereby improve their classroom teaching. However, very little research has been done on the development and design of such training. Therefore, the purpose of this study was 1) to explore the possible benefits of using a new approach to teach clinical interviewing skills, and 2) to investigate how an expert commentary feature can help prospective teachers to learn clinical interview methods. Forty prospective teachers participated in this study and completed a series of carefully designed lessons involving video case analysis with expert commentary. A mixed of qualitative and quantitative method was used to analyzed all the data. The results showed that the prospective teachers found this new approach helpful in learning to conduct clinical interviews.

Keywords: clinical interview, video analysis, case-based learning, early childhood education, mathematics education, teacher education, instructional technology

1. Introduction

Clinical interview has played an important and powerful role in helping pre-service and in-service teachers to gain deep insight into their students' mathematical thinking, and has thereby enabled teachers to assist their students in constructing meaningful mathematical knowledge and enhancing their mathematical learning potential (Baroody & Ginsburg, 1990; Buschman, 2001; Confrey, 1980; Doig & Hunting, 1995; Ginsburg 1981; Ginsburg, 1997; Ginsburg, Jacobs & Lopez, 1998; Ginsburg, Kaplan, & Baroody, 1992; Ginsburg, Kossan, Schwartz & Swanson, 1983; Hunting, 1997; Kaplan, King, Dickens, & Stanley, 2000; McDonough, Clarke & Clarke, 2002; Peck, Jencks, & Connell,1989; Rowland, 1999; Schorr & Ginsburg, 2000; Schorr & Lesh, 1998; Wright & Ellemor-Collins, 2008; Zazkis & Hazzan, 1999.)

Although there is growing evidence of the need to teach clinical interviewing techniques to both pre-service and in-service teachers, very little research has been done on the development and design of such training. With the use of videos, case-based discussion has been found to be a useful tool for both engaging pre-service teachers in observing, noticing, interpreting, discussing, reflecting, and transforming ideas about a complex or ill-structured teaching situation and capturing such change (Hatch & Grossman, 2009; Kinzer & Risko, 1998; Sherin & Han, 2004; Sherin & Van Es, 2005; Van Es & Sherin, 2002). Yet studies documenting the ways and means of teaching and training clinical interview methods are still very limited, and there is no literature on how the instructional design of a video case-based learning environment could help pre-service or in-service teachers to develop clinical competency of this type.

2. Methods

2.1 Participants

Participants in this study were 40 graduate students from Teachers College, Columbia University enrolled in "Development of Mathematical Thinking" course. The students in this course were pre-service teachers pursuing Master degrees in early childhood or mathematics education.

2.2 Measures

The task (which is also referred as clinical interview expert commentary video lesson) used in this study was embedded in a Web-based application called VITAL (Video Interactions for Teaching and Learning), which was designed to support prospective teachers in studying early childhood mathematical education. Therefore, this study used 3 different kinds of data: (a) students' analyses of videos and comments on the lesson (which will refer to pre and post-commentary answers later on); (b) students' answers to the survey questions; and (c) students' clinical interview final project scores from the class.

2.3 Analysis Procedure

A mix of qualitative and quantitative data analyses were used in this study. The development of the coding schemes began with an extensive content analysis of the pre and post commentary answers as well as the open-ended answers in the end-of lesson survey, for a total of 400 video analysis answers. Eight sets of coding schemes were developed to analyze the participants' responses. There are three major categories: a) coding schemes for all pre-commentary responses, b) coding schemes for all post-commentary responses, and c) coding schemes. All the responses were coded by the author and another independent researcher. The kappa value for the inter-rater reliability of all the pre-commentary questions is k=0.80 and of all the post-commentary questions is k=0.81.

3. Results

3.1 Helpfulness of the Expert Commentary Video

Table 1. Helpfulness of the Expert Commentary Video

Rating scale	Frequency (counts)	Percentage
Very Helpful	20	50%
Helpful	16	40%
Somewhat Helpful	4	10%
Total	40	100%

The data was obtained at the end of the video lesson, when the participants were asked to give a four level rating (very helpful, helpful, somewhat helpful, and not helpful). Overall, all the participants (N=40) in this study reported that this video case analysis lesson with expert commentary, was helpful in varying degrees in learning clinical interview. Specifically, 50% of them rated it as "very helpful", 40% said it was "helpful" and only 10% reported it "somewhat helpful". "Not helpful" was not found (see Table 1).

3.2 Video Case Analysis Comparison Within Pre-Commentary Questions

Table 2. Paired T-Test Results For the Two Pre-Post Question Sets Within the Pre Commentary Questions

Question set	Mean	SD	df	T
Q1 vs. Q3	-3.00	.853	39	-2.223**
Q7 vs. Q9	-3.50	1.027	39	-2.156**

** p<.02 (1-tailed)

Since Q1 and Q3 need participants to apply similar concepts to evaluating clinical interviewing techniques, and Q7 and Q9 require similar clinical reasoning methods for the case, we were able to use Q1 and Q3 as a pre-post comparison set, and Q7 & Q9 as another set. As shown in Table 2, the results suggest that there was a significant difference in these prospective teachers' pre and post video analysis levels for both question sets (Q1& Q3, t=-2.223, p<.02 and Q7 & Q9, t=-2.156, p<.02). Therefore, the prospective teachers did show significant differences in their video case analysis in these two pre-post question sets within the pre-commentary questions (Q1 vs. Q3 & Q7 vs. Q9). More details about what factors could contribute to this difference will be discussed later on.

3.3 Video Case Analysis Level on All Pre-Commentary Questions

Table 3. Frequency and Percentage of Video Analysis Level on All Pre-Commentary Questions (N=40)

	Q1	Q3	Q5	Q7	Q9
Level 1	6 (15%)	2 (5%)	6 (15%)	11 (27%)	3 (7%)
Level 2	16 (40%)	12 (30%)	33 (83%)	14 (35%)	16 (40%)
Level 3	18 (45%)	26 (65%)	1 (2%)	15 (38%)	21 (53%)
Total	40 (100%)	40(100%)	40(100%)	40(100%)	40(100%)

The frequencies and the percentages of each video analysis level for individual questions are shown in Table 3. As we can see, Q3 has the highest frequencies (and percentages) in Level 3 among all the pre-commentary questions; followed by Q9. Most of the prospective teachers did not do so well in Q5, but did better in Q3 and Q9. The frequencies (and percentages) in Q1 and Q7 are spread out among all three levels, though the Level 2 and Level 3 have a slightly higher percentage in Q1 than in Q7.

Acknowledgements

I would like to thank Prof. Herbert Ginsburg and Prof. Charles Kinzer for their advising on this study, and also for the funding support from National Science Foundation in USA , the Ministry of Education in Taiwan, and Prof. Chin-Chung Tsai from National Taiwan University of Science and Technology.

References

Baroody, A. J., & Ginsburg, H. P. (1990). Children's mathematical learning: A cognitive view. In R. B. Davis, C. A. Maher, and N. Noddings, (Eds.), *Journal for Research in Mathematics Education* (Vol. Monograph No. 4, pp. 51-210). Reston, VA: National Council of Teachers of Mathematics.

Buschman, L. (2001). Using student interviews to guide classroom instruction. *Teaching Children Mathematics*, *8(4)*, 222-227.

Confrey, J. (1980). *Clinical interviewing: Its potential to reveal insights in mathematics education.*

Paper presented at the Proceedings of the 4th International Conferences for PME Berkeley, CA.

Doig, B., & Hunting, R. (1995). *Preparing teacher-clinicians in mathematics.* Paper presented at the Proceedings of the 19th conference of the international group for the psychology of mathematics education. Recife, Brazil.

Ginsburg, H. P. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the Learning of Mathematics, 1*(3), 4-11.

Ginsburg, H. P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. New York: Cambridge University Press.

Ginsburg, H. P., Jacobs, S. F., & Lopez, L. S. (1998). *The teachers' guide to flexible interviewing in the classroom: learning what children know about math*. Needham Heights, MA: A Viacom Company.

Ginsburg, H. P., Kaplan, R. G., & Baroody, A. J. (1992). *Children's mathematical thinking: Videotape workshops for educators*. Evanston, IL: Everyday Learning Corporation.

Ginsburg, H. P., Kossan, N.E., Schwartz, R., & Swanson, D. (1983). Protocol methods in research on mathematical thinking. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 7-47). New York: Academic Press.

Hatch, T., & Grossman, P. (2009). Learning to look beyond the boundaries of representation. *Journal of Teacher Education. 60* (1*)*, 70-85.

Hunting, R. P. (1997). Clinical interview methods in mathematics education research and practice. *Journal of Mathematical Behavior, 16*(2), 145-165.

Kaplan, R. G., King, B., Dickens, N., & Stanley, V. (2000). Teacher –clinicians encourage children to think as mathematicians. *Teaching Children Mathematics,* (February), 406-411.

Kinzer, C., & Risko, V. (1998). Multimedia and enhanced learning: Transforming preservice education. In D. Reirlking, M. McKenna, L. Labbo, & R. Kieffer (Eds.), *Handbook of technology and literacy: Transformation in a post typographic world* (pp. 185-202). Hillsdale, NJ: Lawrence Erlbaum.

McDonough, A., Clarke, B., & Clarke, D. M. (2002). Understanding, assessing and developing children's mathematical thinking: the power of a one-to-one interview for preservice teachers in providing insights into appropriate pedagogical practices. *International Journal of Educational Research, 37,* 211-226.

Rowland, T. (1999). *The clinical interview: Conduct and interpretation.* Paper presented at the 23rd conference of the International group for the psychology of mathematics education. Haifa, Israel: Israel Institute of Technology.

Schorr, R. Y., & Ginsburg, H. P. (2000). Using clinical interviews to promote preservice teachers' understanding of children's mathematical thinking. In M. L. Fernandez (Ed.), Proceedings of the 22nd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp. 599–605). Columbus, OH: ERIC Clearinghouse for Science Mathematics and Environmental Education.

Sherin, M. G., & Van Es, E.A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*(3), 475-491.

Sherin, M. G., & Han, S. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education, 20,* 163-183.

Van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education, 10*(4), 571-596.

Wright, R. J., & Ellemor-Collins, D. L. (2008). Assessing student thinking about arithmetic: Videotaped interviews. *Teaching Children Mathematics*, *15*(2), 106-111.

Zazkis, R., & Hazzan, O. (1999). Interviewing in mathematics education research: Choosing the questions. Journal of Mathematical Behavior, 17(4), 429-439

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Effectiveness of Reducing State Anxiety by Digital Counseling Tool - Mind Collage

Yu-Jen Hsu [*], Ju-Ling Shih
[a] *Department of E-learning Design and Management, National University of Chiayi, Taiwan*
[b] *Department of Information and Learning Technology, National University of Tainan, Taiwan*
*yuren925@gmail.com

Abstract: This research aims to investigate the effectiveness of digital counseling tool, Mind Collage, to reduce participants' state anxiety. Mind Collage is a digital tool designed based on expressive art therapy, Landscape Montage, to project users' inner self image onto the paintings they made. STAT-S inventory was used to compare the participants' change of state anxiety. From behavioral observations and interviews, the researcher found that Mind Collage can provide a relaxing space for the participants, and effectively reduce their state anxiety. At the same time, they highly trust the tool, and are more willing to reveal themselves on the digital system.

Keywords: Digital Counseling Tool, State Anxiety, Mind Collage

1. Introduction

In the modern society, human beings are living in a much faster pace, receiving more information, facing more challenges, and feeling more pressures. As a natural consequence, invisible anxieties exist deep inside each individual. From the counseling perspective, expressive art therapy is a way to restore emotions, and thus reduce anxiety.

In order to approach the individuals with ease, this study strive to seek a commonly used tool to do the work. Reviewing the nature and features of technology, from the 20th century on, Internet and information technology have brought about the formation of virtual world which intrigued thousands of young people to be its loyal citizens. The virtual environment has greatly influenced people's living and learning style since it has opened the possibility of linking people with others and their environment.

Counseling is a professional work that counselors aim to assist their clients to effectively deal with their problems with effective ways (Egan, 2007). It is based on assisting the clients' to develop into mentally healthy individuals who can work independently and self-guidedly. Ttraditional counseling model provides personal psychological assistance to the client face-to-face, but it often discourages some young people to seek for help due to the mental confrontation of revealing oneself to strangers. Therefore, many counseling groups have begun to provide Internet counseling services, and have led counseling into a new era.

Existing digital counseling tools are becoming more diverse. Many counseling groups have begun to provide Internet counseling services, but limited to mechanisms such as message boards, E-mail, online psychological tests, real-time chat, and so on. Related studies showed that the Internet has advantages such as anonymity, virtuality, convenience, escape, and so on. Therefore, web-counseling can provide clients a high degree of freedom and privacy so that they can directly express their feelings; at the same time, the counselors can create better counseling relationship with the client, and easily record and manage the process of treatment.

With the advantages of counseling in the digital forms, the purpose of this study is to use a self-developed digital counseling tool, Mind Collage, to reduce clients' anxiety. The development of Mind Collage integrated digital technology and counseling theories, with two special features. First, digital expression of mental images. Users project their mental images onto the pictures, which is a non-verbal communication medium. Second, mental images as treatment. As the users conduct the painting, they enter a space where they face themselves, without external pressure or burden. They have little emotional confrontations, and enhance perceptions and awareness. Both features strengthen Mind Collage as the bridge to connect the users' inner self with the outside world.

2. Literature Review

2.1 Anxiety

Anxiety can be categorized as state anxiety and trait anxiety in terms of its nature (Spielberger, 1971). Trait anxiety refers to the anxiety that comes from inside of an individual. It is a long-lasting and fixed personality that no matter what circumstances the person is in, trait anxiety would stay. On the other hand, state anxiety refers the emotional responses when the individual faces certain situations, for example, work pressure; however, as long as the situation does not last, the anxiety disappears as well. Thus, state anxiety has temporary and transitional characteristics (Pintrich & Schunk, 2002).

Anxiety makes people feel tension and nervous. It comes from intangible and unobvious sources and threatens. It is a defensive emotion that helps people deal with the environment, and stay away from danger; at the same time, it transmit communicative message of acquiring help and protection. It stimulates the individual's willingness to take action, and do things better. Nevertheless, it influences the work of cognitive function. Pekrun (2000) pointed out that emotion has three major sources, namely genetic dispositions, physiological processes, and cognitive appraisals. The most commonly seen emotions seen in the psychological therapy include anger, sadness and distress, fear and anxiety, shame, as well as the pleasant emotions. Anxiety is generated from the uncertainty and fragile inner self, which make one feels repressed from self-expectation, afraid of rejection and failure. One would be influenced by the stereotype of the society, and experience anxiety through anger or sadness.

Expressive art therapy can reduce the anxiety that is generated from the first encounter of counseling. Expressive art therapy also helps the participants to establish relationships with the counselors, and be submerged in the counseling process. Since mental image projection create a psychological safe distance of the participant and the issue, the participants can reveal their subconscious feelings and thoughts from the project and make it a part of their self-perception.

Ellis (1998) thought emotion is the result of complex interaction of human behavior responses and perceptions; therefore, emotion is a kind of cognitive perception. There are three levels of emotion expression guidance (Greenberg, 2006). The first level is Dialogue Level which refers to the process of using simply language with experience-focused questions to guide the clients to express their emotions. Most traditional counseling stay in this level. The second level is Processes-oriented Level in which counselors direct the client's emotional focus onto a specific experience or behavior, and increase his awareness to the aroused emotions. The last level is Stimulate Level which is to create a new experience through activities for the client to generate new emotions.

From this point of view, the participants' emotion responses monitored in this research involve those of process-oriented level and stimulation level. Therefore, the aim of this research is to use a digital tool to guide them through specific emotions, and even through the activities to create more peaceful emotions.

2.2 Expressive Art Therapy

Art activities can provide a concrete and non-verbal medium for people to deliver their subconscious thoughts and to raise their awareness. It can be the drive for people to make therapeutic changes (Dalley, 1984). Because that we are very isolated, inhuman, and over-intellectualized that it is rather important to increase the connection with their inner selves (Moreno, 1975). Therefore, non-verbal counseling methods such as art therapy are used more in the counseling arena in the non-verbal domain. Expressive art therapy is a medium of expression, which can present visual image from sub-consciousness. In this way, people can display outside their hidden thoughts and affections. thinks that the expressive art therapy emphasizes players' creative process. This treatment assumed that the individual's inner reality of can be presented by creation. Therefore, the expression of the creative process makes the players experience expanding human experiences, self-balancing, whole life, and increasing self-awareness, and get feelings from inner world.

Expressive art therapy is a medium of expression, which can present visual image from sub-consciousness. In this way, people can display outside their hidden thoughts and affections. Egan (2007) thinks that the expressive art therapy emphasizes players' creative process. This treatment assumed that the individual's inner reality of can be presented by creation. Therefore, the expression of the creative process makes the players experience expanding human experiences, self-balancing, whole life, and increasing self-awareness, and get feelings from inner world. As Withrow (2004) said, art therapy can more effectively reached the therapeutic goals than dialogue because it provides experiences exceeding language. Expressive art therapy allows the clients to recover from their negative emotions through images, and reached the deeper self.

Expressive art therapy is a medium of expression, which requires the individual to present his visual images from the sub-consciousness (Malchiodi, 2007). Landscape montage technique, also called landscape composition, was proposed by Mr. Nakai in 1969. It is one of the methods of expressive art therapies. There are 12 landscape objects, such as river, mountain, farmland, road, house, tree, people, flower, animal, stone, bridge, and sun. Each one of them has its counseling implications. (1) River: unconscious life progress, the flow of energy. (2) Mountain: focus, direction. (3) Farmland: integration of the picture. (4) Road: the shape, form, type, and so on. (5) House: personal perceptions and feelings of the environment. (6) Tree: process of self-growth. (7) People: sensitivity, maturity, flexibility. (8) Flower: vitality, hope, and life which means the subjects' implicit sense of life and affection. (9) Animal: The outside ways getting energy and life. The performance of inside energy and drives from individual's sub-consciousness. (10) Stone: awareness of self-enhancement and protection. (11) Bridge: contact and communication. (12) Sun: vitality or strength.

3. Digital Counseling Tool System Description

Mind Collage was built with Unity 3D game engine. Adobe PhotoShop was used to create and generate various graphic styles for the users to use. From the previous counseling experiences, Landscape Montage drawing is difficult to people with lower painting skills. They would worry about whether they can complete the picture as the counselor hoped. Since this research wanted to solve this problem, the digital system was designed to use easy-to-manipulate techniques. The features include, 1) there is graphical user interface (GUI) to increase user friendliness; 2) objects are provided to minimize users' worriness to painting fright; 3) drag-and-drop is enabled to allow minimal manipulation complexity; 4) expert knowledge database analysis to allow users to receive instant feedbacks.

Expert Knowledge Base refers to the database to generate counseling feedbacks. It is the brain of the system which was built with the Sequence of the 12 elements for Landscape Montage, Definitions of the elements' placements, and Content Experts' Interpretations.

During the process of working on Mind Collage, the participants were asked to draw a landscape following counselor's step-by-step instruction. (Figure 1). Users can choose their own picture frame and background. The sequence and definition of the 12 elements of Landscape Montage were defined according to Akira Kaito's book (2010) Landscape Montage: A Pictorial Psychological Therapy. Our digital counseling game modifies such a method and allows users to perform it on a platform. On the game menu, clients place 12 objects onto the canvas. They place the visual images by simply drag-and-drop the pictorial icons in the zoom-in/zoom-out mode in the virtual environment. The completed landscape can not be changed or modified after they complete the project. The 12 Landscape Montage elements were grouped into three clusters, including background group, center group, and foreground group. Then the system would generate automatic counseling responses according to the definitions of the kinds, styles, sizes, quantity, and placements as the user finishes the game. These steps will constitute the picture that can sufficiently represent users' inner thoughts and feelings. In this way, we can help players to fine their hidden selves and diagnose the situations they are in.

When the client finishes the picture, the counselor discusses with the client, and gives suggestions to the client in aspects of the sense of integration, richness, uniqueness, spatial depth, and so forth; and at the end, concludes the feedbacks with the interpretations of individual meanings (Figure 2).

Figure 1. Mind Collage Interface and Product Example

Figure 2. Counseling feedbacks of the system

4. Methods

This research invited twelve adults randomly chosen, 2 males and 10 females, aged 25 to 40 years old, to participate the experiment. Pretest and posttest of participants' state anxiety were conducted for comparison to understand the effectiveness of expressive art therapy. Instant feedbacks to the participants' Mind Collage final product were given. Desktop recordings were done for the analysis of participants' inner image projections. The experiment process is as Figure 3.

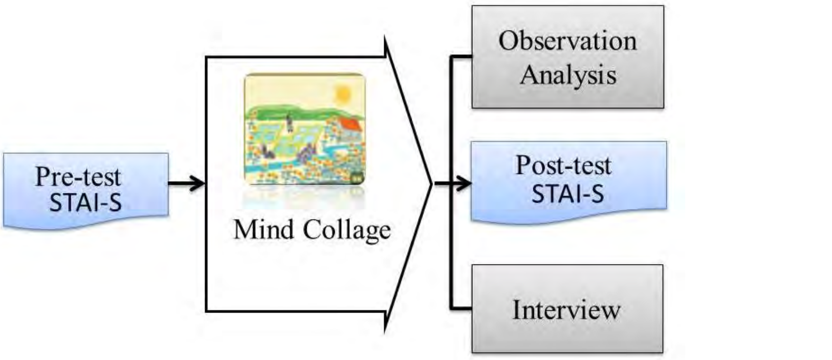


Figure 3. Experiment Process Flowchart

In order to investigate whether expressive art theory based digital counseling tool can allow the participants to calm their minds, be more peaceful, relaxed, and obtained positive action goals. In order to know the effects, the research tool used in this research is State-Trait Anxiety Inventory (STAI) proposed by Spielberger (1983). The inventory is appropriate for measuring the anxiety levels of teenagers to adults. The inventory involves two parts, State Anxiety and Trait Anxiety, with total of 20 questions. This research only uses state anxiety for measurement. State anxiety refers to the temporarily emotional state, which includes nervous, anxious, autonomous nerve system excitements, and conscious perceptions. Therefore, state anxiety would be different according to time and context. Questions in the inventory are mostly about current emotional feelings and self-perceptions. For example, I now feel peaceful; I feel I am confident; I am always alerted; etc. The point of Mind Collage is to allow participants to place their focus on making the pictures, and during this time, to calm down, lower down the defense, forget about the pressures from the outside world, look into their inner selves, and reduce state anxiety.

The scale uses four-point Likert with 4 points for strongly disagree, 3 points for disagree, 2 points for agree, and 1 point for strongly agree. The lowest score is 20, and highest score is 80. The lower the scores represent that the participants have more anxiety. The inventory has a retest reliability of 0.737 with Cronbach's α of 0.898.

The Mind Collage was done in a small room which had only the researcher and one participant and no external interferences. Every participant had 40 minutes of time. During the process, behavior observations were conducted to document participants' digital counseling process in order to find out the factors of participants' emotional changes. After the experiment, focus group interviews were made with the participants to know their use experiences, feelings, counseling effectiveness, and motivation changes. During the process, the researcher would place a mirror behind the participants which would reflect the actions of the computer desktops. In the front, there was a camcorder which recorded the participants' facial expressions. The setup is to conduct synchronous observations, document the working process. The experiment setup is as Figure 4.

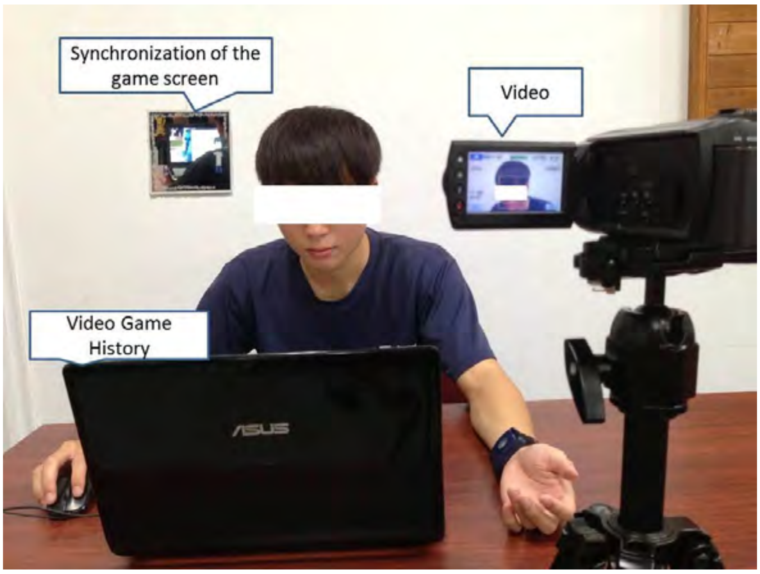


Figure 4. Experiment set up

5. Results

The occurrence and intensity of state anxiety is related to individual's subjective perception to the stimulation, and changes as the situation change. As the stimulation of state anxiety disappears, state anxiety would return to normal. STAT-S inventory shows the results of individual's responses to the certain environmental condition, and whether they feel safe, nervous, or anxious. The score of state anxiety of all twelve participants have risen after using Mind Collage (Table 1). All the posttests are higher than the pretests (Figure 5). It shows that this digital counseling tool has positive influence on the participants' state anxiety that their feeling of anxiety is reduced.

To further investigate the results, the participants can be categorized into two major types: those with obvious variation of state anxiety (Variation > 5) and those of stable variation of state anxiety (Variation < 5). Type I include 7 participants: W01, W02, W05, W06, W07, W08, W10; Type II include 5 participants: M01, W03, M02, W04, W09.

Table 1: STAI-S data

Subjects	Number	Gender	Pretest	Posttest	Variation	Variation
01	M01	M	58	59	Rise	+1
02	W01	F	49	65	Rise	+16
03	W02	F	57	66	Rise	+9
04	W03	F	68	72	Rise	+4
05	M02	M	60	61	Rise	+1
06	W04	F	74	76	Rise	+2
07	W05	F	45	56	Rise	+11
08	W06	F	56	71	Rise	+15
09	W07	F	65	73	Rise	+8
10	W08	F	56	67	Rise	+11
11	W09	F	69	70	Rise	+1
12	W10	F	43	50	Rise	+7

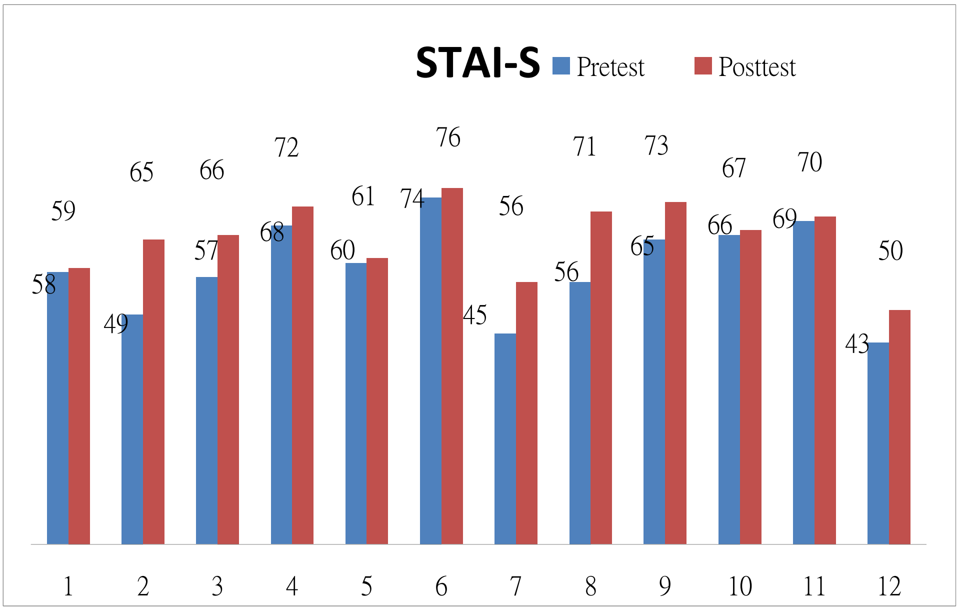


Figure 5. Graph of STAI-S Results

Type I: Obvious variation of state anxiety

A typical example of Type I is W02. The complete picture made by W02 in Mind Collage is as Figure 6. Her pretest score of STAT-S is 49 points, and posttest score is 65 points. It shows that Mind Collage has obvious effectiveness to reduce the participants' state anxiety. From behavior observations, W02 held her chin during the whole process and was immersed in thinking. When she encountered the river placement, she moved her eyes around the canvas and showed uncertainty. As she was placing mountain, she asked whether she can try again, can turn the object around, or can move the object forward or backward. As she is placing the field, she leaned forward; and when it came with the tree step, she was focused and seemed very determined. When she was dealing with the people and bridge, she showed smiles. Overall speaking, she was indulged in the process, and felt pleasant for this activity.

Other participants had similar experiences. They all felt more peaceful after doing Mind Collage, that completing the picture is to project their inner mind without consciously aware of emotional changes. They felt more satisfied (W06), calmer (W01), felt joy instead of cranky (W02), relaxed (W05) (W04) (W06) (W09), comfortable (W07) (W08), etc. Other than that, they felt that they understand themselves more than before (W02), can now see themselves from the other aspects (W08), and feel surprised to the results (W10). It shows that Mind Collage has spoken for itself about having the functions of visualizing self-images and creates emotional changes.

Figure 6. Mind Collage complete diagram of W02

Figure 7. Mind Collage complete diagram of M02

Type II: Stable variation of State anxiety

A typical example of Type II is M02. The complete picture made by M02 in Mind Collage is as Figure 7. His pretest score of STAT-S is 60 points, and posttest score is 61 points. From behavior observations, M02 had been immersed in the activity showing not much facial expression changes. He was much focused, and was indulged in thinking. From the whole process, he chose to block the objects or move them to outside of the frame he found the results to be inappropriate, especially during the steps of mountain and road placements. In the tree placement step, he pondered on the positions and number of trees; and when he moved the house, he showed slight movement of his mouth; as he finished placing the stones, he took a deep breath; last, he pouched on bridge construction and moved his body slightly during the sun placement. When he reviewed the system feedbacks, he held his chin and sunk into deep thoughts. He said after the Mind Collage activity that he completed the analysis in the natural and fun atmosphere, and he felt calm and relaxed without burden (M02).

Although the participants in Type II have not shown big variations in state anxiety, they expressed highly trust to the digital counseling tool, and acknowledged Mind Collage to stimulate positive emotions, and lead to fair adjustments. Other participants also gave positive feedbacks during interviews. W03 expressed that exploring oneself through playing and doing activities is innovative and interesting. Her mind stayed in the pictures she composed and wanted to do more (W03). W04 said she saw the system feedbacks as self-reminders to make her better (W04). The system had gained trust from the users. Also, W01 used Mind Collage as self-examiner for psychological status (W01), so as W03. W05 stated that Mind Collage gave her something for self-reflection, and would encourage her to do or think more for the future (W05). Mind Collage can bring to the users so many functions that show it to have made a big leap to digital counseling.

6. Conclusion

The goal of digital counseling is to assist the participants to deal with emotional issues. Since the digital tools can allow the users to immerse in the environment, be relaxed, release the sub-consicousness, and explore the true personality, their emotions can be relieved in the meanwhile. Therefore, the participants of this research stated that they were relaxed, felt interesting, and thought the tool is innovative (W03)(W04)(W07). Also, the participants both said that they can be more focused on the counseling process. The interactions between the participants and the system was quite smooth, especially the background music helps them to be relaxed (W03). Since the interactions with the digital technology was safe and full of trust, they can easily recognized with the system feedbacks about mental analysis (W10). W02 thought the system feedbacks can guide her to positive thinking, and she felt that the system was very professional (W02). As the positive counseling relationships can be established, they had less emotional connections between each other, and can focused on the counseling goals and tasks. The participants generally felt they were more in control of the counseling process, and were surprised at the freedom they have, and felt empowered. They had more equal relationship with the counselors, and can deal with their emotions they hardly touch. These were the advantages they would not have in the traditional face-to-face counseling sessions.

In the research, the participants also gave suggestions to the future improvement of the system. They thought the design of the graphics and paintings can be more delicate (W03), and can satisfied with the need of revealing one’s inner self in some way. One can seriously face to self alone (W06). Also, for those who are not good at painting, using graphical representations to reach the goal is a nice way (W07). It saves time (W08), but hoped the graphics in each categories can be increased to suffice more preferences.

In the future, the development of digital counseling tool can be more truly project one’s mental images, create a channel for the integration of inner mind and outside world. More importantly, expressive art therapy has the functions to reduce one’s emotions by reaching to the inner mind due to its ability of privacy. Their emotions can be completely accepted by the digital tool, and can calm down their negative emotions during the process. It is of great importance to found that digital counseling tool can has such functions that it can be used for everyone to create a daily happiness through various kinds of expressive art therapy techniques.

Acknowledgements

This study is supported in part by the National Science Council of the Republic of China, under contract numbers NSC101-2511-S-024-009-MY3.

References

Dalley, T. (1984). *Art as therapy: An introduction to the use of art as a therapeutic technique*. London: Tavistock.

Egan, G. (2007). The skilled helper: A problem-management and opportunity-development approach to helping (8th). Belmont: Thomson.

Ellis, A. (1998). Rational emotive behavior therapy: A therapist's guide. San Luis Obispo, CA : Impact .

Greenberg, L.S. (2006) Emotion-focused therapy: a synopsis. *Journal of Contemporary Psychotherapy*, *2*, 87-93.

Malchiodi, C. A. (2007). *Handbook of Art Therapy*. NY: The Guilford Press.

Moreno, J. L. & Moreno, Z. T. (1975). *Psychodrama third volume: Action therapy and principles of practice. New York: Beacon House.* 。

Pekrun, R. (2000). *A social-cognitive, control-value theory of achievement emotions: Developing motivation and motivating development*. In J. Heckhausen (Ed.), Motivational psychology of human development (pp. 143-163). Oxford, England: Elsevier

Pintrich, P. R., & Schunk, D. H. (2002). Motivation in education: Theory ,research, and applications (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Riley, S. (1999). Brief therapy: an adolescent invention. *Art Therapy: Journal of the American Art Therapy Association*, *16*(2), 83-86.

Spielberger, C. D. (1983). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Richard, L., Gorsuch, & Lushene, R. E. (1971). STAT Manual: for the State-Trait Anxiety Inventory, *Consulting Psychologists Press*.

Wang, C. H., & Fei, Y.C. (2006). Single Session Counseling Model for Clients with Anxious Mood, *Guidance Quarterly*, *42*(1), 1-11.

Withrow, R. L. (2004). The Use of Color in Art Therapy. *Journal of Humanistic,Counseling Education and development*, *43*(1), 33-40.

Wright, J. (2002). Online counseling: Learning from writing therapy. *British Journal of Guidance & Counseling*, *30*, 285-298.

Math Island: Designing a Management Game of Primary Mathematics for Facilitating Student Learning

Charles Y.C. Yeh[a], Hercy N.H. Cheng [b], Zhi-Hong Chen [c] & Tak-Wai Chan [d]

[abd]*Graduate Institute of Network Learning Technology, National Central University, Jongli, Taiwan*
[c] *Information Communication, Yuan Ze University, Jongli, Taiwan*

[abd]Charles, Hercy, Chan@cl.ncu.edu.tw
[c]zhchen@saturn.yzu.edu.tw

Abstract: In Taiwan, primary mathematical education currently adopts teacher-centered learning. Teachers are the main source of course and students must to be engaged in the same activity in the same place at the same time. In this study, we designed a digital game-based learning system which called “Math Island” to help students’ mathematics learning by themselves. In the system, we produced 488 mathematics interactive materials, videos and games into a management game and combing “knowledge map” help students to understand the relationships of mathematic concepts. In the system, students can learn mathematics by their own pace by a variety kinds of learning tasks to learn different types of mathematical materials. Moreover, we also collection all user data toward the system and examine how the Math Island system influences on students learning performance and progress. The preliminary results showed that the Math Island may help students make a good progress. In the future, we will In-depth analysis in the further research and implement a diagnosed mechanism to help students learning.

Keywords: knowledge map, knowledge structure, mathematics learning

1. Introduction

In recent years, digital learning has become a more and more popular issue. Digital learning is the use of electronic media, information and communication technologies in education. In addition, it includes game-based learning, intelligent systems, virtual classrooms and digital collaboration. Digital learning will bring new revolutions and learning approaches to education (Evans, 2008; Liaw, 2008). Among all digital learning approaches, game-based learning is a famous one. This may be caused by the fact that there are many research indicating that game-based learning has some benefits to improve learner’s motivation and learning effects (Papastergiou, 2009;Huizenga, Admiraal, Akkerman, & Dam, 2009; Liu & Chu, 2010; Paraskeva, Mysirlaki, & Papagianni, 2010 ;Huizenga, 2009; Paraskeva, 2010; Liu & Chu, 2010; Prensky, 2001; Kiili & Ketamo, 2007; Van Eck, 2006;Kiili, 2005). More specifically, Traditional learning is absorbing knowledge passively. Conversely, digital games provide control and challenge that can make player feeling fun (Prensky, 2001; Crawford, 1982). Some findings are in agreement with the result of Kiili (2005) which indicating that game-based learning environments can inspire students’ learning and provide students with a great deal of learning opportunities to improve their learning. Additionally, many studies have shown that game-based learning has positive effects on student’s spatial abilities and attention (Barlett, 2009). Moreover, a digital game also has a specific goal which motivates players to complete their task (Terborg, 1976). Furthermore, some research such as situated learning, problem solving, and learning motivation are also reported in the digital game-based learning (Li & Tsai, 2013; Tüzün, Yılmaz-Soylu, Karakuş, İnal, & Kızılkaya, 2009). Due to such diversity, there is a possibility that digital game-based learning will bring new learning methods.

Digital game-based learning can occur in the classroom or out of campus because learners can use mobile devices. As a matter of fact, digital game-based learning can also be employed to support education in any environment and give an opportunity for students to choose what they want to learn, and understand how well they learn. Thus, the learning and assessment method need to be changed because the information technology introduction in education. There are many previous studies indicate that mathematics is a very useful skill in daily life and impact learning on science (Espy et al., 2004). However, students learning mathematic are passives (Downes, 2005). In Taiwan, primary mathematical education currently adopts teacher-centered learning. Teachers are the main source of information, while students are required to be engaged in the same activity in the same place at the same time (Hwang et al., 2012). However, personal differences exist in every student. It is important that let students choose what they want to learn (Barr & Tagg, 1995). Among a variety of mathematical knowledge method, knowledge map is a useful method because it can affect how students associate the relationship with their prior knowledge. There are some prior studies showed that knowledge maps can help learners to memorize knowledge and connect the relationship with concepts (Davenport, 1998). A famous learning platform which called "Khan Academy" is also used knowledge maps and portfolio to help learners learn. Knowledge maps could help students to understand the relationships of mathematic concepts between course units and chapters. For this reason, knowledge maps could improve student's learning performance (O'Donnell, 2002).

If we imported digital learning into learning environment, it may be brings many possibilities. Students can control their own learning progress and use all kinds of tools for learning. Therefore, teachers can be converted to the role of mentor and learning materials can be shared for each teacher (Hwang et al., 2012). Moreover, in the digital environment, it can record all students learning portfolio to show students learning effects. To this end, we try to design a digital game-based learning system which allows students control their learning by their own pace and provide a complete learning content with game-based model to enhance students' motivation for learning. Based on the above background and motivation, this research develops a digital game-based learning system which called "Math Island" to help students learning. This system provides a variety of mathematical learning materials in a game-based knowledge map that allows students to choose learning contents and adjust learning directions by themselves. In other words, students can decide what they want to learn by their own-pace and reach self-regulation in this system. In addition, students' learning situations are all recorded in the system. The system not only provides a visual city which using visible surface features such as building, people, and money to motivate students learning, but also have a learning portfolio to show students learning effects. In this way students can guide their own learning and achieve the learner-centered design. Finally, this study wants to explore Influence of students' mathematical learning effect and motivation in the "math island" system.

2. Literature Review

2.1 Knowledge Map

The knowledge maps include concepts and relationships between concepts indicated by a connecting line linking two concepts. It can provide guidance for students to explore concept (Davenport & Prusak, 2000). A study by Patterson et al. (1992) indicates that training benefited students'' performance with knowledge maps. Novak (1990) investigated the concept maps are graphical tools for categorizing and representing knowledge. The other study by "SDC Learning and Networking" indicating that "A knowledge map is a useful tool for presenting what knowledge resides where (e.g. people, media, organizational units or sources of knowledge) and for demonstrating the patterns of knowledge flow". Casonato（2000）consider that that knowledge map is a guidance that shows what resources can be used to help users in the shortest time to find the necessary knowledge. The main purpose of knowledge map is to present a variety of knowledge sources and guide users to use. Holsapple (2002) indicating that using a graphical presentation of a knowledge map can help users to understand and explore knowledge. In addition, O'Donnell (2002) research on the use of knowledge maps has consistently shown benefits for students with low verbal ability or skills. The aforementioned research shows that knowledge maps could improve student's learning performance. To this end, this study attempt to transfer the curriculum to the knowledge map and evaluating the knowledge map design for student's mathematic learning.

2.2 Self-regulation

Self-regulation refers to individual monitoring, controlling and directing aspects of their learning for themselves. Nowadays, students catch their own learning more effective with learning technology. After digital content has been imported in learning, students have the opportunity to decide what to learn. As a result, student Self-regulation learning becomes a new learning way. Self-regulation means that students can refer their own personal learning portfolio to decide personal learning goal and plan learning plan, then do self-paced learning. After learning, they can improve and adjust their individual learning (Bandura, 1986). In general, self-regulation is widely used in the learning situations because it can explain individual learning portfolio, monitoring individual effectively, and catch corrective self-reactions during learning. On the one hand, using technology to develop self- regulation learning is helpful for students to re-organize knowledge (Winne et al., 1998). As a result, students' learning and motivation are interdependent via self-regulation (Zimmerman, 1990). Furthermore, developing students' ability and skill about improving learning performance with self-regulation in school education and family education is an important thing. Moreover, the Self-paced learning also related to the self-management learning. Self-paced means that there are no group assignments or peer review and that learners are able to move with their own speed. Accordingly, using technology to help developing self-paced learning allows students to explore new forms of self-regulation learning and helps students re-organizing knowledge (Winne and Hadwin, 1997). In addition, another important approach is that learning and motivation of students is interdependent (Zimmerman, 1990). The study by Pintrich and Schrauben (1992) examined students are active while learning rather than being passive to accept information. Indeed, it's the best learning method for students that students are active in looking for learning opportunities and after finding learning opportunities, they learn with passion. Moreover, they try to find solutions themselves like setting goals and self-reflecting to adjust their learning method. As in section 1, students were often seemed as passive role in learning process and their learning methods and learning outcomes had a parallel relationship with the teacher teaching way (Kember, 1997). In brief, teacher not only was an assistant but also the protector of class order. Thus, students should leaded their own learning and it was meaning that they has a choice in their learning.

2.3 Learning Portfolio

Traditional approaches of learning, teaching, and assessment need to change after technology imported (Birgin & Baki, 2007). In the current learning environment, learning process should be more emphasized because the traditional test assesses student learning difficulty. Portfolios can be a valuable research tool to gain understanding about personal accomplishments (Tillema, 2001) and displaying student problem solving process, learning attitude, and the growth after learning (Lin, P.J. & Tsai, W. H., 2001). Student learning portfolio is a complete record which collecting information about learner purposely and keeps all student activity processes to understand the student's efforts, growth and achievements (Arter & Spande, 2005). More specifically, learning portfolio can be a useful tool to help student reflection and display their learning outcomes. In addition, the extracted portfolio data can present students' learning performance clearly for teachers, parents, and themselves.
Digital learning has developed rapidly in recent years, which is digitizing learning processes and use internet to store in a database to help recording, searching, organizing, and analyzing the learning portfolio (Chen, 2002). Learning portfolio let students grasp learning processes, as well as help students reflecting and reviewing all learning processes and learning outcome. Furthermore, the learning portfolio gave appropriate feedbacks to support student learning (Chang, 2001). Hence, the Math Island integrates learning portfolio concept to give students immediate feedback and reflection. Moreover, the Math Island also provides teachers with information to focus students' problems. So the Math Island wishes to help students reaching in meaningful learning.

3. System Design: Math Island

3.1 Learning process

The object of this system mainly used in elementary school's mathematics learning. All students play the role of a master who is the owner of the "Math Island". In the system, every student must manage their virtual island and attracts other virtual resident to living in the island. The business model of the Math Island built from a new ground. Each ground has a variety of construction projects (or learning task). The learning process of the Math Island as showed in figure 2. Students can plan their own construction projects which indicate that students could choose a learning task by themselves. In other words, students can do a mathematical learning task to advance their house. Students will get reward form system when they are completing a learning task. Then, the building will grow up by their learning effect. The visible surface features on building represent a portfolio for students. Finally, the island will be personalized because students can purchase some unique building to develop theirs island.

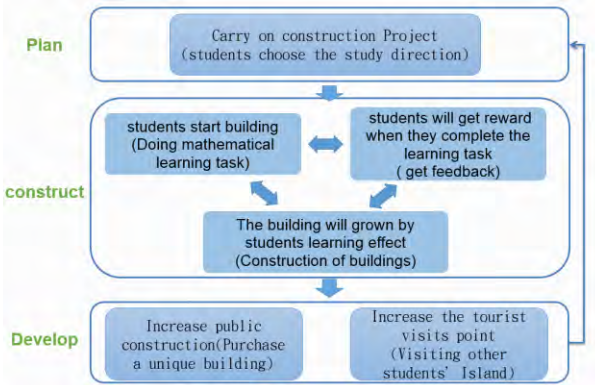


Figure 1. The Learning process of the Math Island system.

3.2 *Learning content design*

In recent years, the educational institution which called "Ministry of Education" promotes slight regulation for nine-year joint curriculum of elementary school and junior high school in Taiwan. The Ministry of Education indicators five-ability index of mathematics which are "Number and quantity", "geometric", "algebra", "Statistics and Probability", and "link". For this reason, the interface and learning content of the Math Island system is according to this five-ability index. However, the index about "algebra" and "link" are not adapted to be an index of the math island system. As a matter of fact, among the contents of the elementary school and junior high school, most of the "algebra" units are in the junior high school's curriculum. Furthermore, the index of "link" units contains many elements such as perceived transformation, problem solving, communication, and analysis. More specifically, it is difficult to define "link" and "algebra" be an index that these elements relate to many mathematic units. Accordingly, the index of "algebra" and "link" units are integrated with other units in this system. Besides, the index of "Number and quantity" in the system are divided into two indexes which called "Numeral and Computing" and "Quantity and Measure" because the "Number and quantity" are covering a wide range of curriculum in elementary school.

To this end, this study draws a mathematical knowledge map which was including four indexes: "Numeral and Computing", "Quantity and Measure", "geometric", and "Statistics and Probability" (see Figure 2). To begin with, these four indexes are depending on the nine-year joint curriculum and divided into a variety of units and sub-units. Then, each unit contains many sub-units and has principles to connect these sub-units. Finally, these units and sub-units have been classified into some roads in the system. Hence, there are many different roads which named by relate units in the math island system. Thus, the indexes of "Numeral and Computing" are containing a very large part of the Math Island because it plays an important role in the primary mathematic curriculums. Basically, the student can do learning tasks in different areas to perform mathematics learning. Table one show the design of knowledge map's structure on the Math Island. We are implementing a great amount of learning task in the system.

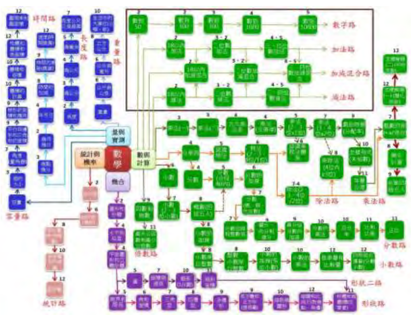


Figure 2. The knowledge map structure.

Figure 3. System interface in the system.

There are much different learning content in the math island. In the system, we design three types of task, such as prior concept review, concept teaching, and mastery learning for students' learning. However, the subject of mathematics learning was usually limited by student's prior knowledge (ex. If a student wants to learn" single-digit addition", he/she must understand the concept of "single-digit" and "addition"). Therefore, this course has been structured to operate in a self-paced manner. In addition, in order to confirm that students really learn a knowledge in self-paced, we set a mechanism which classified the prior concept review task and concept teaching task as "key task" to regulation on students learning. Students must have been completed the "key task" for fear that students have enough prior knowledge to do the next task. In other words, Students must get a passing "key task" to move to the next course.

Figure 4. Snapshot of the learning content interface.

The learning content of each task has referred to a variety of publishing house and re-design by many experts. The learning contents are included graph, text, and instructional videos to guide students learning. Besides, it provides a scratch paper and visual numeric keypad for students to answer questions. More specifically, the system also has a variety of materials which including crosswords, multiple choice, fill in the blank, matching, true or false of questions. These different kinds of questions can support students to learn different concepts. Finally, when students are answering a question, the system will calculate accuracy immediately. If students can get higher accuracy they will earn more rewards. Furthermore, the system will give students positive feedback when students respond a correct answer. Conversely, the system will give students hint to guide students when they respond a wrong answer. There are some learning content examples shown in the figure 4.
The other tasks are mastery learning task which using two digital games to appearing in the system. One is "Happy Farm", and the other is "Fishing Expert". These games have some game elements to enhance students' motivation for their learning. All of learning tasks in the math island develop in accordance with this mechanism.

Figure 5. Snapshot of A mastery learning games.

3.3 System mechanism

When students login into the system, an arrowhead would be displayed on the monitor and guide them to choose a learning unit (see Figure 6). In addition, if students click a building, the scenario will transfer to "learning lobby". The lobby will present many learning tasks for students to choose. The task with blue color means that it can be performed. On the contrary, the gray means students need to complete the prior key task to open the task.

Figure 6. Some arrowheads guiding students' learning.

Figure 7. A task list in the building.

In addition, as shown in figure 8, the system also provides immediate feedback mechanism to help students answer a question. More specifically, when students input a correct answer, the system will show "√" message to represent the correct answer. In contrast, when students input a wrong answer, the system will give immediate feedback or hints to help student's reflection. The rewards for students are according to their accuracy rate in the task. Besides, in the system, each student can construct the map individually and giving feedback for teacher on the level of understanding of every student. Within the option of individual construction of the map, the students can be allowed to visit the other student's island where students are able to share information. The system provides this mechanism for students to interact with others and enhance their learning motivation.

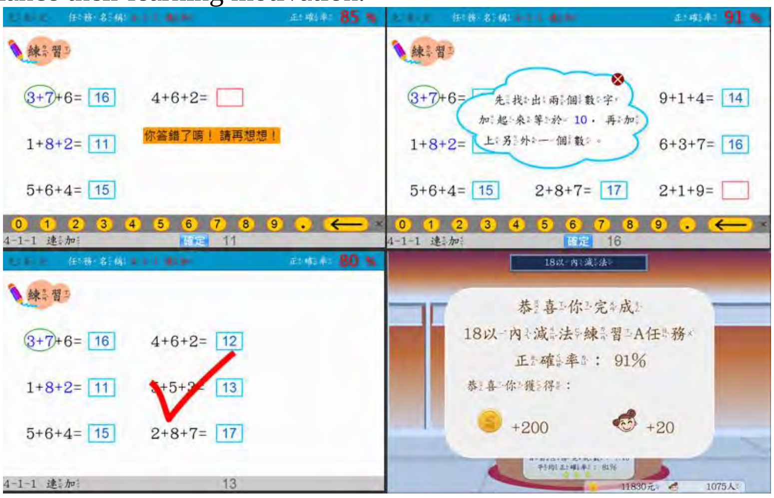


Figure 8. Snapshot of the learning task with immediate feedback.

According to task completion, the buildings can be divided into four grades in total. As show in figure 9, students' learning outcomes will be presented in these buildings to show students learning progress in the Math Island system. In addition, students can understand their learning situation and reflection themselves by their own-paced.

Figure 9. Four grades in the buildings.

4. Case Study

4.1 Participants and Data collection

The participants were 217 second-grade students of a primary school in Taiwan. All participants have a tablet PC to conduct a digital environment. At the beginning, we held a typing training program and system introduction course in order to make sure that all participants could control the prerequisite skill. At the end of the training program, the participants' can login the Math Island system to do a learning task by their own-pace. The systems will collection all user data in the database.

4.2 Preliminary results

According to the system data, some students not only reached the standard progress in semester 1 but also reached the second-grade progress by using the system. Furthermore, the data in Table 1 demonstrates the progress in different mathematic units of all students. The standard curriculum pace is 101-1 and more than 210 students can reached the next semester curriculum. The preliminary shows that some students have a good progress by self-learning but it also need to In-depth analysis in the further research.

Table 1: The results of the student's task data.

	2th Grade Class 1	2th Grade Class 2	2th Grade Class 3	2th Grade Class 4	2th Grade Class 5	2th Grade Class 6	2th Grade Class 7	2th Grade Class 8	Total
Participants	26	28	27	28	27	27	27	27	217
Complete task number	123	138.1	121.1	134.3	130	129.7	128.0	141.6	130.72
Average accuracy	88.26	90.28	88.33	89.52	87.83	88.26	89.16	87.97	88.701
Less than 101-1	1	2	3	0	0	0	0	1	7
Reached 101-1	25	26	24	28	27	27	27	26	210
Reached 101-2	9	22	11	17	18	15	16	22	130
Reached 102-1	1	1	0	2	0	0	0	4	8
Reached 102-2	0	0	0	1	0	0	0	1	2

5. Conclusion and discussion

In this study, the Math Island system with knowledge map and self-regulation design was developed for students in learning mathematics by their own-paced and self-reflection in an easier way. We produced 488 mathematics interactive materials, videos and games and conduct a management game to motivates students' learning. Meanwhile, the aims of this study not only examined whether the Math Island system is useful to improve students' learning motivation, but also investigated how students' learning progress in the system. The preliminary results show that elementary students may be able to learn in the self-regulation learning environment and learning well themselves. However, although students' progresses were increasing at this time, how the influence on the system should be further evaluated and discussed. In the future, we will do more detailed data analysis of the system, support more learning materials and implement a diagnosed mechanism in the system to help students learning.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (MOST 101-2511-S-008 -016 -MY3, MOST 100-2511-S-008 -013 -MY3) and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

Arter, J. A., & Spandel, V. (2005). Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice*, *11*(1), 36-44.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Clinical and Social Psychology*, 4, 359-373.

Barr, R. B., & Tagg, J. (1995). From teaching to learning—A new paradigm for undergraduate education. *Change: The magazine of higher learning*, *27*(6), 12-26.

Barlett, C. P., Anderson, C. A., & Swing, E. L. (2009). Video game effects confirmed, suspected and speculative: A review of the evidence. Simulation & Gaming, 40(3), 377-403.

Birgin, O., & Baki, A. (2007). The use of portfolio to assess student" s performance. *Journal of Turkish Science Education*, *4*(2), 75-90.

Casonato, R., Furlonger, J., & Harris, K. (1999). What Are the Steps Needed to Create a Knowledge Map? . GartnerGroup Research Note, 26(July).

Crawford, C. (1982). *The art of computer game design*. Retrieved from http://www.vic20.vaxxine.com/wiki/images/9/96/Art_of_Game_Design.pdf

Chang, C-C. (2001) “A study on the evaluation and effectiveness analysis of web-based learning portfolio”, *British Journal of Educational Technology*, Vol 32, No 4, pp 435-458.

Chen, D. L. (2002) .Design And Implementation of a Web-Based Portfolio Supporting System. (Doctoral dissertation or master's thesis). Retrieved from http://etd.lib.nsysu.edu.tw/ETD-db/ETD-search/getfile?URN =etd-0625102-190739&filename=etd-0625102-190739.pdf

Downes, S. (2005, October 17). E-learning 2.0. eLearn Magazine [Web blog message]. Retrieved from http://elearnmag.acm.org/featured.cfm?aid=1104968

Davenport, T. H., David, W., & Beers, M. C. (1998). Successful knowledge management projects. *Sloan management review*, *39*(2), 43-57.

Davenport, T. H., & Prusak, L. (2000). *Working Knowledge: How Organizations Manage What They Know*, Harvard Business Press.

Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Computers & Education*, 50(2), 491–498. doi:10.1016/j.compedu.2007.09.016

Gartner Group (2000)."Mapping knowledge to the process: directions to success". Advisory Research Note.

Holsapple, C. W., & Joshi, K. D. (2002). Knowledge management: A threefold framework. *The information society*, (18), 47-64.

Huizenga, J., Admiraal, W., Akkerman, S., & Dam, G. ten. (2009). Mobile game-based learning in secondary education: engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning*, 25(4), 332–344. doi:10.1111/j.1365-2729.2009.00316.x

Hwang, G. J., Su, J.M. & Chen, N. S. E-learning: Theory & practice (pp. 2-29). Taiwan：Drmaste.

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education*, *8*(1), 13-24.

Kiili, K., & Ketamo, H. (2007). Exploring the learning mechanism in educational games. *Journal of Computing and Information Technology*, *15*(4), 319-324.

Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and instruction, 7*(3), 255-275.

Liaw, S.-S. (2008). Investigating students’ perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, 51(2), 864–873. doi:10.1016/j.compedu.2007.09.005

Li, M.-C., & Tsai, C.-C. (2013). Game-Based Learning in Science Education: A Review of Relevant Research. *Journal of Science Education and Technology*, 22(6), 877–898. doi:10.1007/s10956-013-9436-x

Lin, P.J. & Tsai, W. H. (2001). Using research-based cases to enhance prospective teachers’ understanding of teaching mathematics and their reflections. In F. L. Lin (Ed.), *Common sense in mathematics education (pp. 231-272). Proceedings of 2001 the Netherlands and Taiwan Conference on Common Sense in Mathematics Education*, 19-23, November 2001. Taiwan: Taipei.

Liu, T.-Y., & Chu, Y.-L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education*, 55(2), 630–643. doi:10.1016/j.compedu.2010.02.023

Novak, J. D. (1990). Concept maps and Vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional science, 19*(1), 29-52.

Patterson, M. E., Dansereau, D. F., and Newbern, D. (1992). Effects of communication aids and strategies on cooperative teaching. J. Educ. Psychol. 84: 453–461.

Pintrich, P. R., & Schrauben, B. (1992). Students’ motivational beliefs and their cognitive engagement in classroom academic tasks. *Student perceptions in the classroom*, 149-183.

Prensky, M. (2001). *Digital Game-Based Learning*. New York: McGraw-Hill.

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education:Impact on educational effectiveness and student motivationa. *Computers & Education*, *52*(1), 1-12.

Paraskeva, F., Mysirlaki, S., & Papagianni, A. (2010). Multiplayer online games as educational tools: Facing new challenges in learning. *Computers & Education*, 54(2), 498–505. doi:10.1016/j.compedu.2009.09.001

Tillema, H. H. (2001). Portfolios as developmental assessment tools. *International Journal of Training and Development*, *5*(2), 126-135.

Terborg, J.R.(1976). The motivational components of goal setting. *Journal of Applied Psychology*, 61, 613– 621.

Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., İnal, Y., & Kızılkaya, G. (2009). The effects of computer games on primary school students’ achievement and motivation in geography learning. *Computers & Education*, 52(1), 68–77. doi:10.1016/j.compedu.2008.06.008

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE review*, *41*(2), 16.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 279-306). Hillsdale, NJ: Erlbaum.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist*, *25*(1), 3-17.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Change of Interpersonal Relationship for Group Development in Digital Game-based Adventure Education Course

Chang-Hsin LIN*, Yu-Jen HSU, Ju-Ling SHIH, & Chia-Chun TSENG
Department of Information and Learning Technology
National University of Tainan, Taiwan
*arthur1379@hotmail.com

Abstract: The development of physical adventure education activities has existed for a long time. However, there was little research concerning its applications and developments with technology. This research designed a digital game-based adventure education course with Tuckman's stages of group development. Six traditional adventure education activities were chosen to be developed into digital forms by Unity3D and were developed into three different platforms: desktop computers, tablets, and motion-sensing devices. Bales interaction process analysis was used to observe the development of group interactions of the participants during the experiment. The results showed that the interpersonal relationships between participants in groups can be effectively improved after the course.

Keywords: Adventure education, Digital game-based learning, group development, interaction process analysis.

1. Introduction

Adventure education is formed by risky and challenging activities experienced by students. The connotations of adventure education were thought and internalized through the personal experiences to the activities and group reflections after the activities led by the facilitators. For several decades, the activities of adventure education have been implemented in the physical form. Students have been grouped up and experienced the activities in the wild for a few days or even weeks. Group members had face-to-face contacts, went through the courses with members, and experienced their missions of activities from being strangers to partners. Goals for the activities were learned in the reflection sessions with the facilitators after each activity. Group members could use the lessons they learned in the daily life situations. However, few studies have presented insights of adventure education with digital games.

With the advance of technology, more and more teenagers are interested in digital games. Traditional adventure education activities are seldom their learning or leisure options. Conversely, teenagers are more apt to show themselves in the virtual world of games in terms of their real personalities, creative ideas, and internal emotions. Therefore, this research aims to investigate whether digital games can effectively become the medium for adventure education, and whether teenagers can establish interpersonal relationships more successfully.

The study aims to design a digital game-based adventure education course base on Tuckman's stages of group development. Six traditional adventure education activities that are difficult to be implemented due to location or weather limitations were chosen to be digitized into digital games and were integrated in accordance to the five stages of Tuchman's stages of group development. Bales interaction process analysis was used to observe the development of group interactions of the participants during the course experiment.

The game engine Unity3D was used as the development tool. With the advantages of cross-platform, high-simulation, and easy-manipulation, researchers transformed the six traditional adventure education activities into two computer games, two tablet games, and two motion-sensing games. After the course, students can not only learn positive interactive relationship of adventure education, but also have fun with digital games.

2. Literature review

2.1 Adventure Education

The courses of adventure education which takes days or even months to finish were originally long and extensive. In 2001, the essences of adventure education were devised into small-size courses which only took a few hours (Glass & Myer, 2001). Glass and Myer (2001) thought the learning effectiveness of adventure education courses should not only be analyzed with questionnaires. Research should increase observations and interviews to investigate students' learning effectiveness. They thought students' behaviors in groups would be affected by the individual state of mind. Consequently, they observed the performance of each student in the reflection session from the individual psychology side. For several years, studies in adventure education area were based on thematic course or activities. In company with Baldwin, Persing, and Magnuson (2004), Brown (2006) indicated that the difference between pre-test and post-test have been over-focused in those studies. Those studies ignored the change of students' behaviors and learning styles in the process of activities. In 2009, a detailed account of using qualitative research method, such as interviews and observations, was presented to explore the change of peer relationships, interpersonal skills, and students' self-concept during the process (Zmudy, Curtner-Smith, & Steffe, 2009).

In recent years, the design of adventure education course has been a fixed process in Taiwan and other countries. The course design was always based on Tuckman's stages of group development (Tuckman & Jensen, 1977). The five stages of group development were combined with suitable activities into the adventure education course. Through those activities, every group was developed to be a high-performance team. However, there were only a few studies that investigated the interaction details in the process. Therefore, this study aims to present a detailed analysis to the interaction process.

2.2 Digital games in counseling area

Digital game-based learning (DGBL) has been gradually emphasized in recent years. The purpose is to improve students' learning motivation and effectiveness through digital games. Face to face interviews were usually used in the traditional counseling which may build up intangible stress to the participants that were unhelpful to bridge the communication gap. It always took lots of time to drop one's guard that only can be done by counseling psychologists. Without proper temptations, teenagers kept losing their interests when doing activities in a confined room. It is supposed that digital games can help the counseling activities to be carried out smoothly. Taking the advantages of digital games can lower down the stress to them and help them to focus on the counseling issues. To help people to have in-depth understanding to themselves and are willing to make necessary changes are the purposes of counseling. It is worth to try to unload the defenses of participants and achieve the purposes of counseling.

In the last two years, many digital games were developed based on adventure education and counseling activities were developed by our team (Hsu & Shih, 2013), and were used in the individual and group counseling. Their research results show that the learning effectiveness of digital counseling games perform equal or even better than the traditional ones, as well as the attention and motivation of the students in participating of the adventure education activities. It is proved to be feasible to implement the teaching methods of digital game-based learning in the courses of adventure education or counseling area.

2.3 Interaction process analysis

Bales interaction process analysis (IPA), which is widely used and is especially suitable to analyze problem-solving groups, is an analysis tool about group interactions. Task dimension and socio-emotional dimension were the two dimensions of interactive behaviors that IPA considers. The two dimensions are opposite to each other (Bales, 1950). In the recent years, studies range from group efficacy, parent-child relationships, and counseling were focused on observing the physical group activities. Some studies explored about the use of technology to help researchers to record game dialogs. Interaction observations were recorded by computer-assisted tools (e. g. Nam, Lyons, Hwang,

& Kim, 2009; Severino & Messina, 2010). Those studies all indicated that communication is easy to observe and record with computers, but the on-line group interactions are less effective than face-to-face ones. There were some studies of adventure education using IPA to observe group members' interaction process. However, observations to the group interactions in the digital games were rare to find.

3. The digital game-based adventure education course

Tuckman (1977) thought the groups development go through stages sequentially, thus, the theory described the process in five stages. However, members' negative behaviors such as bad communications and misunderstandings in the storming stage would hinder group development. In this situation, group members had to face and solve their problems in order to move forward to the next stage, and thus become high-performance groups. Table 1 describes the individuals and groups situations for each stage of group development.

Table 1: Situations of group development in Tuckman's five stages (Tuckman, 1977)

	Individual situations	**Group interaction situations**
Forming	1. Explore, feel strange to everything. 2. Both individual and group goals are not clear. 3. Look for his position in the group.	1. All members are not familiar with each other. 2. Finish personal works without group work.
Storming	1. Look for the individual goals 2. Have personal opinions, dealing with works from their own angles. 3. Have influence in group.	1. Form small groups. 2. Groups have abilities to divide the works to each member, and can complete simple missions.
Norming	1. Understand group goals. Individuals are team-oriented. 2. Accept others' opinions and views, and make adjustments. 3. Personal motivation. Members take roles in groups.	1. Cooperation. Members finish all missions successfully. 2. Groups work on the same rhythm. Understand personal roles and values.
Performing	1. Identify with group goals. Have personal ideals. 2. Look for further growths and changes. 3. Help partners to grow from each other.	1. Have great interactions and work skills. 2. Have great group climate. Cover and support each other.
Adjourning	1.Share self-reflections and experiences. 2.Produce common memories in the development process. 3.Say goodbye. Start next cycle.	1. The cycle ends. Review and share the process of group development, common experiences, and memories.

Two books about adventure education, "Experiential education- learning from 150 games" (Hsieh, Wang, & Chuang, 2008) and "Experiential education: theory and practice" (Kuo, Liao, & Shih, 2009), were used to identify suitable activities for the digital counseling games. The activities were designed to be conducted in groups of five members. The design of the digital game-based adventure education course was based on Tuckman's stages of group development. The five group development stages are described as follows.

Forming: Members meet each other for the first time. Everything is strange for them. The warm-up activities are chosen to let them share, communicate, and give opinions frequently so that the members can be familiar with each other. The initial group relationship is built by group members. The purpose of the first stage activity is to create group relationships, breaking fixed conceptions, and making reflections. The game which is chosen to digitize for this stage is Polar Bear and Hole (Figure 1). At the beginning, a story which provides hints for the task is presented in the game. The facilitator throws dices three times for giving hints. After that, members observe, discuss, and induce for correct answers following the logistics of the story.

Storming: The activities which need deep communications between members to finish the missions are contained. In this stage, all members have to discuss the group problems with their internal dialogues and work together to complete the tasks. The group problems and personal weaknesses are improved by themselves. After this stage, groups can become high-performance teams. The purposes of the activities in this stage are cooperation, communication, respect, breaking fixed conception, problem-solving, trust, and reflection. Cooperative Puzzle (Figure 2) and Chessboard Maze (Figure 3) are used in this stage. In Cooperative Puzzle, three puzzle pieces out of total of fifteen puzzle pieces are distributed to each member. All five members in the group have to work together to assemble five equal-size squares with puzzle pieces. Members can exchange puzzle pieces with others without discussion. They can only give out puzzle pieces and may not ask from others. The purpose of the game is to make members pay attention to others' needs and realize their roles in the group. In Chessboard Maze, only one person can play at a time. Each member has to pass the 9x5 chessboard which is full of landmines with only one path. The member has to start over again when he goes through the wrong path. Every failure leads to point deduction. To generate gaming strategies together through members' experience, challenge obstacles, break fixed conceptions, and trial-and-error, and to build the trust between members are the goals of the game.

Norming: After the stage of storming, members start re-thinking about their relationships in the group. With communications, members start to trust each other. Members create greater effectiveness and achieve group goals more rapidly. The goals of the activities in this stage are communication, leadership, reflection, and cooperation. Moon Ball (Figure 4) is used in this stage. The ball is patted by group members to keep it in the air as long as they can. To train members to use strategies to cooperate with each other is the purpose of the game.

Performing: Members now are in great interactions and have skilled cooperative strategies after going through the frustrations and difficulties in the previous activities. The goals of the activities in this stage are trust, cooperation, leadership, communication, and reflection. Group Balance (Figure 5) is used in this stage. Three people are required to play the game. One member keeps his balance on the board and the other two members keep pushing the board upward on the two opposite sides until reaching the top. To increase members' cooperation strategies, reliance, and face the challenges are the purpose of the activity.

Adjourning: Group development process comes to the end. Members encourage each other and share their experiences. Having existing members to leave, new members to join, and start a new cycle, are the situations the members in the group must face. The goals of the activities are communication, trust, leadership, cooperation, and reflection. Calculator (Figure 6) is used in this stage. Thirty ladybugs are marked from numbers 1 to 30 in a garden on the screen. The goal is to catch the bugs serially as fast as possible. Five seconds would be added as punishments rules were violated. Four chances were given to each group and the best one would be the final score of the group. To train members to solve problems and communicate are the purposes of the activity. It strengthens group concepts and identifies their roles.

Those games in this study were devised for adventure education which required highly interactions and simulations. Therefore, Unity3D game engine was used to develop those games. With its great performance of simulation, members can become immersed as if they were in the real scene. The game engine also provides cross-platform publication such as IOS, XBOX360, Android, Web, Wii, and PC. Because physical activities vary, the digital games designed for this course had to be presented in different formats and platforms to fulfill the needs. Using Unity3D to make cross-platform games also saved time.

Figure 1. Polar and Hole

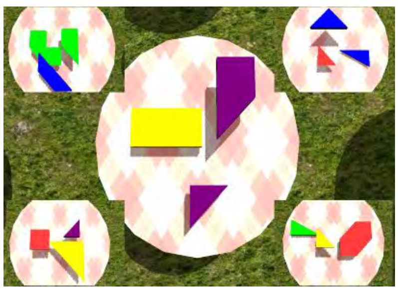


Figure 2. Cooperative Puzzle

Figure 3. Chessboard Maze

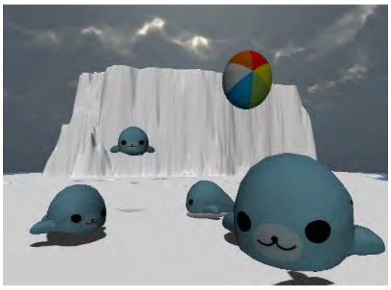


Figure 4. Moon Ball

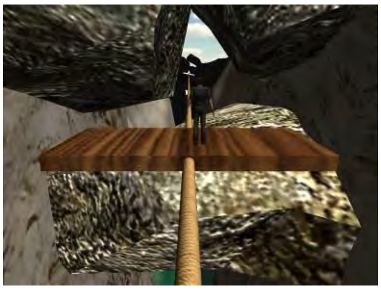


Figure 5. Group Balance

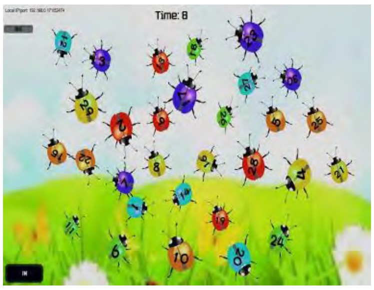


Figure 6. Calculator

4. Experimental Design

In this research, six digital games were included in the course that adults or teenagers were ideal targets for. Therefore, thirty college students aged between 19 and 25 from Taiwan were randomly recruited to participate in the course. The experiment was divided into three sessions in three different days. Two groups, with five members in each group, were guided by a facilitator to take the course in a day; the whole course lasted for six hours. Audio and video recordings were made during the experiment for later observations with Bales interaction process analysis (Figure 7).

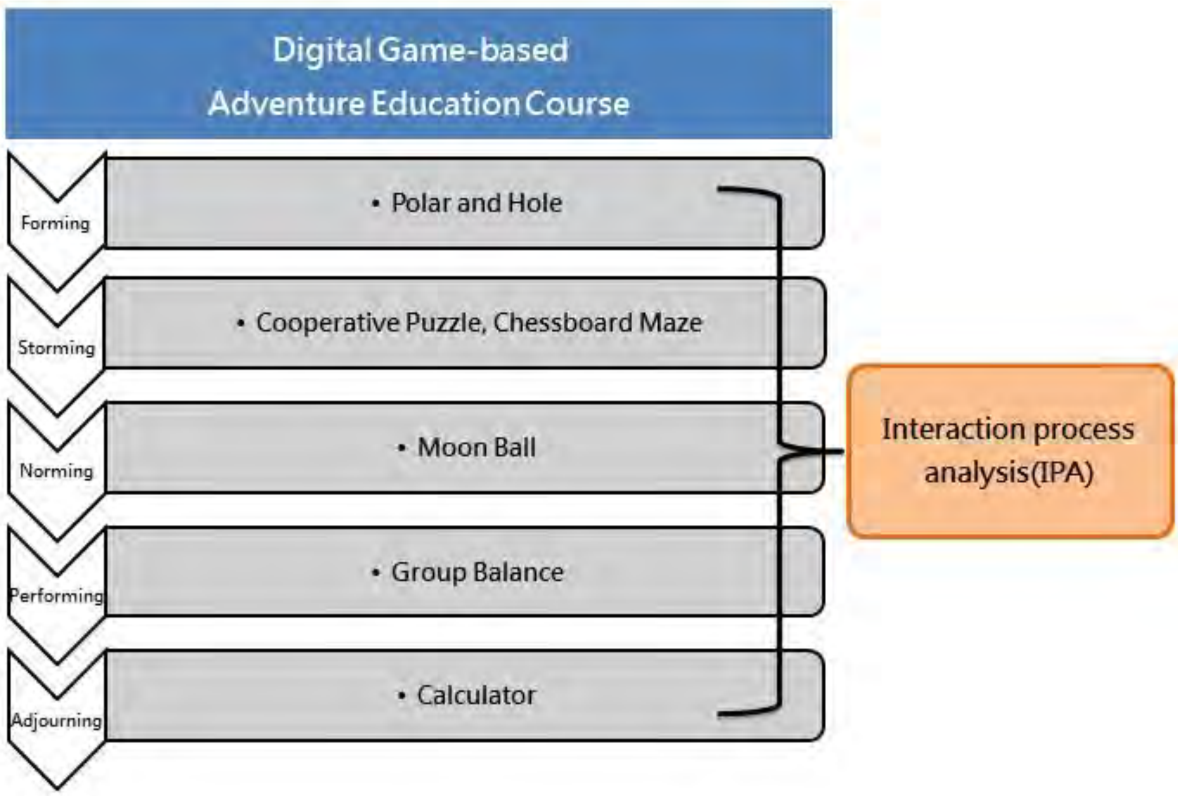


Figure 7. The flowchart of experiment procedure

According to IPA, members have to solve group problems in two dimensions. 1. Task dimension. The contents of dialogues were the situations which group members have to ask, solve, guide, order, and suggest. Task-oriented interactions were classified in this dimension. 2. Socio-emotional dimension. The contents of dialogues were the situations which concern the acceptances, coordination, conflicts, and opinions in group members. Socio-oriented interactions were classified in this dimension. During the process of group development, all interactions and dialogues were recorded according to the schemes stated in Table 2. With the data which are recorded by IPA, the group development problems can be resolved and identified (Bales, 1950).

Table 2: Bales interaction process analysis (Bales, 1950).

Functional dimension	Content categories	Code

Task	Attempted answers	Gives opinion, evaluation, analysis, expresses feeling and wish.	C
		Gives suggestion, direction, implying autonomy for other.	B
		Gives task orientation, information, repeats, clarifies and confirms.	A
	Questions	Asks for orientation, information, repetition or confirmation.	A
		Asks for opinion, evaluation, analysis and expression of feeling.	B
		Asks for suggestion, direction and possible ways of action.	C
Socio-emotional	Positive reactions	Shows solidarity, raises other's status, gives help and rewards.	F
		Shows tension release, jokes, laughs, shows satisfaction	E
		Agrees, shows passive acceptance, understands, concurs and plies.	D
	Negative reactions	Disagrees, shows passive rejection, formality, with olds help.	D
		Shows tension, asks for help, with raw out of field.	E
		Shows antagonism, deflates other's status, defends or asserts self.	F
Code	A: Problem of orientation. B: Problem of evaluation. C: Problem of control. D: Problem of decision. E: Problem of tension-management. F: Problem of integration.		

5. Research Results

The changes of interpersonal relationship of the group members were shown in this section. All the results were analyzed by encoding the dialogues occurred during the experiment sentence by sentence with Bales interaction process analysis.

According to Tuckman's thesis, groups had to go through four group development stages before becoming high-performance teams. Differences of values and concepts that caused conflicts between members occurred during the group development process. Those conflicts were solved after personal adjustments, interpersonal communication, leader's coordination, and members' compromises, and the groups would start to become effective ones. The needs for coordination and communication increased when the groups undergone the stage of storming. Facilitators should encourage the group timely when the groups hesitated. Facilitators played an important role in every stage of group development. Members reached consensus on mission completions or problems solving after the stage of storming. Therefore, the frequency of group interactions peaked at the stage of storming and started decreasing when the members were on the same wavelength at the stage of norming and performing.

The interactions of Group 1, 2, 3 and 4 were mainly distributed into the dimensions of positive reactions, attempted answers, and questions defined in Bales interaction process analysis in every stage of the course. The groups faced and solved the problems positively with positive communications, and caused less negative interactions that might affect the group developments. Negative reactions mostly occurred when the groups were solving the Chessboard Maze in the stage of storming. The frequency of negative reactions in stage of storming was more than other stages. It was because the game was designed to let the members experience frustrations and failures.

Although the distributions of interaction contents of all groups were similar, the frequency of interactions of Group 5 and 6 were much less than the former groups. These two groups finished the mission of Chessboard Maze in the stage of storming smoothly without a great amount of trials and errors that they had no chance to build up the cornerstone of positive interactions and group understanding. Consequently, the frequency of interactions of the activities in the stage of adjourning which need great understandings and strategies to get better scores sprang up to finish the work that should be done before.

The interaction frequencies of group 1, 2, 3 and 4 were shown in Figure 8. The frequencies of interactions of the four groups reached a peak in the stage of storming. After a great amount of adjustments, communication, and coordination, better scores were reached easily with excellent understandings to each other. The interactions of group 2 and 3 in Calculator in the stage of adjourning were more than in Chessboard Maze in the stage of storming. From the observations of the researchers, the reasons that caused the differences might be the lack of the explicit strategies and the guidance of a leader. The members needed further discussions on the execution of strategies to get better score in the stage of adjourning.

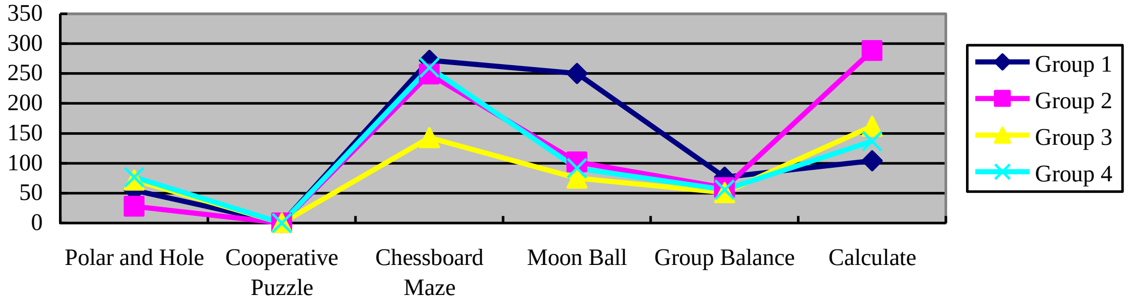
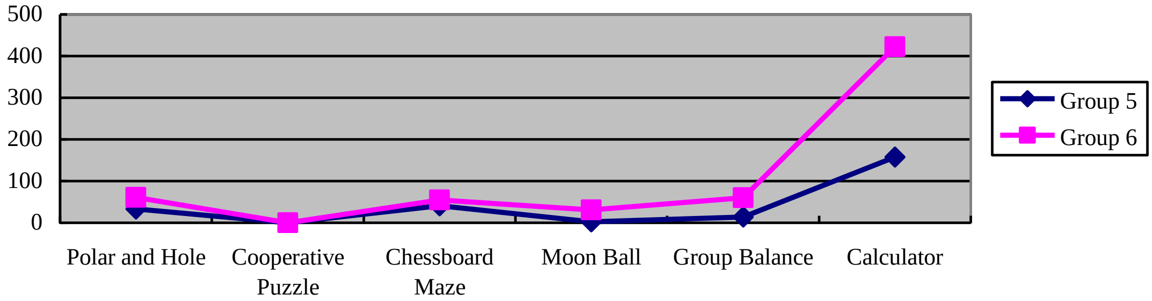


Figure 8: the line graph of interaction frequencies of Group 1, 2, 3 and 4.

As shown in Figure 9 below, the frequencies of interactions of group 5 and 6 in the stage of storming were lower than the former groups. The performances of the later stages would be affected without sufficient adjustments, coordination, and compromises in the stage of storming. Therefore, that would be worth discussing.



.

Figure 9: the line graph of interactive frequency of Group 5 and 6

6. Conclusion

Six theory-based traditional adventure education activities that were difficult to be implemented due to limitations of bad weathers, restrictions of rules, and documenting problems were developed into digital games with Unity3D. Unlike commercial games, those digital games which were combined with Tuchman's stages of group development aimed to put considerations to provide joy to the learners, and at the same time to reach the purposes of adventure education. The research questions were set to observe the process as digital games were used as the medium of adventure education, and the interaction patterns of interpersonal relationships in digital game-based adventure education course.

The purpose of this study was to investigate whether the members in each group can turn strangers into a high-performance team after taking the course designed in this study. After the analysis by Bales interaction process analysis, the results showed that no matter how fast the group developed, all six groups had become positive high-performance teams after the course. Group development efficacy was heavily depended on the adjustments, communication, and coordination they experienced in the game.

Acknowledgements

This study is supported in part by the National Science Council of the Republic of China, under NSC 98-2511-S-024-006-MY2, NSC 100-2628-S-024-002 -MY3, and NSC 101-2511-S-024 -009 -MY3.

References

Baldwin, C., Persing, J. & Magnuson, D. (2004). The role of theory, research, and evaluation in adventure education. *Journal of Experiential Education, 26*(3), 167-183.

Bales, R. F. (1950). *Interaction Process Analysis.* Chicago: University of Chicago Press.

Brown, M. (2006). Adventure education and physical education. In D. Kirk, D. Macdonald, & M. O'Sullivan (Eds.), *The Handbook of Physical Education* (pp. 685–702). London, England: Sage Publications Ltd.

Glass, J. S., & Myers, J. E. (2001). Combining the old and the new to help adolescents: Individual psychology and adventure based counseling. *Journal of Mental Health Counseling, 23*(2), 101-114.

Hsieh, C. M., Wang, J. Y., & Chuang, H. W. (2008). Experiential education-learning from 150 games. Taiwan, Taoyuan: AAEE.

Hsu, Y. J., & Shih, J. L. (2013). Developing computer adventure education games on mobile devices for conducting cooperative problem–solving activities. *International Journal of Mobile Learning and Organisation, 7*(2), 81-98.

Kuo, T. Y., Laio, S. H., & Shih, H. Y. (2009). *Experiential education: theory and practice*. Taiwan, Taichung: Far Du.

Nam, C. S., Lyons, J. B., Hwang, H. S., & Kim, S. (2009). The process of team communication in multi-cultural contexts: An empirical study using Bales' interaction process analysis (IPA). *International Journal of Industrial Ergonomics, 39*(5), 711-782.

Severino, S., & Messina, R. (2010). Analysis of similarities and differences between on-line and face-to-face learning group dynamics. *World Journal on Education Technology, 2*(2), 124-141.

Tuckman, B., & Jensen, M. (1977). Stages of small group development revisited. *Group and Organizational Studies*, *2*(4), 419-427.

Zmudy, M. H., Curtner-Smith, M. D., & Steffen, J. (2009). Student participation styles in adventure education. *Sport, Education, and Society, 14*(4), 465-480.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Designing Educational Computer Game for Human Circulatory System: a Pilot Study

Jatuput LOKAYUT[a] & Niwat SRISAWASDI[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
* niwsri@kku.ac.th

Abstract: Recently, digital games have been becoming popular and integral part of our society. The benefit of digital games is not only for fun but also for supporting learning by adding educational purpose into the games called educational games. In this study, we identified 12th grade students' alternative conceptions of human circulatory concept by using two-tier test to design a game served as their alternative conceptions-based learning material. The tests consist of multiple-choice and confident scale. The results show that the students hold alternative conceptions (false-positive and false-negative group) in every concept of human circulatory system, especially homeostasis. From students' alternative conception findings, in this study, an educational computer game was designed and developed, for example the Blood Donor game. To study the effectiveness of Blood Donor game, we explored the 11th grade students' motivation and perception of learning before providing them game (pre-test) and perception after taking game (post-test). The results show that game can enhance students' perceptions. Moreover, students' motivation did not correlate to their perception, indicating that the developed game could be used for students who have low or high motivation in science.

Keywords: Game-based learning, digital games, alternative conceptions, science motivation, perception to learn

1. Introduction

Digital games or computer games are programs or software that created for entertainment purpose (Rollings and Adams, 2003). Recently, digital games have been becoming popular and integral part of our society, especially children or younger generation who like to play game as a favorite activity. To apply the game to educational system, teachers and educators have attempted to find the new ways of teaching by adding educational purpose into the games which is called educational games, this approach namely game-based learning (Tang, Hanneghan and Rhalibi, 2009).

The benefit of digital games is not just for fun, it can be used to support learning. Papastergiou (2009) reported that using game in high school can promote students' of computer concepts and more motivational than non-game approach. Another study also reported that educational games can promote engagement and learning for students with special learning needs (Ke and Abras, 2013). The comparison effect of games indifferent course showed that students and teachers in the game-based course provided more reasons for student motivation along with more desirable, more helpful and less hindering aspects than students and teachers in the non-game-based course (Gunter and Hess, 2013). In addition, students who play an educational video game expressed all of the projected twenty- first century skills while being engrossed in the embedded science content (Annetta, Cheng and Holmes, 2010).

Human circulatory system, which is one topic in biology course, is difficult to learning due to its complex, too much information, and sometimes students' alternative conception come from textbook illustrations (Buckley, 2000). This is a reason why educators have attempted to seek a new approach to teach biology, such as active learning and student-centered pedagogy (Armbruster, Patel, Jhonson and Weiss, 2009). Oblinger (2003) suggested that actively engaged students insist that education must be more than the conventional PowerPoint classroom lecture where information is poured into their heads and regurgitated onto worksheets. To this end, knowing students' alternative conception is useful when designing the learning material. Consequently, students' alternative conceptions of human circulatory system are needed to investigate for served as an input to designing a digital game. The developed game is served as a tool for supporting students' learning on human circulatory system topic. Therefore, the students' perception and the correlation between students' science motivation and perception are challenging to examine in this study.

2. Study 1

2.1 Participants

A total of 31 students in the twelfth grade from local public school at the northeastern region of Thailand were recruited to respond their conceptual understanding about human circulatory system by taking a conceptual test including 16 items for 40 minutes. All of them are women and age ranging from 18-19 years old. They learned about human circulatory system before participating in this study. They did not prepare themselves for taking this test. Such that, it is reasonable for seeking their alternative conception on the topic.

2.2 Research Instrument

The two-teir concept test about human circulatory system was used in this study. We developed frist teir as a multiple-choice by applying the study of Sungur and Tekkaya (2003) to investigate students' alternative conceptions. The second teir is confidence scale (McClary and Bretz, 2012), ranging from 0% (just guessing) to 100% (absolutely confident), to examine students' confident in their response and facilitated an analysis about students' alternative conceptions. Our concept test consist of 16 items and main concept is classified into 4 concepts including blood, heart, blood vessels, and homeostasis.

2.3 Data Collection and Analysis

To design and implement the games, we investigated students' alternative conceptions about human circulatory system at first step. Students were tested by taking two-tier conceptual test including 16 items for 40 minutes without preparing themselves for taking an examination. The students have to select only one correct answer form four options and make "X" mark on the scale to assign their confidence about answer in each item. Students' answer and their confidence were classified into 4 groups, true-positive (TP) for correct answer with confidence more or equal 50%, true-negative (TN) for correct answer with percent of confident below 50%, false-positive (FP) for incorrect answer with confidence more or equal 50%, and false-negative (FN) for correct answer with percent of confident below 50%.

2.4 Results and Discussion

The aim of developing two-tier test is to identify alternative conceptions in which students have about human circulatory system concept. The results from this test show that the students hold alternative conceptions for every concept such as blood, heart, blood vessels, and homeostasis. List of students' alternative conceptions was shown in Table 1. Considering Table 1, students' alternative conceptions about blood concept shows that 80.65% of them hold alternative conceptions in which biconcave shape allows red blood cells to be in close contact with body cells, 32.26% of them believed that some of the blood stays inside blood tube and some of it leaves the blood tubes and baths the cells. In other concepts, 38.71% of them thought that blood goes into the heart on one side and leaves the other sides and goes to all parts of the body, low blood velocity in capillaries is due to material exchange through capillaries (58.06%), in systemic circulation, percentage of blood volume in the arteries and capillaries is equal, which is greater than that of veins (29.03%), all plasma proteins catalyze reactions in blood under normal physiological conditions. Similar to the study of Özgür (2013), he found that student hold alternative conceptions about blood circulatory system, such as heart, blood, blood circulation, and blood transfusion.

Table 1. List of students' alternative conceptions about human circulatory system.

Concept	Students' alternative conceptions	% (n=31)
Blood	Fat is not found in plasma.	22.58
	Vitamins uric acid is not found in plasma.	29.03
	Uric acid is not found in plasma.	16.13
	Biconcave shape allows red blood cells to hold more hemoglobin.	16.13
	Biconcave shape allows red blood cells to be in close contact with body cells.	80.65
	Serum is plasma to which necessary nutrients for a patient.	16.13
	Serum is storage from of plasma.	29.03
	Serum is house's plasma.	12.09
	Blood stays inside blood tubes that go to and from the cells.	22.58
	Blood leaves the blood tubes and bathes the cells.	3.23
	Some of the blood stays inside blood tube and some of it leaves the blood tubes and baths the cells.	32.26
Heart	Ventricular filling occurs mostly during atrial contraction.	35.48
	Ventricular filling occurs mostly during contraction of atrioventricular valves.	25.81
	Blood goes into the heart on one side. Blood leaves the other side and goes to all part of the body.	38.71
Blood vessel	Low blood velocity in capillaries is due to their small diameter.	9.68
	Low blood velocity in capillaries is due to material exchange through capillaries.	58.06
	Low blood velocity in capillaries is due to their long distance from the heart.	22.58
	In systematic circulation, percent of blood volume in the arteries, capillaries, and veins is equal.	9.68
	In systemic circulation, percent of blood volume in the arteries and capillaries is equal, which is greater than that of veins.	29.03
	In systemic circulation, the percent of blood volume in the arteries is the highest, while the blood volume in the veins is the lowest.	25.81
Homeostasis	Under normal physiological conditions, all plasma proteins are used to meet cells' amino acid needs.	6.45
	Under normal physiological conditions, all plasma proteins help material transport across capillaries.	19.35
	Under normal physiological conditions, all plasma proteins catalyze reactions in blood.	38.71
	Glucose leaves the blood in capillary mainly by diffusion through endothelial cells.	29.03
	Glucose leaves the blood in capillary mainly by fluid movement through endothelial cells at arteriole end.	41.94
	Glucose leaves the blood in capillary mainly by diffusion through narrow opening between endothelial cells.	19.35

From analysis of students' answers shown in Table 2, we found that 18.55% and 33.07% were classified as FP and FN group respectively, this suggests that half of students hold alternative conceptions and 14.49% hold incomplete scientific conceptions about blood concept. In the same way, more than half of students hold alternative conceptions about heart, blood vessel, and homeostasis concept.

Table 2. Students report self-confidence in each item on multiple-choice and scale of 0% to 100% of confident.

Content	Item	TP		TN		FP		FN	
		N	%	N	%	N	%	N	%
Blood	1	4	12.90	6	19.25	9	29.03	12	38.71
	3	22	70.97	3	9.68	3	9.68	4	9.68
	4	5	16.13	7	22.58	6	19.35	13	41.94
	16	11	35.48	2	6.45	5	16.13	13	41.94
		Mean	33.87	Mean	14.49	Mean	18.55	Mean	33.07
Heart	8	8	25.81	4	12.90	11	35.48	8	25.81
	9	6	19.35	3	9.68	10	32.26	12	38.71
	13	3	9.68	8	25.81	3	9.68	17	54.84
	15	17	54.84	2	6.45	11	35.48	1	3.23
		Mean	27.42	Mean	13.71	Mean	25.81	Mean	30.65
Blood vessel	5	1	3.23	1	3.23	25	80.65	4	12.90
	10	6	19.35	3	9.68	10	32.26	12	38.71
	11	2	6.45	1	3.23	11	35.48	17	54.84
	12	3	9.68	8	25.81	3	9.68	17	54.84
		Mean	9.68	Mean	10.49	Mean	39.52	Mean	40.32
Homeostasis	2	1	3.23	10	32.26	4	12.90	16	51.61
	6	1	3.23	1	3.23	11	35.48	18	58.06
	7	3	9.69	3	9.68	7	22.58	18	58.06
	14	1	3.23	3	9.68	8	25.81	19	61.29
		Mean	4.85	Mean	13.71	Mean	24.19	Mean	57.26

3. Study 2

3.1 Participants

To examine student's science motivation and perception, the participants were 50 eleven-grade students who never learned about human circulatory system before and lack of experience of learning through educational computer game.

3.2 Material about Game

To construct a game as a learning material, we classified the main concept into 3 concepts such as heart, component and flowing of blood, and blood group. We have developed a digital game for assisting learning in each concept. The Blood Donor is an example of game in this study. The overall of this game is shown in Figure 1. In this game, the students get the mission to find who can give their blood for the patient in the scenario. Patient's blood group is random at the beginning of the game. Students have to type blood group everybody in the hospital to know and decide whether the patient can receive blood from this person by dropping blood into test tube of which antibody. Students can click at the test tube for get more details of molecular level. After typing blood group, the students have to check a list of name of people in the hospital with the nurse. The mission will be completed until the students can check list of name for all blood donor correctly.

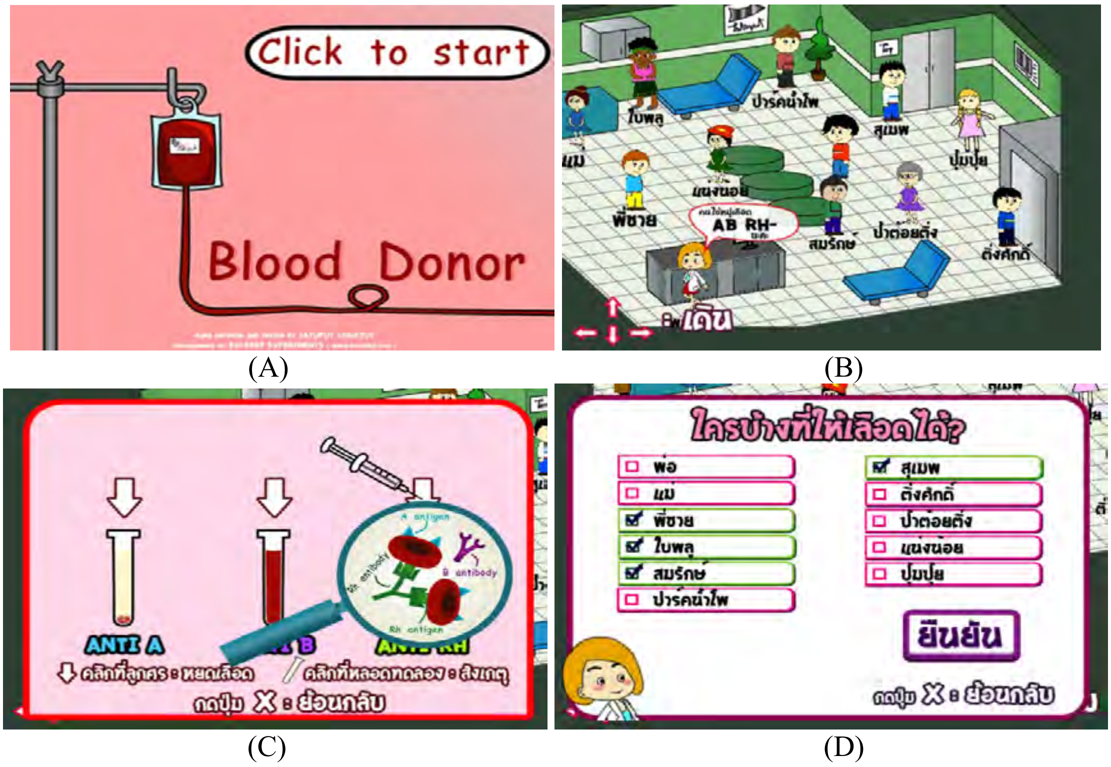


Figure 1. Illustrate example of game “Blood Donor” for teaching about concept of blood type: (A) shows home screen of the game; B shows people in the hospital that need to know their blood group; (C) shows typing blood by dropping blood into test tube; (D) shows checking list of people in the hospital.

3.3 Research Instruments

3.3.1 Student’s Perception

To study students’ perception of the educational computer game, we develop the questionnaire with 6 scale, such as perceived learning (PL), flow (FL), enjoyment (EJ), ease of use (EU), usefulness (UF), and satisfaction (SA) (Cheng, 2014, Barzilai and Blau, 2014), see Table 3.

Table 3: Scale and sample items of the perception questionnaire.

Scale	α	Example of item
Perceive learning (PL)	0.80	I learned a lot from the game. The game added to my knowledge.
Flow (FL)	0.82	I learned a lot from the simulation. The game added to my knowledge.
Enjoyment (EJ)	0.75	I enjoyed the game. I had fun playing the game.
Perceive ease of use (EU)	0.74	It is easy for me to learn how to use game. The user interface of game is easy to use.
Perceive of usefulness (UF)	0.84	Game can help me learn more effectively. Game can improve my course performance.
Perceive of satisfaction (SA)	0.77	I feel comfortable to use game. I enjoy the experience of using game.

Participants were required to consider each possible reason for educational game and rate how important it was for them by using a 5-point Likert scale (1-strongly disagree; 2-disagree; 3-neutral; 4-agree; 5-strongly agree). For overall items have a very good reliability.

3.3.2 Student's Science Motivation

A science motivation questionnaire was developed from the study of Glynn, et al. (2011) by translated into Thai version. They are 5 categories of this questionnaire and 5-point Likert scale, such as intrinsic motivation (IM), self-determination (SDT), self-efficacy (SEC), career motivation (CM), and grade motivation (GM). Reliability of the questionnaire was tested Cronbach's alphas in each category which were 0.79, 0.81, 0.89, 0.81 and 0.85 for IM, SDT, SEC, CM and GM, respectively, indicating that Thai version's Science motivation questionnaire was found a good reliability.

3.4 Data Collection and Analysis

3.4.1 Students' Perception

Students were investigated by perception questionnaire before providing the educational games (pre-test). After that, students were asked to play the game Blood Donor for 20 minutes and were investigated perception again by using the same questionnaire (post-test). To compare both pre- and post-test, the pre- and post-test scores were analyzed by using pair *t*-test using SPSS program.

3.4.2 Students' Science Motivation

To explore students' science motivation, the questionnaire was used before providing the educational games. There are 25 questions on the questionnaire and each item rates the students' perceptions of the game using five-point scale ranging from "least" (1 point) to "most" (5 points). Students' science motivation was determined the correlation with post-test of students' perception by using the educational computer game.

3.5 Results and Discussion

3.5.1 Students' Perception

After playing game, students' perception before and after were analyzed by using pair *t*-test. The result shows that post-test score is higher than pre-test significantly in each scale as shown in Table 4.

Table 4: Students' perception in each scale.

Scale	IM	CM	SDT	SEC	GM	PL	FL	EJ	EU	UF	SA
IM	1										
CM	0.49^{**}	1									
SDT	0.56^{**}	0.52^{**}	1								
SEC	0.63^{**}	0.34^{*}	0.43^{**}	1							
GM	0.52^{**}	0.42^{**}	0.22	0.43^{**}	1						
PL	-0.26	-0.12	-0.20	-0.09	-0.11	1					
FL	-0.10	-0.23	0.02	0.10	0.15	0.61^{**}	1				
EJ	-0.22	-0.07	-0.25	-0.01	-0.36	0.77^{**}	0.57^{**}	1			
EU	0.07	0.18	0.14	0.08	-0.03	0.24	0.43^{**}	0.16	1		
UF	-0.10	-0.28	-0.14	0.07	0.03	0.42^{**}	0.55^{**}	0.47^{**}	0.20	1	
SA	-0.24	-0.26	-0.22	-0.01	-0.14	0.33^{*}	0.40^{**}	0.56^{**}	0.15	0.57^{**}	-
Mean	17.36	16.28	15.72	14.00	16.76	13.90	15.96	10.60	9.45	10.14	10.82
SD	3.15	2.87	2.42	2.55	3.01	1.81	2.10	1.77	1.73	1.44	1.76

$^{*}p < 0.05$, $^{**}p < 0.01$

This indicates that the progression of students' perception increase after playing the developed educational computer game. This finding conforms to the study of Srisawasdi (2012) that using technology can improve students' perception of learning.

3.5.2 Correlation between Students' Science Motivation and Perception

Pearson's correlation was used to investigate correlation between science motivation (IM, SDT, SEC, CM, and GM) and perception (PL, FL, EJ, EU, UF, SA) in this study as shown in Table 5. Regarding Pearson's correlation analysis of each scale in science motivation, we found that all scale related together except self-determination which do not relate to grade motivation. This results reveal that grade in science subject is important for them even though they have different level of self-determination. From the analysis of perception, perceived learning showed significantly related to flow, enjoyment, usefulness, and satisfaction. In addition, students' perception of flow was positively related to their change in enjoyment, ease of use, usefulness, and satisfaction.

The analysis of correlation between students' science motivation before playing the developed educational computer game and perception after playing the game showed that intrinsic motivation, self-determination, self-efficacy, career motivation, and grade motivation were related to their perceived learning, flow, enjoyment, ease of use, usefulness, and satisfaction. These results indicate that the developed game, The Blood Donor, could be used for all students who have low or high motivation in science. In addition, game tasks can improve students' motivation, engagement, and problem-solving performance (Eseryel, Law, Ifenthaler, Ge, & Miller, 2014).

Table 5: Descriptive statistic and correlation for perception and science motivation

Scale	Mean (SD)		*t*
	Pre-test	Post-test	
Perceived Learning	11.52 (2.37)	13.90 (1.81)	5.79*
Flow	14.04 (2.59)	15.96 (2.10)	4.41*
Enjoyment	9.06 (1.92)	10.60 (1.77)	3.92*
Ease of use	8.18 (1.99)	9.52 (1.79)	3.60*
Usefulness	8.70 (2.10)	10.14 (1.44)	4.01*
Satisfaction	9.04 (2.47)	10.82 (1.76)	4.12*

$^{*} p < 0.01$

4. Conclusion

This study surveys students' alternative conception about human circulatory system. After that the educational computer game was designed and developed based on those alternative conceptions. The finding of this study show that (1) students still hold alternative conception about the concept of human circulatory system even if they learned before; (2) our educational computer game, the Blood donor, can be used for improving students' perception to learn; (3) students' perception dose not relate to science motivation. To this end, the Blood Donor game which is designed based on students' alternative conceptions can assist students who have low or high motivation in science to learn. However, to address students' learning performance we are going to study about the effect of using educational computer game on students' conceptual understandings.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Annetta, L.A., Cheng, M., & Holmes, S. (2010). Assessing twenty-first century skills through a teacher created video game for high school biology students. *Research in Science & Technology Education*, *28*, 101-114.

Armbruster, P., Patel, M., Jhonson, E., & Weiss, M. (2009). Active Learning and Student-centered Pedagogy Improve Student Attitudes and Performance in Introductory Biology. *Life Sciences Education, 8*, 203-213.

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, *70*, 65-79.

Buckley, B. (2000). Interactive multimedia and model-based learning in biology. *Int J Sci Educ, 22*, 895–935.

Cheng, G. (2014). Exploring students' learning styles in relation to their acceptance and attitudes towards using Second Life in education: A case study in Hong Kong. *Computers & Education*, *70*, 105–115.

Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An Investigation of the Interrelationships between Motivation, Engagement, and Complex Problem Solving in Game-based Learning. *Educational Technology & Society*, *17*(1), 42–53.

Glynn, S., Brickman, P., Armstrong, N. and Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159-1176.

Gunter, G., & Hess, T. (2013). Serious game-based and nongame-based online courses: Learning experiences and outcomes. *British Journal of Educational Technology*, *44*, 372-385.

Ke F., & Abras, T. (1013). Games for engaged learning of middle school children with special learning needs. *British Journal of Educational Technology*, *44*, 225-242.

McClary, L.M., & Bretz, S.L. (2012). Development and Assessment of A Diagnostic Tool to Identify Organic Chemistry Students' Alternative Conceptions Related to Acid Strength. *International Journal of Science Education,* 1-25.

Oblinger, D. G. (2003). Boomers, gen-xers, and millennials: Understanding the new students. *Educase, 38*(4), 37-47.

Özgür, S. (2013). The Persistence of Misconceptions about the Human Blood Circulatory System among Students in Different Grade Levels. *International Journal of Environmental & Science Education*, *8*, 55-268. Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computer & Education*, *50*, 1-12.

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education*, *52*(1), 1-12.

Rollings, A., & Adams, E. (2003). *Andrew Rollings and Ernest Adams on Game Design*: New Riders Publishing.

Srisawasdi, N. (2012). Introducing Students to Authentic Inquiry Investigation Using an Artificial Olfactory System. In K.C.D. Tan & M. Kim (Ed.), *Issues and Challenges in Science 93 Education Research: Moving Forward* (pp. 93-106). *Issues and Challenges in Science 93 Education Research: Moving Forward.*

Sungur, S., & Tekkaya, C. (2003). Students' Achievement in Human Circulatory System Unit: The Effect of Reasoning Ability and Gender. *Journal of Science Education and Technology, 12*, 59-64.

Tang, S., Hanneghan, M., & Rhalibi, A.E. (2009) Introduction to Game-based Learning. In T. Connolly, M. Stansfeild, & L. Boyle (Ed.), *Game-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices* (pp.1-17). New York: Information Science Reference

The Curriculum Design of Nutrition and Food Safety Game for Elementary School Students

Chun-Heng Lin[a], Tsung-Yen CHUANG[a*], Chung-chiann Chuang[b], Fang-Ying Tu[a] & Hua-Hsiang TSENG[a]

[a] *Department of Information and Learning Technology, National University of Tainan, Tainan*

[b] *Department of Early Childhood Education, Taiwan Shoufu University*

*chuangyen@mail.nutn.edu.tw

Abstract: Because of food serves energy and nutrition for human, human beings extremely rely on food. Furthermore, micronutrients help our bodies function normally and keep us healthy. Without micronutrients, we may lose our lives. Recently, as many food safety issues revealed, we really should pay more attentions on how to eat healthily. Because tainted food caused by dishonest or ignorant men may damage our health even severely. On account of those reasons above, this study basically aims at K-5 and K-6 elementary school students, the important stage of human lifestyle building, to develop a RPG game played on mobile device. Hope the children can enjoy the game and easily applied what they've learned in the game to their daily lives.

Keywords: Nutrition education, food safety education, game-based learning, mobile learning, RPG game

1. Introduction

Food is essential to keep us alive. Recently, the scandals about illegal ingredients and additives of food are exposed one after another in Taiwan. In addition to the tainted food, there are also many dishonest food business were revealed. In order to make a profit, the food industry owners reduce the production cost by replacing high-quality elements with low-quality ones. Yet, even worse, some of them use harmful substitutes to reach their goals. Those food safety incidents passed down a message that people who run the food business have put morality far behind. This is a serious problem because human beings can't live without food. The contaminated food affects children more greatly (Meysenburg, Albrecht, et al., 2014). Under the threat of food safety problems, children should equip with sensible discernment in their choice of food and good nutrition education to fight for their own healthy growth.

2. Background

2.1 Nutrition Education for Elementary Students

Except for energy, human begins also needs to get nutrients from food to stay healthy. According to the amount of nutrition which human body needs, the nutrient falls into two groups. One is macronutrient, the other is micronutrient. Macronutrient, which provides calories, mainly contains carbohydrate, protein and fat. Micronutrients, including vitamins and minerals, are essential for human body (Crowe, et al., 2011). Human being only needs a little micronutrient. They can regulate physical function, but cannot be produced by our own bodies. The knowledge of micronutrient is extremely important for us to live healthily. Micronutrients can be found in variety of food. Nevertheless, only intake of micronutrients from fruit and vegetable leads to less burden to human bodies (Herrero, 2011).

Most of nutrition curricula are in elementary school for K-5 and K-6 Students. Those students are at the mental and physical growth stage (Wu, 2012). They have tenacious memory at the same time.

Also, they have enough capability to understand the systematic structure of nutrition knowledge. Some research indicated that the habit or thought students built in this stage will affect and last the rest of their lives most (Fu & Jien, 2012). There are many types of micronutrients with various sources and different functions. In existing materials, micronutrient knowledge is represented by text description, table with text or adding pictures in table beside the text. However, neither of them is easy for students to memorize and build the concept of micronutrients. In order to get the students' attention, this study is trying to develop the effective way to teach them micronutrient knowledge and to ease their fear with fruits and vegetables.

2.2 Food Safety Education

However, there're some studies indicate that it's risky to get nutrients form natural crops (Martin-Belloso & Fortuny, 2011). According to these studies, there may be toxic residue left on the plants. As we have the fruit or vegetable for health, those contaminants could come with the food to enter our bodies. Some of them may make us feel physically uncomfortable. Even worse, they may cause death. In light of the situations above, food safety is a significant problem when people try to eat healthily. It is possible that people who are not in the food producing process pollute food the most. Some factories were built near by the farmland. But they didn't deal with the pollution they made well. This is a serious problem because pollution form factory usually contain heavy metals, which can't be metabolized by human body (Marshall et al., 2007). The factory managers may not notice that they also have the responsibility to agricultural production, yet their emissions cause immeasurable damage to the crops. Therefore, we think everyone should pay attention on this issue and treat this issue with morals.

According to the study of Haapala and Probart (2004), they found the middle school students are not with enough knowledge, perceptions and behaviors of food safety. Once students get into adolescence, their chance of accessing to the drugs, tobacco and alcohol will increase (Choquet et al., 2004). That's why we think K-5 and K-6 students are the ideal learners of food safety curricula. Before they become teenagers, they should learn how to live a healthy life. After all, everyone only have one chance to become adult. So growing with health is crucially significant. Instead of abstruse knowledge about food safety, we think that the upstanding mind is which they need the most. They should learn to put themselves in anyone else's shoes. The characteristics mentioned above could be cultivated by simulation. Students can get their own experience when undergo the simulations in the game story. Moreover, we hope that they may adopt the experience on their real lives.

2.3 Game-based Learning

Micronutrient knowledge is something abstract to children. Food safety is also something hard for them to realize. To overcome the abstraction and complication of learning content, we must find the way to concretize it and maintain learner's attention on it at the same time. Before solving those problems above, there is one thing we would face first. How to attract K-5 and K-6 students to learn these contents? The "motivation" leads us to game. There are many cases we can follow (Papastergiou, 2009; Prensky, 2003). Combine teaching with game can always gain more interests from children. Using game to teach can bring them fun and reduce their stress when learning (Kanthan & Senger, 2011). Especially, the RPG game can pull students into the game world easily. They will be the super mayor character of a game. They'll treat what they face in the game extremely important. Therefore, we think the RPG game is better than other candidates. Like a simulation, children meet some events that happened in real life when playing game. They learn how to deal with them by the guidance of assignments. It is a good method to keep danger away and save resource such as time, money and so on when simulating in the game (Pidd, 1992).

The main point of instructional game is the learning content. There is a model of games and learning showed in Figure 1 (Garris, Ahlers & Driskell, 2002). At first, we should put learning content and game features together. Design a reasonable and joyful game. Once students get into the game, they start the game cycle. They receive stimulus from game and produce their own views. Then, they'll reflect the thought in mind on their behavior. After catching the stimulus from player's behavior, the game also responses back. Students will undergo the game cycle again and again and get their own experience. We evaluate their performance to reach their learning outcomes. This is the model gave us confidence to teach our subject by game. Although the content in the subject is not friendly at all, we still can make it adorable to children. Blend them into the game and intersperse some charming stories with lovely game scene. In contract of the fantastic image, the knowledge students would learn in game is true in real world. Instead of memorizing the contexts in the book, we believe this is a good way for students to realize numerous micronutrients and their various source and function (Kiili, 2005).

Figure 1. The games and learning model.

3. Method

3.1 Game Design

Our goal is to equip children with enough nutrition knowledge and clear-sighted view of food safety issues. At the same time, attracting children to keep learning the content is our challenge. That is why we tried to pass off the instruction as game. Children can enjoy the fun and feel less pressure when learning by game. Especially when they played the RPG games, they only focused on their missions and the role they played. Totally immerse in the game, enjoy characteristics of game. Children will see the game as their second life. Therefore, we developed the game described below to meet our goal. In the game system part, we'll talk more about the entertainment which can distract people's attention from learning. The game story is the background we create to match the subject. The last is the part of integration, explaining those ideas we use in embedding the learning content.

3.1.1 Game System Development

We build a RPG game by Unity 4.3.4. In addition to basic RPG game features such as role training, level bonuses, interactions with NPCs and game tasks solving, we added duel and collecting book. Basic features can challenge players and invite them to play. The collecting book is another shot at the game for players. Once children meet the fruit or vegetable in the game and prevail it, it will be recorded in the book automatically. Collecting the items in the book all is the motivation for players to continue the game. Besides, the duel gives the game more appeal and excitement. We need to sustain children’s interest, so the RPG game is our choice to involve players in learning content. We want children to forget they’re learning and enjoy totally fun. For this purpose, we create a lovely world which can be seen in Figure 2 and decide to run our game on mobile devices. M-learning (mobile learning) means that learning will no longer be limited by time or location (Hwang & Chang, 2011). There are two main properties in mobile learning. One is arbitrary learning place, the other is short time interval. With these two properties, children can play the game anywhere and anytime.

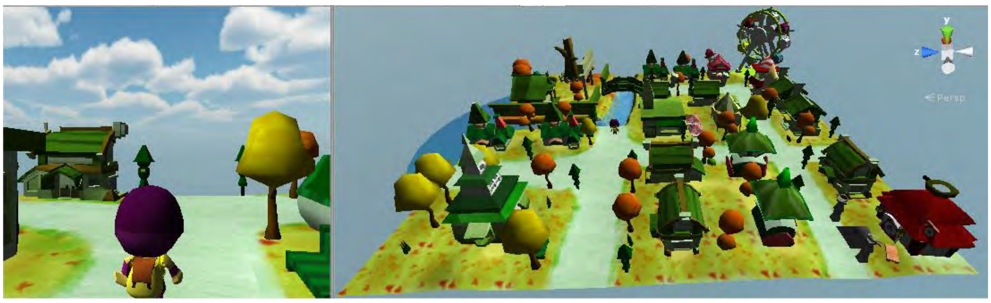


Figure 2. The world in the game.

They don't need to stay steadily. With mobile devices, they could learn anytime, anywhere. They could lie down or sit casually when holding light device and learn. They play the game in their daily lives and easily connect what they've learned to their real lives (Chu, Hwang, Tsai, & Tseng, 2010). They may be interrupted by others when playing game on mobile device. But the features of RPG game could pull the players back fast. The familiarity of plants could urge students to try them for food. Meanwhile, the short time interval could also reduce the learning burden of children. For example, they could play the game when taking a break in the school or waiting for dinner prepared by mother. Give them more flexibility to connect the content of the game to real world.

3.1.2 Game Story Design

To close the distance between children and plants, the background of our game is a kingdom lived on fruits and vegetables. There are soldiers called "armed-grape", whose duty is saving people's life and health. They're trained by the government since the incredible event changed the kingdom dramatically. Some of plants can move like animals. They started to protect their own kind from human. People couldn't get essential micronutrient from them, so the weakness and disease rage on. "A-Ni" is the hero of our game. At the beginning, A-Ni is one of new recruits undergoing training and sees fighting for people as his or her own destiny. After the training, A-Ni steps on the journey to help people and enhence the ability. A-Ni will be promoted as more and more assigned tasks being solved. For this reason, A-Ni has more power to deal with classified problems. In the meanwhile, some secret and evil plans of government are available to A-Ni, too. At the end, we could only rely on players to lead A-Ni to make wise decision and do sensible action.

We use the tasks to guide players. For instance, players may receive three assignments first. One is to collect the material, such as vitamins. Another might ask them to satisfy the NPC's needs. And then they meet some plant monsters on the road. They have to prevail the monster to take the material and record the plant into the collecting book. The material form that monster may just meet the NPC's need. Therefore, players solve the tasks at the same time and continue their trip. Of course, they still receive another assignment in the following adventure. Analyze the example mention above, there're three types of assignments, material or plant collecting, the requirements of NPCs and achievement making. We cross these assignments to build the situation in the game. Players will be required to fight or pick the plants on the road and follow the guidance to interact with NPCs. Players will need to collect something for NPCs, or heal them by knowledge and experience. As for achievement making, players can know their own state, like the ability of the role, the progress of collection and journey.

3.1.3 Learning Content Integration

Children will play the hero of the game to help the kingdom back to placidity. Follow the assignments, we set micronutrient knowledge and food safety issues in the story. Children need to organize the information and use those items in their hands to solve the problems. Every assignment is designed for learning contents. The assignments of material or plant collecting are mainly related to micronutrient knowledge. The interaction with NPCs is the most flexible part. It can embed in the food safety cases and the events of micronutrient deficiency or excess. In addition, children should understand their learning pace or condition in the game. That's why we design the assignments of achievement making for a short-term target. The achievement reminds children that they have already had some knowledge or experience. By the similar cases in the story, players can realize the reason behind and know the key points they should put their eyes on when facing food safety problems. The simulations are merged into game tasks to teach children the skills to find out the solutions of food safety cases step by step.

Further, we would prefer to see them adopt those skills on their realistic life and prevent people from the threat of unsafe food. To increase micronutrient learning effects, collecting book and fighting effect are designed. Players can record the plants they have met. There is much information in the collecting book, such as the name and the short introduction of the plant, the micronutrient in the plant and the effects they provide to human body. The description above can be displayed more clearly by Figure 3. The content of collecting book is edited by the real information of those plants. In the game, it could also help player to look for solutions of the assignments. By this way, children could know much about fruits and vegetables and understand how to get healthy with them. The other design to help children learning micronutrient knowledge is the fighting effect, like Figure 4. Players should fight to the plant monsters to get the micronutrients they need. As fighting, once the player hit the plant monster, the micronutrients which the plant contains will appear and drop down from it. The micronutrients are showed by text. The times of every text showed in fight are set on their percentage in the plant.

Figure 3. The schematic drawing of collecting book.

Figure 4. The schematic drawing of fighting effect.

3.2 Design of Experiment

The target of our game is K-5 and K-6 students. We want to get more details about children's learning in our game, so the questionnaire comes to us. There are two parts in the questionnaire. One is about the performance of students' learning, the other is the system usability scale. In the game, we try to build the knowledge and concepts of micronutrient and food safety. Therefore, the questionnaire of students' learning performance is divided into two themes, micronutrient and food safety. Students will do this questionnaire before and after playing the game to evaluate the effect of our learning game. The system usability scale will be done only after the game to collect students' thoughts of the game and understand that if the game actually can help students to learn something. We want to know whether or not the results meet our previous expectation.

The experiment will be implemented in the Health and Physical Education class of K-5 and K-6 students. The duration is about one month. We'll occupy two time HE & PE class time. One is at the beginning of the experiment. We must to introduce our game, tell students how to play the game and let them do the pretest. Then, they can play the game whenever and wherever they want in the next month. After one month, we'll let them do the posttest in the class. In addition to questionnaire, we'll do some interviews with students who had participated in experiment. We believe that children will give us more feedback in the interview.

4. Conclusion

K-5 and K-6 students are at the critical stage of lifestyle building. As we mentioned previously, our plan is to equip children at that stage with the knowledge and concept of micronutrient and food safety. To reach the purpose, game becomes our conclusion. Game can make fun for people. There are many types of games. The RPG game is our choice to combine with the learning content we prepared. Students can solve assigned tasks and learn something through game. Base on the real food safety cases, we design tasks in the game story. Students can do the simulation in the game world. They enrich their own experience when playing the game. Give them the capability to protect themselves and anyone else. The information of fruits and vegetables is also something true in the game. We hope students to be close to plants for food and have correct micronutrient knowledge about plants. There are some ideas being developed to aim our target, such as collecting book and fighting effect. Help students to improve their learning effects.

Everyone lives on food. It is important for us to care more about something we get into our bodies. This game is the brainchild of everyone in our team. We try to deliver some useful knowledge and information about food to K-5 and K-6 students. It is our hope that they could learn these serious things, realize them and deal with them. Further, learning by mobile devices could help them practice all they learn in game to daily life more flexibly. The final goal of ours is to help people to guard their own health and others. At the end, we all have a clean, safe and healthy life.

Acknowledgements

The research reported in this paper has been supported in part by the National Science Council in Taiwan under the research project number NSC 102-2511-S-024 -006, and MOST 103-2511-S-024-009.

References

Pidd, M. (1992). *Computer simulation in management science (3rd Ed.)*. New York: John Wiley and Sons, Inc.

Haapala, I., & Probart, C. (2004). Food safety knowledge, perceptions, and behaviors among middle school students. *Journal of Nutrition Education and Behavior, 36*(2), 71-76.

Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*(4), 1023-1031.

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, J. C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science course. *Computers & Education, 55*(4), 1618-1627.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming, 33*(4), 441-467.

Meysenburg, R., Albrecht, J. A., Litchfield, R., & Ritter-Gooder, P. K. (2014). Food safety knowledge, practices and beliefs of primary food preparers in families with young children. A mixed methods study. *Appetite, 73*, 121-131.

Martin-Belloso, O., & Fortuny, R. S. (2011). *Advances in fresh-cut fruits and vegetables processing*. CRC Press.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1-12.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE), 1*(1), 21-21.

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education, 8*(1), 13-24.

Kanthan, R., & Senger, J. L. (2011). The impact of specially designed digital games-based learning in undergraduate pathology and medical education. *Archives of pathology & laboratory medicine, 135*(1), 135-142.

Marshall, F. M., et al. (2007). Contaminated irrigation water and food safety for the urban and peri-urban poor: appropriate measures for monitoring and control from field research in India and Zambia. *Incpetion Report DFID Enkar, 8160*.

Choquet, M., Morin, D., Hassler, C., & Ledoux, S. (2004). Is alcohol, tobacco, and cannabis use as well as polydrug use increasing in France? *Addictive Behaviors, 29*(3), 607-614.

Fu A. H., Jien J. J. (2012). Perception of Body Shape and Factors that Affect Dietary Behaviors of Fifth and Sixth Graders in Taiwan. *Nutritional Sciences Journal, 37*(1), 36-50.

Wu, G. Y. (2012). Research on the Nutrition Knowledge, Nutrition Attitude, and Dietary Behavior of Elementary School Children at Upper Grades in Taichung City.

Crowe, F. L., Roddam, A. W., Key, T. J., Appleby, P. N., Overvad, K., Jakobsen, M. U., ... & Riboli, E. (2011). Fruit and vegetable intake and mortality from ischaemic heart disease: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Heart study. *European heart journal, 32*(10), 1235-1243.

Herrero, M., Gerber, P., Vellinga, T., Garnett, T., Leip, A., Opio, C., ... & McAllister, T. A. (2011). Livestock and greenhouse gas emissions: The importance of getting the numbers right. Animal Feed Science and Technology, 166, 779-782.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Thinking as a Pleasure: Tactics to Design Digital Educational Games from the Perspective of Board Games

Hercy N.H. CHENG
Collaborative & Innovative Center for Educational Technology, Central China Normal University, China
*hercycheng.tw@gmail.com

Abstract: Board games provide players with not only pleasure but also thinking. More specifically, it is observed that digital games tend to ask players to make quick judgment, demanding less consideration, while board games allow players to think more thoroughly in reasonable time. For this reason, this paper attempts to capture the features of board games, from which most modern digital games originate. When designing digital educational games, designers may consider the features so that players can have both pleasure and opportunities for thinking. Finally, this paper also suggests three design tactics to design digital educational games.

Keywords: Digital educational games, board games, choices

1. Introduction

Salen and Zimmerman (2004) defined a game as a system in which players engaged in an artificial conflict, resulting in a quantifiable outcome. Furthermore, the conflict was purposely produced by game rules with a goal that players attempted to achieve. Crawford (1984) described that conflicts were obstacles preventing players from easily achieving their goals. A puzzle, for instance, provides static and passive conflicts, while games have more dynamic and actively responsive conflicts, which come from the interactions between players and the other entities (e.g. interactive objects, non-player characters, human opponents or collaborators) in the game world.

From the perspective of education, however, a difficult conflict does not guarantee players' learning. As a matter of fact, most digital games adopt immediate conflicts, which may encourage players' fast reactions but hinder their thinking without sufficient time. Such digital games require players' skills of hand-eye coordination (Crawford, 1984), which can be improved by gaming over and over again rather than strategic reflection.

Unlike digital games, although board games lack well-simulated environment, most players still consider board games full of interactivity and interesting conflicts. On the other hand, board games have been widely adopted in classrooms as a part of curriculum (Hinebaugh, 2009), implying the essence of educability. Mayer and Harris (2010) also indicated that the board games might fit the abilities of the 21st century learners defined by American Association of School Librarians (2009). In practice, many recent studies have showed that board games may help students learn in various courses, for instance, mathematics (Ramani, & Siegler, 2008), physics (Smith, & Munro, 2009), health education (Lennon, & Coombs, 2007), and financial accounting (Gamlath, 2007).

Board games, according to their game goals, can be roughly classified into three categories: war games, race games and alignment games (Hinebaugh, 2009). The first one is war games, in which players aim at capturing and/or destroying opponents' units (Woods, 2012). Obviously, *chess* and *go* are both war games; so are some digital games like *Age of Empires*. Just like a war, the conflicts emerge if the units of two parties fight against each other. At this time, survivability becomes the first thing that players should consider. Moreover, players have to figure out how to avoid being attacked effectively and make choices to overcome enemies' defense.

The second category is race games, in which players aim at reaching a destination the fastest, such as *Chinese checkers* and *Chutes and Ladders*[1]. Alternatively, a variation of race games is to score the highest when the game ends. For example, *Monopoly*[2] can be considered as a race game because the player who has the most money in the end wins the game. Most strategy board games adopt such a game type, such as *Puerto Rico*[3] and *Agricola*[4]. These games usually allow players to obtain resources and transform them into scores. Such a transformation mechanism is basically a model of economy (Rollings, & Adams, 2003), facilitating players to consider which actions may provide an efficient way to transforming the least resources into the highest scores.

The third category of board games is alignment games, which require players to make a particular pattern, such as a line of pieces in *tic-tac-toe* or *Gomoku*. Another typical alignment game is *Mahjong*. For winning, players have to collect resources and decide whether the resources are needed or not. If necessary, players have to discard some resources, which involves lots of decisions.

Among the three categories, there is a common characteristic, which is to provide choices so that players may make decisions for their own. With choices, players may start to think which option is better than the others. After making a decision, players are motivated to expect its consequence and attached to the gameplay. In other words, the provision of choices may make players enjoy the thoughtful pleasure of board games

However, media somehow limits the game type. Most digital games are skill-and-action games rather than strategy games because of their aims at massive popularity (Crawford, 1984; Rollings, & Adams, 2003). Although digital games allow players to make decisions immediately, players usually do not have sufficient time to think. On the contrary, lots of board games, especially German games or eurogames (Woods, 2012), are strategy games, taking advantage of turn-based mechanism and allowing players to think more when they make decisions. It turns out that well-designed choices potentially facilitate players to think and have fun at the same time in games.

Therefore, this paper aims to study the feature of board games, choices, which may facilitate both pleasure and learning. The feature may be helpful to design digital educational games. Perhaps digital media may tempt or deceive game designers into overemphasizing fantasy without noticing the real enjoyment and potential educational functions. It is a good idea to explore the possibilities of board games in education.

2. Choices

Sid Meier, the game designer of *Civilization* series, believed that a good game was a series of interesting choices (see Rollings, & Morris, 2000), suggesting the importance of choices in a game. This section will introduce the concept of choices and discuss how choices may facilitate both thinking and pleasure.

2.1 Choices facilitate thinking

An opposite concept of choices is linearity. A game with no choices is merely a story with linear facts (Crawford, 1984). Furthermore, without meaningful choices, the outcome of a game becomes predetermined (Salen, & Zimmerman, 2004). Choices imply the possibility of different results after players invest their efforts, making a game dynamic. When players make choices in a game, they actually explore the relationship between the chosen actions and their results. In terms of computer science, the rules of a game construct a state machine, which describes the aforementioned relationship (Juul, 2005). If all possibilities of choices are considered, a game may be like a tree structure with a

[1] *Chutes and Ladders* originated from an India game, *Snakes and Ladders*, in the 16th century. In the game, on the way to the destination, ladders made players approach the destination while chutes made them away from it.

[2] *Monopoly*, originating from *The Landlord's Game* in 1904, was patented in 1935 by Charles Darrow and published by Parker Brothers.

[3] *Puerto Rico* was designed by Andreas Seyfarth and first published in 2002. The game required players to produce resources, which may be transformed into money and scores.

[4] *Agricola*, a farm-theme board game, was designed by Uwe Rosenberg and published in 2007. Players as farmers use resources to build their own farms, grow crops and keep animals, which may be transformed into scores in the end of a game.

result on every leaf node (Crawford, 1984). In a sense, playing a game is interacting with the state machine as well as exploring the tree of game results (Juul, 2005).

Such a game structure suggests that players should predict the results before choosing an action. Prediction may involve a series of cognitive process: a first and straightforward thought is to evaluate the value of every legal action. In other words, without other consideration, players tend to choose an action that may bring the greatest benefits. In an extreme case, if a game only offer obvious choices, the game has actually no choices at all and thus no challenges for players. For this reason, a game should provide players with choices that need think.

For offering comparable choices, a game can assign different costs to every action. By doing so, players may start to consider the values and their respective costs. An action with high value and high cost is not a necessarily better choice than that with lower value and lower cost. The costs include risks, which are outcomes aware of likely happening (Epstein, 1977; Salen, & Zimmerman, 2004). Risks are usually negative and so inexplicit that players need to deliberate any possible situations before making decisions.

Besides, human opponents provide more factors that should be considered in decision-making. In a competitive multiplayer game, more specifically, players also need to conjecture opponents' actions, which likely influence choices. An experienced player can successfully predict and avoid the harm from opponents' actions. With the consideration of opponents, even obvious choices may change. Especially in a strategy game, players usually need make a long-term plan, consisted of a series of actions. Human opponents with similar abilities may easily interfere a player's plan, resulting in unexpected results. When this happens, the player may need to adjust his/her plan. Therefore, by providing various difficult choices with dilemma, a game may become non-linear and start to facilitate players to think.

2.2 Choices facilitate pleasure

Previous research has shown that the provision of choices will increase the feelings of intrinsic motivation because choices allow people a feeling of autonomy (Ryan, & Deci, 2000). Choices, as a concept against constraint, may free people's mind and satisfy what they need under rules. For this reason, Malone and Lepper (1987) have advocated providing learners with explicit choices may facilitate their perception of control, an individual element of intrinsic motivation. In particular, learners can construct, select and name the features of fantasy to have a personalized learning and playing experience. As a matter of fact, Cordova and Lepper (1996) have showed that, in an education game, learners with choices and personalized fantasy may report obvious pleasure, including higher favor of the game and more willingness to stay after class.

The previous paragraph concluded that choices influenced pleasure. Interestingly, research also found that pleasure may influence choices. More specifically, Mellers, Schwartz, and McGraw (1999) showed that choices were associated with anticipated pleasure. In other words, people tend to select an option with greater pleasure they perceived. Furthermore, Mellers and McGraw (2001) also identified an outcome effect that anticipated pleasure rose with the increase of anticipated outcome. These results suggested that people may make decisions based on the higher anticipated outcome as well as pleasure.

Although it seemed that the aforementioned research upheld as more choices as possible, some studies also revealed the negative consequences of too many choices. For example, Iyengar and Lepper (2000) showed that people with extensive choices (up to 30 choices in this case) reported less satisfaction than those with limited choices (6 choices in this case). Furthermore, when those people made a decision with too many choices, they felt both pleasure and frustration at the same time. This phenomenon was termed as "choice overload" or "paradox of choice" later by Schwartz (2005).

The possible reasons were that too many choices brought large-scale comparisons as well as the regret of the second best options (see Scheibehenne, Greifeneder, & Todd, 2010). Comparison the anticipated outcome of the chosen options with possibly positive outcomes of an un-chosen option may result in regret, yet if the outcome of an un-chosen option was negative, such a comparison would be pleasure (Mellers, & McGraw, 2001). Therefore, it should be more careful to design choices in a game-based learning environment, as Malone and Lepper (1987) suggested.

3. Design Tactics to Facilitate Thinking and Pleasure

From the perspective of board games, this section enumerates three design tactics to facilitate both thinking and pleasure of learners in a game-based learning environment. The three tactics are introduced in order of complexity.

3.1 Tactic 1: Choices

As a basic design tactic, choices, referring to selecting one option from several ones by definition, are widely adopted in board games. Board games usually allow players to choose one action in their own turn. A recent board game *Love Letter*[5] requires players to play a card from two cards in their hands. Although the action is very simple, players may focus on deducing which is the best option with a little chance. Because the game brings so much fun, the game received best card game, best family board game, best innovative board game, and best party game awards in 2013.

For enhancing strategic thinking, board games usually limit the amount of resources. Another board game *Puerto Rico*, for example, requires players to choose one from six actions in order to obtain limited buildings, plantations, workers, and spaces for selling. The player may execute the chosen action first with a privilege (*e.g.* to pay 1 dollar less or to get 1 resource more, depending on the chosen action); the rest of players execute the same actions in order without the privilege. Without sufficient resources, even simple selection may become seriously careful and tactical.

However, without choices, insufficient resources do not necessarily lead to strategic thinking. For example, *Monopoly* is a well-known family board game without sufficient resources and choices. In the game, a player plays as a landlord, aiming at buying properties to make a fortune. In one's turn, one has to roll a dice, which may randomly determine one's moves from 1 to 6. If one stays at an unoccupied property, one may buy it as one's own asset; however, if one stays at an opponent's property, one has to pay the rent. Although the game provides limited properties, players' actions are essentially determined by randomness (no choices), making the game linear and non-strategic.

In order to facilitate players to think and have real pleasure, the game should empower players to decide their actions. A good example is a digital educational game, Joyce, designed to incorporate educational quizzes in *Monopoly* (Chang, Yang, Yu, & Chan, 2003). Although the general game rule of Joyce was based on *Monopoly*, the designers modified the rule of moves for enabling players to choose their moves. More specifically, when players collect two random numbers by rolling dices, they were allowed to move with three choices: (a) the sum of the two numbers, (b) the larger one minus the smaller one, or (c) the smaller one minus the larger one (*i.e.* moving backward). Instead of moving randomly or freely, such a choice may force players to figure out optimal solutions, to take responsibility and to enjoy any possibilities of a non-linear game.

3.2 Tactic 2: Combination

Combination refers to mixing several parts for creating an assigned pattern according to the game rule. Some board games adopt the design of combination even if one action is chosen at a time. *Mahjong* is a typical example to demonstrate the combination tactic in a single action. In the game, players are required to draw a tile and then to choose a tile to discard from seventeen titles in hands in their own turn until anyone makes a legal combination of hands and wins the game. Although the action is simple, the action is actually a difficult decision to discard the most useless tile for increasing his/her probability of making a legal combination as well as decreasing the others' probabilities.

Thinking about combination is never easy for players, because it involves enumerating all possibilities. In other words, combination demands the ability to recognize, specify, and generalize patterns. In *Mahjong*, players first have to classify the tile types and sort the tiles of the same type physically or mentally. Then, as the rule of legal combination requires, they need to group every three tiles of the same or successive figures into a pattern. Experienced players can create several grouping ways on purpose in order to increase their own wining probabilities. Sometimes players need to give up one of the grouping ways because these ways are mutually exclusive.

[5] *Love Letter* is a card game, designed by Seiji Kanai and published in 2012.

On the other hand, combination can also be considered as a complex set of choices, requiring players to make several decisions at the same time. A popular poker game, *Big Two*, uses this tactic. Players in their turn have to decide to play single card, pairs, triples, or five-card hands. In a sense, the action can be decomposed as multiple decisions on whether they play each card or not. Alternatively, it can also be considered as decisions on how many and which types of cards should be played. In either way, players have to evaluate all conditions of the combination and their consequence.

The tactic of combination can also be adopted in educational games, such as *Numbers League*[6], a mathematical card game. In this game, players take the role of superhero leagues, whose goals are to capture aliens. More specifically, each alien has a certain number (*e.g.* from 3 to 26 in the easiest mode) while the value of each superhero is the sum of the numbers on their heads, bodies, and legs. If a superhero is equipped with a weapon (*e.g.* +5, -10, or ×4), the player may decide to use it or not. In order to capture an alien, the player has to specify he/she use which one or more superheroes and whether the heroes use their weapons. For doing so, the players have to consider as many combinations of arithmetic expressions as they can. For example, if a player has a superhero of 4 with a weapon +10 and another superhero of 8 with a weapon -5, he/she could capture aliens of number 3, 4, 7, 8, 12, 14, 17 and 22 in all combination of his/her superheroes with/without their weapons.

3.3 Tactic 3: Creation

In a spectrum of choices, from controlled to free actions, creation is perhaps the most implicit but imaginative way to making decisions. It is true that creation involves a plenty of choices. In drawing, for example, people actually have to decide the theme, perspective, tools, colors and so forth. Besides drawing, creation can be a writing, a song, a dance, or an artifact. Although real world enables people to produce creations at will, a game never allows players to create works totally freely. Instead, it provides rules to constrain players' creation to a certain extent.

There are many board games about creation. *DIXIT*[7] is one of them and popular. This game is consisted of a deck of cards with undefined pictures, so that every player may interpret these cards in their own way. In one player's turn, he/she plays the role of a question poser, who need to secretly select a card in hand and to give it a subjective meaning. The meaning can be a word, a phrase, a sentence, a story, a song, or even on a motion. According to the meaning, the other players also have to secretly select a card in their own hands as a distracter. After revealing all selected cards, the players except the question poser score if they can correctly choose which card belongs to the question poser. More interestingly, the question poser cannot score if all of the other players choose the correct or wrong cards. This scoring rule facilitates the question poser to pose a moderately ambiguous question, not too straightforward or too difficult to guess.

Without doubts, DIXIT may facilitate players' creativity and imagination. Besides understanding the relation between the meaning and the correct card, the pleasure of playing *DIXIT* is devising an ambiguous meaning. The aforementioned scoring rule for question posers may transform the creativity of the question poser into a simple choice. As a matter of fact, such a rule is also used in several board games of creativity, such as *Barbarossa*[8] and *Ask Anything*[9]. While the former requires players to create clay artifacts, the latter requires them to pose yes/no questions.

Not very many digital games take advantage of human's creativity as a part of games. A mobile app game, *Draw Something*, which was downloaded extensively two years ago, allowed every two players to guess each other's drawing. With the connectivity of network, this game may link friends or people across the world regardless of the constraint of time and space. However, although the painter may draw a picture creatively, the guesser actually solves a puzzle according to the hint of the picture the painter made. For the guesser, the conflict of the game is passive, making the game too linear. A possible solution is to provide some choices for guessers.

---

6 *Numbers League*, designed by Ben Crenshaw and Chris Pallace, was published by Bent Castle Workshops in 2007.

7 *DIXIT* was designed by Jean-Louis Roubira and published in 2008.

8 *Barbarossa* was designed by Klaus Teuber and published in 1988.

9 *Ask Anything* was designed by C. C. Hung and published by TwoPlus Games in 2013.

4. Concluding Remarks

In order to design digital educational games, this paper attempts to study the design tactic of board games owing to their potential for facilitating both thinking and pleasure. Historically, digital games originate from board games, and thus they may share the same core about thoughtful pleasure. Unlike digital games, which have additional fantasy, board games have pure design of game rules, worthy of thorough investigation.

This paper preliminarily identifies players' choices in games as the route to both thinking and pleasure. Based on this argument, this paper also proposes three tactics to design digital educational games. Furthermore, when the aforementioned design tactics are applied to a digital educational game, computing power should not restrict the game play. Rather, in terms of Crawford's definition of games (1984), computers ought to reflect the representation of reality or fantasy, to reinforce the interactions between human and computers or among players regardless of time and space, to enrich the conflict in any forms of exciting challenges, as well as to ensure the safety cognitively, emotionally and socially.

References

American Association of School Librarians (2009). *Standards for the 21st-Century Learner in Action*. Chicago, IL: American Library Association

Chang, L. J., Yang, J. C., Yu, F. Y. & Chan, T. W. (2003). Development and Evaluation of Multiple Competitive Activities in a Synchronous Quiz Game System. *International Journal of Innovations in Education and Training, 40*(1), 16-26.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88*(4), 715-730.

Crawford, C. (1984). *The Art of Computer Game Design*. New York: McGraw-Hill.

Epstein, R. (1977). *The Theory of Gambling and Statistical Logic*. San Diego: Academic Press.

Gamlath, L. S. (2007). Outcomes and observations of an extended accounting board game. *Developments in Business Simulations & Experiential Exercises, 34*, 132-137.

Hinebaugh, J. P. (2009). *A Board Game Education*. Lanham, MD: Rowman & Littlefield Education.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 79*(6), 995-1006.

Juul, J. (2005) *Half-Real: Video Games between Real Rules and Fictional Worlds*. Cambridge, MA: MIT Press.

Lennon, J. L., & Coombs, D. W. (2007). The utility of a board game for dengue haemorrhagic fever health education. *Health Education, 107*(3), 290-306.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow, & M. J. Farr (Eds.), *Aptitude, Learning and Instruction III: Conative and affective process analyses* (pp. 223-253). Hilsdale, NJ: Erlbaum.

Mayer, B., & Harris, C. (2010). *Libraries Got Game: Aligned Learning Through Modern Board Games*. Chicago, IL: American Library Association

Mellers, B. A., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General*, 128, 332-345.

Mellers, B. A., & McGraw, A. P. (2001). Anticipated emotions as guides to choice. *Current Directions in Psychological Science, 10*(6), 210-214.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375-394.

Rollings, A., & Adams, E. (2003). *Andrew Rollings and Ernest Adams on Game Design*. Indianapolis, IN: New Riders Games.

Rollings, A., & Morris, D. (2000). *Game Architecture and Design*. Scottsdale, AZ: Coriolis.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions, *Contemporary Educational Psychology, 25*(1), 54-67.

Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, Massachusetts: MIT Press.

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research, 37,* 409-425.

Schwartz, B. (2005). *The paradox of choice: Why more is less*. New York: Harper Perennial.

Smith, R. D., & Munro, E. (2009). Educational card games. *Physics Education, 44*(5), 479-483.

Woods, S. (2012). *Eurogames: The Design, Culture and Play of Modern European Board Games*. Jefferson, NC: McFarland & Company, Inc.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

VocaMono: An Online Multiplayer English Vocabulary Learning Board Game

Jia-Jiunn LO* and Chin-Kun HSIN
Department of Information Management, Chung-Hua University, Taiwan, R.O.C.
*jlo@chu.edu.tw

Abstract: It is imperative to construct a learning environment to make EFL (English as a Foreign Language) learners have an engaged vocabulary learning experience to enhance their learning interests. Recently, digital game-based learning suggests a new learning paradigm of learning by playing within visual and interesting environments. As a consequence, this study developed an online multiplayer English vocabulary learning board game, named *VocaMono* (short for *Vocabulary Monopoly*), for English vocabulary learning. Adapted from the famous games *Monopoly*, *Unscramble*, and *Scrabble*, this study developed *VacaMono* by effectively integrating the gameplay and pedagogy designs for vocabulary learning. *VocaMono* has been designed to have unique educational game characteristics that should be of interest to vocabulary learning. It enhances learners' motivation for learning and memorizing English vocabulary. By playing this educational game, players' are able to enhance their vocabulary acquisition abilities not only for correctly spelling taught vocabulary but also for discovering spelling patterns. This game can fit players' vocabulary levels by selecting words from the word scope. Furthermore, this game is possible to be used to correspond to the class progresses.

Keywords: English vocabulary learning, game-based learning, online game, multiplayer game

1. Introduction

In the globalized society, English communication competence is indispensable so that learning English has become a part of lives in many non-English speaking countries. Learning another language is not easy at all (Turgut and Irgin, 2009). It is important to find the threshold for language learning. Vocabulary is the foundation of a language and the core of learning and communication. They are the cornerstone of a language, from which sentences, paragraphs, and articles are formed. Laufer (1986) noted that there would be no text comprehension without understanding the text's vocabulary. It has been suggested that vocabulary competence has more contributions to language comprehension than the components of reading (Chen, 2011). Increasing vocabulary abilities can effectively enhance reading comprehension and vocabulary learning has been recognized as a central part in language learning (Chen, 2011). Language development, therefore, is highly dependent on learners' vocabulary acquisition abilities (Carol, 2001).

The most serious obstacle is not lack of various reading strategies but insufficient English vocabulary (Chen, 2011). Without knowing enough words, people cannot acquire fluent language competences neither for listening, speaking, reading, nor writing. Therefore, EFL (English as a Foreign Language) learners, particularly beginners, should be given explicit instruction to memorize and practice the basic level words through the repeated exposure so that they can automatically recognize words and enjoy reading (Chen, 2011).

To enhance vocabulary acquisition abilities, practice is indispensable. Shemesh and Waller (2000) even said that for vocabulary learning, "No practice, no learning!" Chen (2011) mentioned that "for beginners, it may be more appropriate and effective to learn English (vocabulary) in a direct and explicit way." However, vocabulary acquisition is usually the most difficult part during the learning processes for most EFL learners.

Most English vocabulary practicing activities are based on monotonous and tedious approaches so that many learners lack motivation to learn spelling words repetitively and do not learn English vocabulary effectively. For most learners in Taiwan, vocabulary learning is a terrible experience even though they spend much time. They usually think it is boring and even hate to memorize vocabulary.

The learning outcome is often limited and even frustrating (Chen, 2011). Therefore, how to improve learners' English vocabulary abilities has become an unavoidable issue and it is imperative to construct a learning environment to make EFL learners have an engaged vocabulary learning experience to enhance their learning interests.

Contemporary students are digital natives who grew up with the computer. Applications of information technologies on educational contexts are increasingly getting popular for them. Among the computer applications, educational games have got more interest by integrating learning materials into computer games to promote student-centered learning activities to help students drill facts, connect ideas, or synthesize discrete knowledge (Nettleton, 2008). Many studies have found that games can stimulate learners' motivations and enhance their learning experiences across multiple disciplines and ages (Chen, 2011). Digital educational games suggest a new learning paradigm of learning by playing within visual and interesting environments to capture learners' interests, encourage active learning, develop learning by doing, and trigger motivation and enjoyment (Regueras et al., 2009). While playing games, learners as players take control of the learning processes so that they are more willing to learn actively (Chen, 2011). One characteristic of computers is their "patience" in repetition and recycling tasks which conform to the repeated exposure and practice requirements of effective learning so that educational games can be particularly effective in learning boring materials such as vocabulary learning (Chen, 2001; Prensky, 2001).

Competition is one of the game characteristics to stimulate players' motivation. Regueras et al. (2009) stated that learners can achieve better comprehension, retain the information longer, and enjoy learning more with active learning methodologies which can be structured to force learners to compete each other in multiplayer games. The nature of multiplayer games can take the advantages of competitive learning to stimulate learners' motivation which in turn promote learning and sometimes shorten the whole learning processes (Burguillo, 2010; Regueras et al., 2009).

As a consequence, this study aims to develop an online multiplayer educational game, named *VocaMono* (short for Vocabulary Monopoly), for English vocabulary learning. It is expected to enhance learners' motivation for learning and memorizing English vocabulary.

2. *VocaMono* Game Design

Garcia et al. (2008) suggested that the easiest way to ensure entertaining value of educational games is to adapt existing popular games to be integrated with the learning goals. *VocaMono* integrates the gameplay and pedagogy designs by being adapted from three popular games, *Monopoly*, *Unscramble*, and *Scrabble*.

Monopoly is selected as the base because it is a well-known, classic, and popular amusing board game with broad target audiences. It is one of the board games into which different content domains can be integrated. Many computer games have been adapted from it and people are familiar with this game. *Unscramble* and *Scrabble* are both spelling games of word knowledge to stimulate players to rapidly retrieve appropriate words from memory through the visualization of spatially aligned alphabets (Halpern and Wai, 2007). They can enhance players' vocabulary abilities not only by recalling words on their current spelling list but also by constructively finding new words. *Scrabble* defines alphabets' points based on their possibilities to form words so that players are encouraged to spell more difficult words. The points of a spelled word are based on the sum of each alphabet's point and its location on the *Scrabble* game board.

VocaMono is developed under the client/server architecture. Being adapted from the famous game *Monopoly*, though being added with vocabulary learning activities, *VocaMono* is designed as a competitive multiplayer game which has similar gameplay rules with which players are familiar. Each player has two attributes: *money* and *credit*. The ultimate game goal of a player is to become the richest player. Figure 1 illustrates the *VocaMono* game board interface.

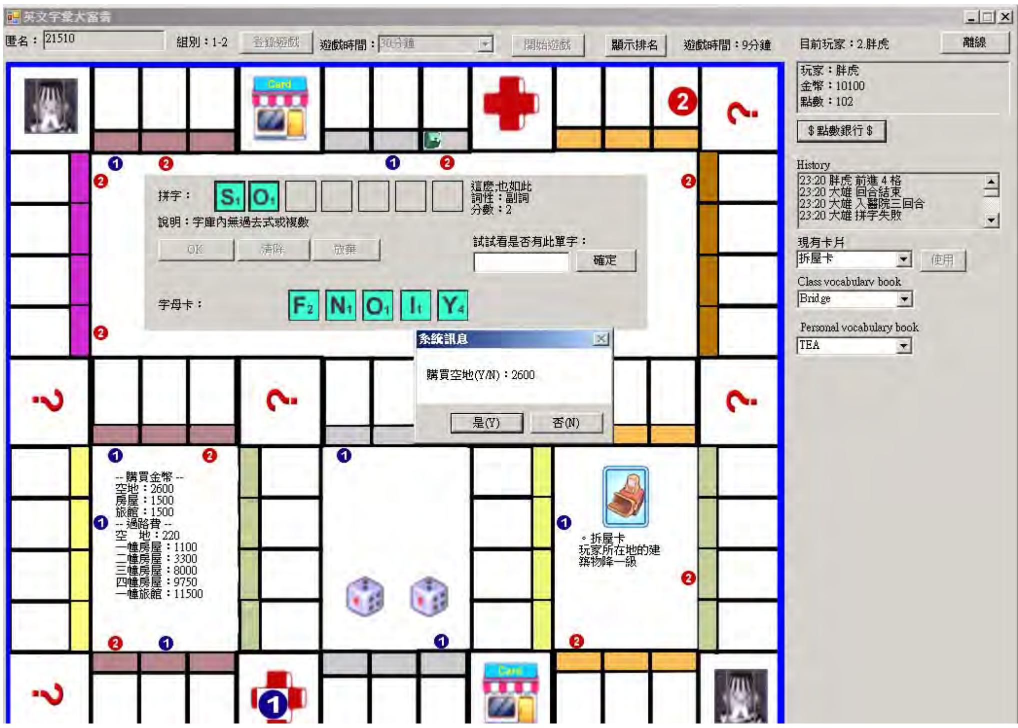


Figure 1. Illustration of *VocaMono* game board interface.

This game is turn-based. Players take turns in order. Each turn, the player complete his/her actions such as using cards or spelling a word within 30 seconds and these actions are sequentially listed in *History* box. The player first click two rolling dices to show the number of steps which his/her token moves forward along the path on the game board. The vocabulary learning activities are integrated into the gameplay rules by requiring the player to find a correctly spelled word by dragging and dropping a series of alphabet tiles which are selected according to the defined pedagogical scopes. Spelling words correctly can increase players' credit points. Players can find words either by recalling from memory or by trial and error with any combination of alphabets. A system embedded dictionary can facilitate players the trial and error process. If a word is correctly spelled, the word's Chinese translation and type (noun, verb,…) will be shown and the player's credit points increase based on the sum of each alphabet's point. Points and selection probabilities of alphabets are adopted from the famous spelling game *Scrabble* (Table 1). In this game, the default number of alphabet tiles is seven. In each turn, the system firstly randomly selects a word from the vocabulary database, whose length does not exceed seven alphabets. If the length of a selected word is less than seven, the remaining tiles are randomly chosen. These selected alphabet tiles are then presented with a random order. For example, if the word **PLAY** is selected and the remaining three alphabets **I, G,** and **E** are randomly selected with the presenting order $\mathbf{A_1}$, $\mathbf{P_3}$, $\mathbf{Y_4}$, $\mathbf{G_2}$, $\mathbf{E_1}$, $\mathbf{I_1}$, $\mathbf{L_1}$ (the subscript is the credit point of each alphabet). Players can find words with any combination of alphabet orders. With this design, players can find the target word, **PLAY** (3+1+1+4=9 points), or other words, e.g. **LIP** (1+1+3=5 points).

Table 1: The alphabet points and selection probabilities.

Point	Alphabet(Probability)
1	A(.09), E(.12), I(.09), L(.04), N(.06), O(.08), R(.06), S(.04), T(.06), U(.06)
2	D(.04), G(.03)
3	B(.02), C(.02), M(.02), P(.02)
4	F(.02), H(.02), V(.02), W(.02), Y(.02)
5	K(.01)
8	J(.01), X(.01)
10	Q(.01), Z(.01)

When a player's token moves forward along the path and stops at or passes a space, several options may be available to him/her depending on the definitions of that space. In *VocaMono*, a space can be a property, *Shop*, *Chance*, *Hospital*, *Jail*, or *Intersection*. Properties are the only spaces that can be bought and developed by players to collect rents as opponents stop at. If a player's token stops at a property that no one owned, this player can buy the property with the listed purchase price. If a player's token stops at his/her own properties, this player can construct a house, or develop the level for a constructed house to this property, and its rent will be increased accordingly. A house can have at most five levels. If a player's token stops at a property already owned by another player, this player pays the owner a given rent, depending on its level of development and whether the property is part of a set. A property set includes properties with the same color. If several properties within the same set are owned by the same player, the total rent to pay is the sum of rents of these properties. During the gameplay, the player can exchange credit and money at *Bank* anytime and trade a variety of cards with different credit points as he/she stops at *Shop*. Two types of cards are designed to increase the playfulness: *Self-reinforcement* and *Trap* cards. *Self-reinforcement* cards are those cards to increase self-competitiveness (e.g., "controlling dice numbers"). On the other hand, *Trap* cards can be used to attack others (e.g., "destroying one opponent's house"). The player can check the list of currently owned cards and use cards before spelling words. If a player's token stops at *Chance*, a random event, with different probabilities, will be triggered. Events include actions needs to occur (e.g., "go to *Jail*", "backward 3 steps"), obtaining bonus cards, etc. If a player stops at *Hospital* or is injured, whether by events triggered at *Chance* or by being trapped by another player, his/her token moves to *Hospital* wherever the token is, and ceases moving for three turns. A player's token moves to *Jail* and ceases moving for three turns when his/her token stops at *Jail*, the "go to *Jail*" event is trigger at *Chance*, or he/she is trapped by another player. The difference between dwelling at *Hospital* and *Jail* is that this player can collect rents while dwelling at *Hospital*, however no rent collection is allowed during dwelling at *Jail*. An *Intersection* is any space that has multiple directions to move forward. If a player's token stops at or passes "*Intersection*", the moving direction will be randomly assigned to increase the uncertainty of gaming to increase players' motivation (Lo et al., 2008).

In addition to luck and gameplay skills for playing games, knowledge skills in pedagogical content must be considered in educational games. *VocaMono* integrates the word scope corresponding to the learning goal into gameplay rules to enhance learners' vocabulary acquisition abilities with joyful experiences. Three pedagogical scaffolding tools, *Class vocabulary book*, *Personal vocabulary book*, and *Dictionary*, are developed to facilitate learning (see Figure 1). *Class vocabulary book* includes target words corresponding to class progress and the current learning goal is to facilitate players to acquire these words. The selection probabilities of words in *Class vocabulary book* are much higher than other words to make players have more opportunities to practice. Furthermore, to encourage players to practice words in *Class vocabulary book* to achieve the learning goal, if a word in *Class vocabulary book* is correctly spelled, double credit points will be rewarded for this word which yields high point values. All players share the same *Class vocabulary book* and they can always look up *Class vocabulary book* as references to find words. *Class vocabulary book* lists the target words to be learnt and raises players' interests by doubling the credit points, hence enhance their adherence to learn target words. Unlike *Class vocabulary book* is shared by all players, *Personal vocabulary book* records the correctly spelled words of an individual player during the gameplay. The player can review his/her own *Personal vocabulary book*. At the end of the gameplay, *Personal vocabulary book* provides summary of players' learning report, which can be used to estimate their competence of vocabulary learning. *Dictionary* is designed to facilitate players to implement the trial and error process in spelling and recognizing words.

3. Pedagogical Contributions of *VocaMono*

VocaMono has been designed to have unique educational game characteristics that are of interest to vocabulary learning. First, Ang and Zaphiris (2006) mentioned that in educational games, there are two types of winning rules: macro and micro. Macro rules define the ultimate goals. Oriented toward macro rules, the player devises individual micro rules to achieve the goal. Micro rules can be important in educational games in that they can function as guidance to steer players toward the learning goals. Word spelling practices in *VocaMono* are embedded micro winning rules to achieve the ultimate goal of becoming the richest player (macro winning rule). Players are motivated to find higher-point words, which are target words to be learnt and are usually more difficult than words with lower points.

Second, players need to simultaneously integrate four cognitive abilities to successfully play the game within limited time: verbal, visuospatial, numerical, and strategic abilities. Verbal ability can be enhanced because *VocaMono* requires players to rapidly retrieve appropriate words from memory or trial and error. Players find words by dragging and dropping alphabets from a series of random ordered titles which is related to visuospatial ability. Players also use their numerical ability to play *VocaMono* because they need to rapidly calculate the points a word can earn. In such a competitive game like *VocaMono*, players need to apply strategic ability to make skillful thinking and planning to defeat their opponents.

Third, one of the fundamental language skills is to organize and remember the correct order of the alphabets in a word (Shemesh and Waller, 2000). In *VocaMono*, the vocabulary acquisition practice is integrated into the gameplay rules by requiring the player to find a correctly spelled word by dragging and dropping from a series alphabet tiles. It is consistent to the results as surveyed by Halpern and Wai (2007): most *Scrabble* players, both experts and novices, think about forming words by physically moving the tiles.

Fourth, requiring players to form words from a set of alphabets can facilitate players to tell apart words and non-words hence find spelling patterns. Shemesh and Waller (2000) proposed that teaching vocabulary with the idea of spelling patterns can be effective and students like to acquire vocabulary through spelling patterns. Through playing *VocaMono*, players enhance their vocabulary acquisition abilities not only for correctly spelling taught vocabulary but also for finding new vocabulary and spelling patterns. As addressed by Halpern and Wai (2007): "the visualization of spatially aligned letters that create partial word combinations using implicit rules of how letters combine in English to create words (e.g., *thr* is a common alignment of letters; *rht* is not)…".

Fifth, Garcia et al. (2008) suggested that a key design concern for spelling games is to avoid showing wrongly spelled words. Therefore, spelling games based on question answering, which remind testing by presenting correct and wrong spellings, could undermine both the entertaining and learning goals. *VocaMono* requires players to find a word from a set of disordered alphabet tiles. The alphabets must be arranged in correct order. Though the alphabet tiles are disordered, incorrect words are not shown to players. It fulfills the design requirement, suggested by Garcia et al. (2008), to avoid visual learning of misspellings.

Finally, it has been recognized that fitting players' knowledge levels and class progresses is a critical criterion to select educational games for maintaining players' motivation and enhancing their learning (Chen, 2011; Uzun, 2009). The design of *Class vocabulary book*, *Personal vocabulary book*, and *Dictionary* conforms to what Halpern and Wai (2007) addressed that in the context of vocabulary learning, most people learn and retrieve words by relying on their stored lexicon of word meanings and they have relatively few words in their lexicon. Furthermore, it fulfills principle of challenge for motivation (Malone and Lepper, 1987) to allow players to be challenged at their current skill levels (Alessi and Trollip, 2001). It also conforms to the flow theory (Kiili, 2005; Lin et al., 2010) in that the balance between players' skills and game challenges are successfully balanced by the predefined word scope. It includes the players' prior knowledge and class progress and gives them appropriate challenges to help them construct new knowledge. During the gameplay, to win the game, players do their best to find correctly spelled words in a very involved, focused state. They can effectively enhance their vocabulary acquisition abilities from the gaming experience.

4. Conclusions

Vocabulary is the foundation of a language and the core of learning and communication and language development is highly dependent on learners' vocabulary acquisition abilities (Carol, 2001). However, vocabulary acquisition is usually the most difficult part during the learning processes for most EFL learners. Therefore, it is imperative to construct a learning environment to make EFL learners have an engaged vocabulary learning experience to enhance their learning interests. Recently, digital game-based learning suggests a new learning paradigm of learning by playing within visual and interesting environments. As a consequence, this study developed an online multiplayer English vocabulary learning board game, named *VocaMono*, for English vocabulary learning. Adapted from the famous games *Monopoly*, *Unscramble*, and *Scrabble*, this research developed *VacaMono* by effectively integrating the gameplay and pedagogy designs for vocabulary learning. *VocaMono* has been designed to have unique educational game characteristics that should be of interest to vocabulary learning. It enhances learners' motivation for learning and memorizing English vocabulary. By playing this educational game, players' are able to enhance their vocabulary acquisition abilities not only for correctly spelling taught vocabulary but also for discovering spelling patterns. This game can fit players' vocabulary levels by selecting words from the word scope. Furthermore, this game is possible to be used to correspond to the class progresses.

Acknowledgment

This research is supported by Ministry of Science and Technology, Taiwan, R.O.C. (NSC 101-2511-S-216-001-MY2).

References

Alessi, S.M. & Trollip, S.R. (2001). *Multimedia for Learning: Methods and Development*, 3rd ed. Allyn & Bacon, Needham Heights Massachusetts, USA.

Ang, C.S. & Zaphiris, P. (2006). Developing Enjoyable Second Language Learning Software Tools: a Computer Game Paradigm. In Zaphiris, P. & Zacharia, G. (Eds.), *User-Centered Computer Aided Language Learning*, Hersley, PA, USA: Information Science Publishing, 1-21.

Burguillo, J.C. (2010). Using Game Theory and Competition-based Learning to Stimulate Student Motivation and Performance. *Computers & Education*, 55(2), 566-575.

Carol, A. C. (2001). *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing and Research*, Cambridge: Cambridge University Press.

Chen, W.-F. (2011). *A Study of the Selection Criteria on Evaluating Game-Based Vocabulary Learning: The Case for Junior High School Students*, Master Thesis, Chung-Hua University, Taiwan, R.O.C.

Garcia, R.M.C., Kloos, C.D., & Gil, M.C. (2008). Game Based Spelling Learning. *38th ASEE/IEEE Frontier in Education Conference*. Saratoga, NY, USA.

Halpern, D.F. & Wai, J. (2007). The World of Competitive Scrabble: Novice and Expert Differences in Visualspatial and Verbal Abilities. *Journal of Experimental Psychology: Applied*, *13*(2), 79-94.

Kiili, K. (2005). Digital Game-Based Learning: Towards an Experiential-gaming Model. *The Internet and Higher Education*, *8*(1), 13-24.

Laufer, B. (1986). Possible Changes in Attitude towards Vocabulary Acquisition Research. *International Review of Applied Linguistics*, 24(1), 69-75.

Lo, J.-J., Ji, N.-W., Syu, Y.-H., You, W.-J., & Chen, Y.-T. (2008). Developing a Digital Game-Based Situated Learning System for Ocean Ecology, *The 3rd International Conference on E-learning and Games (Edutainment 2008)*, Nanjing, China, 51-61.

Malone, T.W. & Lepper, M. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In Snow, R. & Farr, M. (Eds.), *Aptitude, Learning, and Instruction*, *3*, Hillsdale, NJ: Earlbaum, 223-253.

Nettleton, K. F. (2008). Fair Game: Gender Differences in Educational Games. In Miller, C.T. (Ed.). *Games: Purpose and Potential in Education*, *Chapter 4*, Springer, 55-72.

Prensky, M. (2001). *Digital Game-Based Learning*. New York: McGraw-Hill.

Regueras, L.M., Verdu, E., Munoz, M.F., Perez, M.A., de Castro, J.P., & Verdu, M.J. (2009). Effects of Competitive E-Learning Tools on Higher Education Students: A Case Study. *IEEE Transactions on Education*, 52(2), 279-285.

Shemesh, R. & Waller, S. (2000). *Teaching English Spelling: A practical guide*, Cambridge, UK: Cambridge University Press.

Turgut, Y. & Irgin, P. (2009). Young Learners' Language Learning via Computer Games. *Social and Behavioral Sciences*, *1*(1), 760-764.

Uzun, L. (2009). An Evaluative Checklist for Computer Games used for Foreign Language Vocabulary Learning and Practice: VocaWord Sample. *Novitas-ROYAL*, *3*(1), 45-59.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Application of Game-Based Learning in Early Childhood Acquisition

Sanko Lan[a*], Joni Tzuchen Tang[b*] , Yie-Su Hwang[c]

[a&b]*APP Department, Kizi Lab Inc., Taiwan*
[b]*Graduate School of Curriculum and Instruction, National Taipei University of Education, Taiwan.*
[c]*The Department of Early Childhood Care and Education, Cheng Shiu University, Taiwan.*
[a]*lan@allproducts.com
[b]*jttang0@gmail.com

Abstract: In this paper, we addressed the issue on application of game-based learning in early childhood acquisition. Accordingly, the new game design company, Kizi Lab Inc., had implemented the theoretical background of game-based learning and early child acquisition into the developed process to make the games playable, enjoyable and learnable for 2-8 years old children. Kizi Lab Inc. had found children can play kids' tablet, Kizipad, individually and automatically with highly motivation and interests of learning, and the games also showed children are able to learn from the games effectively.
Game Play Video:
https://docs.google.com/a/tradeserv.com/file/d/0B8673eHDKiOPaGRrNFVuT3BIdzQ/edit

Keywords: Game-based learning, Early Childhood Acquisition, Kizipad.

1. Introduction

Kizi Lab Inc. is a new game-design company for early childhood education area. The rationality of this company is to develop mobile games for 2-8 years old children. The newest released product in Kizi Lab Inc. is called Kizipad (http://www.kizipad.com/). Kizipad is a kind of children tablets, which allow 2-8 years old children to have their own responsible toy and motivate children to play and learn automatically.

2. Theoretical Background

2.1 Early Childhood Acquisition

First, we need to clarify the differences between "learning" and "acquisition". According to Brown (1989), learning happens when students have awareness, whereas acquisition is an unconsciousness and automatic process. Different theory showed different ideas of early childhood acquisition.
Table 1: Theoretical Framework

Theory	Belief on early childhood acquisition
Behaviorism	Learning can be acquired by stimuli, reinforcement, and operation (Skinner, 1945).
Cognitivism	Learning acquisition goes through the meaningful process. Learners would revise their intrinsic to understand the external environment (Piaget, 1926).
Humanism	Learning acquisition is entirely innate. It is the potential of human development; it is a spontaneous progress (Chomsky, 1993).
Social constructionism	Society itself would provide the experience; the society can scaffold learners' thought (Vygotsky, 1978).

Kizi Lab believes early childhood acquisition can be nature and innate, learning can occur in a provided game.

2.2 *Game Based Learning*

Game-Based learning (GBL) is a game play which has defined learning outcomes. The idea of Game-Based Learning believe that if we can motivate children and allow them to develop an awareness of consequentiality, children can learn and acquire knowledge and information automatically (Van Eck, 2006).

Why do we need Game-Based learning in early childhood acquisition? There are several reasons. First, children are interested at games. If we can provide games in learning, we can gain children's learning motivation. Second, mobile devices can carry games; therefore, we can provide ubiquity of learning for children. Third, game-based learning believed learning can be automatically, which is related to our belief in early childhood acquisition.

3. The Road Map of Kizi Lab

3.1 *Develop a Curriculum Map*

In order to provide a curriculum map, Kizi Lab Inc. had several educational consultants to develop the map together. The curriculum map followed the learning objectives, predicted outcomes, and related research.

Right now, Kizipad had followed seven main areas from the curriculum map, which is: English as Second Language, Mathematics, Science, Life and Society, Creative Tools, Stories, Songs, and Screening Test.

3.2 *Design the Practical Games*

All of the games have followed the curriculum map, and design the practical and learnable contents. The games had been justified by several tests to ensure 1) Does the game meet the curriculum map? 2) Can children play the games individually? 3) Do children like the games? 4) Can children learn from the games?

All of the justified tests had been recorded, which allowed games designers to revise and rethink how to make the game become more practical and learnable.

3.3 *Test Each Games Reliability and Validity*

By ensuring the games had reliability and validity, all the games went through educational consultants to meet the games' validity, and all the games had been tested by 2-8 years children to meet the reliability needs.

The game went to circle process, and each test had been recorded to allow the games to meet the users' reliability, and educational validity.

3.4 *Released the games*

After the games meet all the criteria, the games can be released; however, the process had not been ended.

3.5 *Adjusted the games*

All of the games had recorded from the players. By the recording statistics, we can know how the players like the games and what are the norm for the players' ability to finish the game. If we find the games are too hard for most of players, we would adjust the games to make it easier.

4. Kizipad

Kizipad is a kind of kids' tablets, which design for 2-8 years children. Kizipad had 83 games, and users can download games monthly. All of Kizipad now is in Chinese, and later may release different language version in the future.

Figure 1. Kizipad's appearances

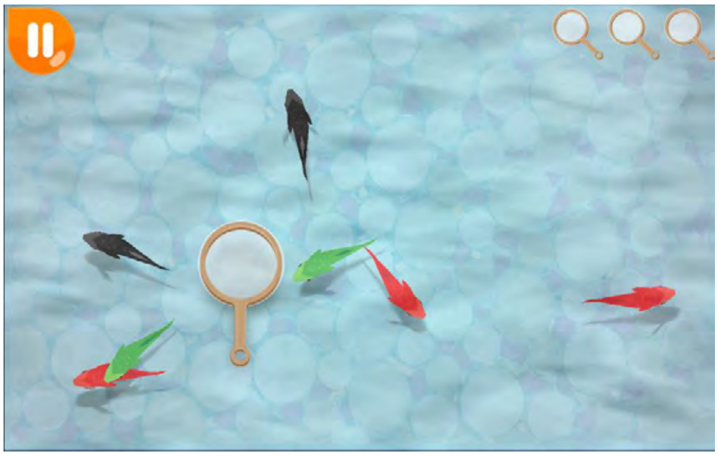


Figure 2. Coloring Games from Kizipad

Figure 3. Sorting Games from Kizipad

Figure 4. Moving Ball Games from Kizipad

Figure 5. users' recording statistics from Kizipad

For more Kizipad's content: www.kizipad.com
Game Play Video:
https://docs.google.com/a/tradeserv.com/file/d/0B8673eHDKiOPaGRrNFVuT3BIdzQ/edit

5. Contribution

Kizipad had opened a new view in early childhood area. By doing the several tests, we found 2-8 children enjoy playing Kizipad, and many of our researches and tests had found games can enhance children learning individually, automatically, and effectively (Tang, Hwang & Lan, 2013).

References

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*(January-February 1989), 32-42.

Chomsky (1993). *Language and Thought*. Wakefield, RI: Moyer Ball.

Piaget, J. (1926). The language and thought of the child. Routledge & Kegan Paul: London.

Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*. 52, 270-277, 290-294.

Tang, J. T, Hwang, Y & Lan, S (2013). The experiment of mobile learning in early childhood vocabulary. *2013 Pacific Neighborhood Consortium (PNC) Annual Conference and Joint Meetings*, Kyoto University, Japan.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *Educase Review*, 41, 2, 1-16.

Vygotsky L. (1978). Mind in society : The development of higher psychological processes. Cambridge, MA : Harvard UP.

Correlation of Professional English Reading VS. Eye Gazing and Frequency of Rereading Eye Movement

***Hong-Fa HO*[a] , Guan-Yu HOU[a] , *Chen-Ku LIN*[b], *Chen-Hsiung LIN*[c], *and Soh O-K*[d]**
[a]*Department of Electrical Engineering, National Taiwan Normal University, Taiwan (R.O.C.)*
[b]*Ph.D. Graduate Institute of Political Science, Taiwan Normal University (R.O.C.)*
[c]*Department of Industrial Education, National Taiwan Normal University (R.O.C.)*
[d]*School of Languages & Linguistics, University Kebangsaan Malaysia*
*3rrr168@gmail.com

Abstract: Over the past decades, English has become one of the major international languages widely used by many countries. Taiwan, being an international partner, has strived to improve its nationwide English proficiency to advance with international settings. This study investigates the correlation between eye movements of both 1) gazing time (fixation) and 2) frequency of re-reading (number of fixation) the vocabulary of familiar and unfamiliar professional English subjects, for those of both technical and vocational professionals of English as Foreign Language (EFL) participants, while reading professional English, to reading fluency. A newly invented eye chasing device has been used to monitor visual reading progress experiment. Fifteen technical high school students, who have completed Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET) in 2014, participated in the experiment. Their reading fluency are measured based on the results of eight questionnaires and answers (Q&A) after reading 8 professional English articles from computer engineering, mechanical, bio-medical, and business subject fields. SPSS 21 was used for descriptive and correlation statistics. Both hypotheses are accepted with results showing that readers from technical background spent more time fixing on reading familiar professional subjects, and mildly inverse relation on frequency of re-reading of professional English contexts. The combination use of multiple visual displays is recommended not only to improve English as foreign language (EFL) users' training and practice but also to enhance quality on professional English readings.

Keywords: Vocabulary of familiar and unfamiliar subject, gazing time (fixation), frequency of reread (number of fixation), eye chasing device (ECD), professional English reading (PER), Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET)

1. Introduction

1.1 Purpose and Objective

Over the past decades, English has grown to become one of the major international languages widely used by many countries (Nunan, 2003). Taiwan, Republic of China (Tw, R.O.C.) being an international partner of the global village, has constantly been striving to improve its nationwide English proficiency to advance ~~alignment~~ with international settings (Lin, 2011). While short of literature studies on expert English reading and eye configuration, this study intends to investigate the correlation between eye movements of both 1) gazing time (fixation) and 2) frequency of re-reading (number of fixation) on the vocabularies of familiar and unfamiliar subjects, when situation occurs on re-reading professional English texts, to reading fluency. A newly invented ECD has been used to monitor this visual reading observation experiment. 15 recently completed their TVE JCEET have been invited to participate in this experiment.

This research aims at finding possible association on reading fluency of professional English subject contents, to human eyes reading pattern of technical and vocational based English as Foreign Language (EFL) participants, particularly on fixation of vocabularies, and on frequency of re-reading the words. This research also attempts to serve as a reference to help bridge future researchers of possible advancement for EFL reading with help of visualization device. To contribute to the development of enhancement, this research has come up with following objectives:

1. To explore the differences between eye reading gaze time and reading fluency among technological and vocational students.
2. To detect the relationship between eyes rereading the words and reading fluency among technological and vocational students.

2. Literature Review

2.1 Background and Reading Fluency

In the past, with lack of proper tools, scholars were unable to fully identify the eye reading pattern of EFL readers when they read professional English materials, especially on those of familiar and unfamiliar fields to build the links with reading fluency. In ~~the~~ recent studies, there have been a growing number of researches conducted on ECDs' applicability, and putting them to test on general education, biological, and psychological studies (Clark & Clark, 2010; Orquin & Mueller Loose, 2013; Rayner, 2009) . With the construct of eye chasing device by the National Taiwan Normal University (NTNU), the authors are able to explore the relationships of eye movements to professional English reading.

Literacy is important for human development and interaction. Therefore, educators have attempted various approaches and methods to advance learners' literacy and knowledge, while the knowledge and literacy acquired are transferable and appropriate to expand on learning scopes and applicable for broad range of skills utility (Koo, Becker, & Kim, 2014; Kuhn, Schwanenflugel, & Meisinger, 2010; Lin, 2011; Peregoy & Boyle, 2000). Vision is an important function of learning literacy by means of visually reading for knowledge absorption (Clark & Clark, 2010; Lin, 2011; Orquin & Mueller Loose, 2013). To increase the speed and fluency of learning and the ability to demonstrate the generalized responses across fields, theory of Instructional Hierarchy (IH) by Haring and Eaton (1978) is further explained by Ardoin and Daly (2007) stated that knowledge and information are stimulus. Hence repeated practice and enhanced accuracy of response and perception to stimuli have greater control over behavior and recognition. Learners are more likely to be able to generalize from their existing knowledge and apply their wisdom and skills to new circumstances. However, numerous exemplars of targeted stimuli in multiple situations need to be implied by variety for generalization of application in order to link with practical situation (Kuhn et al., 2010; Stokes & Baer, 1977). Consequently, with more exemplars and stimuli from previous training and language learning, people of technical backgrounds may exercise a better set score of reading fluency through generalization, despite of reading unfamiliar professional English.

Reading fluency, regardless of various definitions, focuses on the consensus of accuracy, automaticity, and prosody as central components (Hudson, Pullen, Lane, & Torgesen, 2008; Kuhn et al., 2010; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011) . Reading fluency is then defined in terms of rate and accuracy, simultaneously decoding and comprehending, recognition and understanding of what has been visualized and read (Hudson et al., 2008). To syndicate the former studies, and extend on the research, the authors have added reading and fluency with visualization and PER. In the process of finding the relationships on gazing time (fixation), frequency of re-reading (number of fixation) the vocabularies of familiar and unfamiliar subjects and topics from various PER, the authors have adopted a newly invented low cost but high quality ECD to monitor the eye movements.

2.2 Eye Chasing Device References and English as Foreign Language (EFL)

Former studies have used eye chasing device for different studies (Cole, Gwizdka, Liu, Belkin, & Zhang, 2013; Mayer, 2010). The studies have shown that participants who spent more duration on ECD required the greater amount of attention on emphasis and access (de Koning, Tabbers, Rikers, & Paas, 2010; Ho, Tsai, Wang, & Tsai, 2014). English learners, of both native and foreigner, differ from one another, as the classrooms and programs that serve them. Despite of the diversity among second language readers, one common factor, as an equation of the process of reading English is essentially similar among the readers of native or non-native English speakers (Ardoin & Daly III, 2007; Fitzgerald, 1995; Goodman & Goodman, 1978). The process contains decoding of characters to linguistic they represent to thrash out at meaning. What disseminates between EFL and native English readers are the cognitive linguistic and experiential resources of reading, particularly the reading fluency (Fitzgerald, 1995; Goodman & Goodman, 1978; Peregoy & Boyle, 2000). As further noted, the direct correlation to reading fluency in English comprises "(a) English language proficiency, (b) background knowledge related to the text, and (c) literacy abilities and experiences, if any, in the first language" (Peregoy & Boyle, 2000).

2.3 Technological and Vocational Education Joint College Entrance Examination of Taiwan

The Technological and Vocational Education (TVE) system is important to nurture human resources in Taiwan. Under the commission of Taiwan Ministry of Education (MOE), English is one of the mandatory tests. Therefore, the test is a well-respected and legitimate exam. The results and associated backgrounds consequently provide the valid and consistent records for this research perseverance.

2.4 Research Questions

To accomplish the purpose of this study, the following two questions were proposed.

1. What is the relationship between eye movements on fixation, for people of English as foreign language (EFL), with technical background (TB), to professional English reading fluency (PERF)?
2. What is the relationship between eye movements on number of fixation, for people of English as foreign language (EFL), with technical background (TB), to professional English reading fluency (PERF)?

2.5 Hypothesis

Base on the questions proposed and the literatures reviewed, this research intends to test the following hypotheses (Figure 1):

1. There is an inverse correlation between participants of technical background and the time spent on reading unfamiliar professional English. In ~~an~~other words, even though they may be reading unfamiliar subject, with technical background, participants may spend normal or lesser time gazing.
2. There is an inverse correlation between the frequencies of rereading the vocabularies of unfamiliar subject for participants of technical background on contents of professional English. In other words, even though they may be reading unfamiliar subject, with technical background, participants may spend normal or lesser number of times on fixation of words.

3. Methodology

3.1 Research Framework and Process

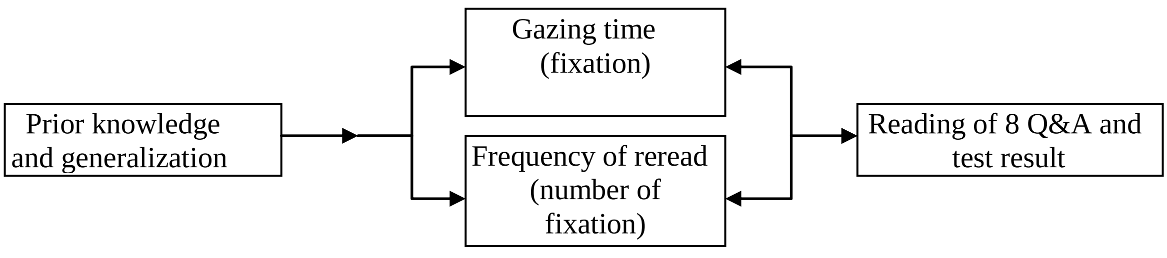


Figure 1. Observation of eye movement through eye chasing device.

3.2 Material

To conduct this experiment, 4 same sets of ECDs, head stabilizers, laptop computers, 8 (predesigned and preloaded into laptop computers) PER articles (with 2 from Mechanic, 2 on Computer Engineering, 2 from Bio-Medic, and 2 on Business), 8 (predesigned and preloaded into laptop computers) questionnaires, of one for each of reading article were used for this experiment, to promptly went through the experiment. To identify the unfamiliar words read, 15 participants were asked to circle out the words of unfamiliarity on physical printouts prepared by authors, after completion of reading the same 8 PER articles and responding to the 8 (Q&A)s electronically. All data were collected and analyzed.

3.3 Participant

Due to availability, the participants (n=15) are from Neihu Vocational High School who are majoring in Computer and Engineering majored students. They completed their TVE JCEET in 2014. This is to ensure that they have sufficient level of English knowledge and technical background for this experiment. This research has excluded the components of gender, age, social and economic status of participants. The sample population were volunteered to take part in the pilot test.

3.4 Design, process, and Procedure

Invitations for this experiment have been sent out to a number of local vocational and technical high schools. There were only 15 volunteers who would like to participate in the pilot test. To identify the focal possible PER outcomes from those with technical and vocational backgrounds, this research has invited the students who just completed their TVE JCEET for observations. On the date of experiment, one NTNU Professor and five research assistants went to Neihu Vocational High School to conduct the investigation. Participants were each equipped with one ECD, positioned individually against each one of the four head stabilizers, and sat in front of the four laptop. Participants were told to read a set of 8 pre-loaded professional English articles (2 from computer engineering base, 2 on mechanical base, 2 on bio medical field, and 2 from business related). Participants did not make known to the reading materials. After reading each article, the participants were required to answer 8 comprehension questions by selecting their answers. Participants were not assigned the time frame to complete the readings and comprehension questionnaires; however, they were accounted for by the proctors for analysis later. Four participants started the test at the same time, but performed the PER and answering the questions separately. Upon completion, participants were then asked to circle out the unfamiliar vocabulary words, of the same 8 professional English articles read electronically, on physical copies. Data were collected and analyzed afterward.

3.5 Instrument

Four recently produced EyeNTNU-180 eye chasers were used for monitoring and recording eye movements of 15 participants while reading professional English articles and responding to 8 Q&As. Each of EyeNTNU-180 set includes a laptop computer; a camera in front to record eye movements. The laptop is responsible for presenting 8 aforementioned professional English articles and questionnaire. In this experiment, a sampling rate is 180Hz (sampling 180 times per second). To avoid errors in eye chasing measurement caused by shaking of head and inconsistency of eye movement, a head stabilizer rack was used to fix head position. The distance between the screen and participants were set to 60cm straight apart. The normal time duration of a fixation of this experiment was set to 80 milliseconds.

3.6 Tools and Data Collection

EyeNTNU-180 was employed to collects eye movements, and calculates average time of visualization and coordination. Through these two parameters and ROI (Region of Interest), the Total Contact Time (TCT) and Number of Saccade (NOS) are able to be generated to provide verification of fixation and number of fixation of the words read. Four laptops were used separately on each one of the 4 separate experimental stations. The screen height was adjusted to individual's visual level, where participants were able to face straight on the screen. To avoid major or sudden physical lurching, which may affect eye positioning; participants were each placed on a head stabilizer comfortably on chin cushions throughout the experiment. Fixation Calculator, software designed from open source was used to examine the data of inspected components. One major focal indicators of the Range of Interests (ROI) were categorized by ROI-splitter software and eye movement analyzer to evaluate eye movement data to generate scan paths, gazing time, frequency of rereading the words, average and total contacts.

	Average time of unfamiliar word fixation (ms / word)	Average time of unfamiliar word fixation standard deviation (ms / word)	Average time of familiar word fixation (ms / word)	Average time of familiar word fixation standard deviation (ms / word)	Average number of saccade (NOS)	Average number of normal direction scanning and reading (ANNDSR)	Percentage of saccade (NOS / (Total Average NOS + ANNDSR))
Mechanic	359.41	79.934	102.81	16.271	791	8217.5	8.78%
Computer engineering	332.61	65.568	102.24	13.588	915	8213	10.02%
Bio-Medic	236.11	50.570	98.81	11.746	645	6406.5	9.15%
Business	201.54	46.319	97.11	12.027	670.5	6092.5	9.91%

Figure 2. Descriptive Statistic for Vocabulary of Familiar and Unfamiliar Subjects: Gazing Time (Fixation), Frequency of Reread (Number of Fixation), and Standard Deviations.

SPSS 21 software was used for descriptive statistics and Pearson correlation. The average time spent on gazing at unfamiliar vocabularies for Business subject is (201.54 millisecond (ms) /word, standard deviation 46.319), Bio-Medic subject is (236.11 millisecond/per word, standard deviation 50.570), Computer engineering is (332.61 millisecond/per word, standard deviation 65.568), and Mechanic is (359.41 millisecond/word, standard deviation of 79.934); while the average time spent on fixation of familiar word for Business subject is (97.11 ms/word, standard deviation 12.027), Bio-Medic of (98.81 ms/word, standard deviation 11.746), Computer engineering being (102.24 ms/word, standard deviation 13.588), and Mechanic subject of (102.81 ms/word, standard deviation 16.271), as shown in Figure 2. The term saccade is defined in this research as re-reading of the words and visually tracing backward at vocabularies of normal reading sequence (contrary to normal sequence of logical reading). Figure 2 shows the average number of times of saccade for Business subject being (670.5), Bio-Medic being (645), Computer engineering of (915), and Mechanic subject being (791); while the average number of normal directional sequence of reading and scanning at the contexts of Business subject is (6092.5), Bio-Medic (6406.5), Computer engineering (8213), and Mechanic of (8217.5). To take NOS and divides it by the denominator of Total Average NOS plus ANNDSR, the authors are able to measure the percentage of frequency on saccade, for each one of the professional English reading subjects, as shown in Figure 2. Figure 3 is a sample diagram of the 4 PER articles' eye maps; the lines are the visual scanning paths traveled, where the different coloring indicates the time of fixation, with no coloring means skipping of the words, and darker colors indicate longer fixation and repetitive saccades accumulated.

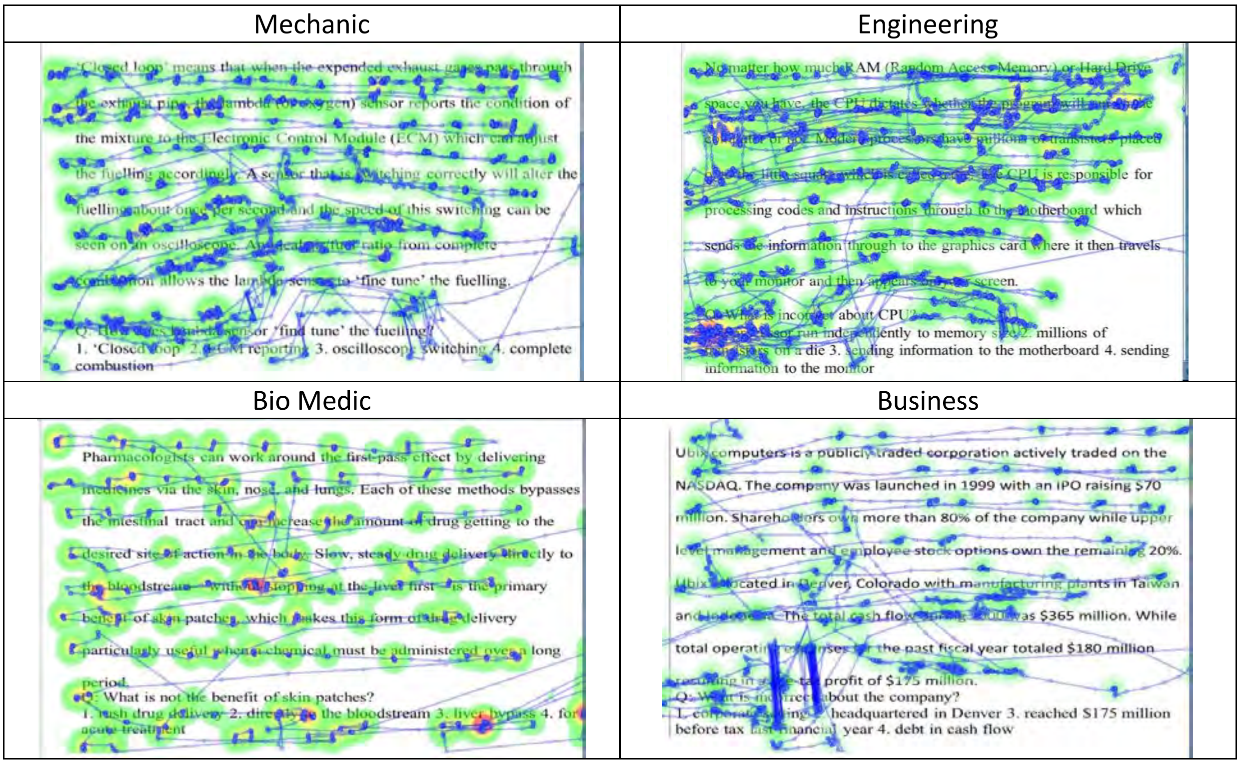


Figure 3. This is a sample eye map, the lines indicate visual scanning paths' of participant, with different colors indicate different degree of fixation and number of fixations.

4. Results and Discussions

It has been observed, from Figure 2 that participants on average spent more time gazing at unfamiliar vocabulary words of Mechanic subject, and their identical background in technical field of Computer Engineering. However, the average time spent on reading the unfamiliar words of Bio Medic and Business fields are about 100 ms per word lesser than the former ~~two~~. Mechanic and Computer Engineering, subjects wise though independent, are closer in academic study, therefore they are often grouped together under Engineering school, but Bio-Medic and Business professional subjects. The reason for longer fixation time may be explained by the similar causes provided by the former studies stated in literature review section that readers tend to devote more time and efforts browsing at recognizable contents. It is also possible because people are more willing to spend extra time reading at the contexts they are able to associate with, and are able to understand fluently than the unfamiliar subjects, i.e. in this experiment of Bio-Medic and Business professional English readings. Figure 2, indicates that the average time spent on fixing at familiar words fall around 100 milliseconds (ms) per word and the time spent is consistent throughout all the 4 professional subjects, regardless of the differences among the 4 professional English reading subjects experimented for this research.

To calculate individual professional reading field's Percentage of Saccade, the authors took the Number of Saccade (NOS) as nominator and divide it by (Total Average Number of Saccade plus Average Number of Normal Direction Scanning and Reading) as denominator. The resulting percentages fall in between 8.78% ~ 10.02% (8.78% for Mechanic, 9.15% for Bio-Medic, 9.91% for Business, and 10.02% for Computer Engineering). Figure 2 indicates that Percentage of Saccade is steady throughout all four different PERs. In Total Average Number of Saccade, this technical background participants' saccade frequency reduced noticeably when reading unfamiliar subjects. It could be explained that the behavior changes because of human rejection and un-anticipation of reading ~~at~~ unfamiliar subjects. These behavioral reflections can be identified through scan path. From the eye map scan path in Figure 3, the participant spent more time and intensity reading identical technical background of Computer Engineering and related Mechanic than Bio-Medic and Business professional English reading.

	Mechanic UNFV	Mechanic FV	Computer Engineering UNFV	Computer Engineering FV	Bio Medic UNFV	Bio Medic FV	Business UNFV	Business FV
Mechanic UNFV	1	0.498	.915**	0.473	0.635	0.298	.668**	0.415
Mechanic FV	0.498	1	0.326	.886**	0.199	0.689	0.282	.434**
Computer Engineering UNFV	.915**	0.326	1	0.301	.755**	0.106	0.777	0.288
Computer Engineering FV	0.473	.886**	0.301	1	0.355	.628**	0.353	0.404
Bio Medic UNFV	.635*	0.199	.755**	0.355	1*	0.091	.908**	0.034
Bio Medic FV	0.298	.689**	0.106	.628*	0.091	1**	-0.044	.576*
Business UNFV	.668**	0.282	.777**	0.353	.908**	-0.044	1**	0.087
Business FV	0.415	0.434	0.288	0.404	0.034	0.576	0.087	1

** Correlation is significant at the 0.01 level, * Correlation is significant at the 0.05 level.

Figure 4. Pearson correlation (2-tailed) table showing the relationships between the familiar and unfamiliar vocabularies of 4 PERs from Mechanic, Computer Engineering, Bio-Medic, and Business subject fields, with N = 15. Familiar vocabulary = FV, Unfamiliar vocabulary = UNFV.

Figure 4 shows the gazing times Pearson's correlation between the familiar and unfamiliar vocabulary of 4 professional English readings performed by Mechanic, Computer Engineering, Bio-Medic, and Business students. It is recognized that regardless of professional reading subjects, when readers come across to unfamiliar vocabularies, there is a uniformity of positive significance on the time speed on reading. For Engineering participants, the time spent on fixation and re-reading the words correlations are much stronger among the Computer Engineering and Mechanic participants, but not for Bio-medic and Business PERs. It is also observed that all significant relationships are positive in trend.

5. Conclusion

There are a growing number of researches conducted on eye chasing devices; however, after reviewing various literature~~s~~, the authors noted that professional English reading and eye movement together from English as foreign language has not been explored. Therefore, this lack of research has prompted the authors' intention to explore the correlation between eye configurations of English as foreign language users / learners with technical background, and their PERs of eye fixation and frequency of re-reading the vocabulary. With English being one of the major international languages widely used throughout most settings, the authors also intend to contribute this study to improve the EFL language usage in Taiwan.

Reading is an important part of learning, as reading focuses on the consensus of accuracy, automaticity, and prosody combined in reading fluency. Theories based on Instructional Hierarchy (IH) believe that instructed and repeated practice and reinforcement improve accuracy and may further develop perceptions through stimulations. Therefore repeated eye fixation and gazing may improve reading fluency through vocabularies and minds association. These associations can be processed through prior knowledge as stated in literature reviews and in this research experiment. With the help of ECD, the authors are able to carefully examine the correlation, with participants' eye movements, among the variables. As the result of research experiment and the literature~~s~~ studied, the two hypotheses are accepted by the authors. The eye movement and visualization serve as important functions of learning, the authors would suggest EFL to combine multiple visual equipment and tools to enhance English as foreign language users' training and practice. It may dramatically enhance the quality of professional English reading.

Acknowledgements

This research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the "International Research-Intensive Center of Excellence Program" of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301. We would also like to thank Hao-Yuan Lin who has prepared and revised previous versions of this document.

References

Ardoin, S. P., & Daly III, E. J. (2007). Introduction to the special series: Close encounters of the instructional kind—how the instructional hierarchy is shaping instructional research 30 years later. Journal of Behavioral Education, 16(1), 1-6.

Clark, R. E., & Clark, V. P. (2010). From neo-behaviorism to neuroscience: Perspectives on the origins and future contributions of cognitive load research. *Cognitive load theory*, 203-228.

Cole, M. J., Gwizdka, J., Liu, C., Belkin, N. J., & Zhang, X. (2013). Inferring user knowledge level from eye movement patterns. *Information Processing & Management*, *49*(5), 1075-1091.

de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and instruction*, *20*(2), 111-122.

Fitzgerald, J. (1995). English-as-a-second-language learners' cognitive reading processes: A review of research in the United States. *Review of Educational research*, *65*(2), 145-190.

Goodman, K. S., & Goodman, Y. M. (1978). Reading of American Children Whose Language is a Stable Rural Dialect of English or a Language Other Than English. Final Report.

Ho, H. N. J., Tsai, M. J., Wang, C. Y., & Tsai, C. C. (2014). PRIOR KNOWLEDGE AND ONLINE INQUIRY-BASED SCIENCE READING: EVIDENCE FROM EYE TRACKING. *International Journal of Science and Mathematics Education*, *12*(3), 525-554. doi: 10.1007/s10763-013-9489-6

Hudson, R. F., Pullen, P. C., Lane, H. B., & Torgesen, J. K. (2008). The complex nature of reading fluency: A multidimensional view. *Reading & Writing Quarterly*, *25*(1), 4-32.

Koo, J., Becker, B. J., & Kim, Y.-S. (2014). Examining differential item functioning trends for English language learners in a reading test: A meta-analytical approach. *Language Testing*, *31*(1), 89-109.

Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230-251.

Lin, H.-Y. (2011). The impact of intellectual capital with knowledge management on Taiwan MICE industry (Master thesis). *National Taiwan Normal University Graduate Institute of International Human Resources Development*, 1-117.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and instruction*, *20*(2), 167-171.

Nunan, D. (2003). The Impact of English as a Global Language on Educational Policies and Practices in the Asia - Pacific Region*. TESOL quarterly, 37(4), 589-613.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: a review on eye movements in decision making. *Acta psychologica*, *144*(1), 190-206.

Peregoy, S. F., & Boyle, O. F. (2000). English learners reading English: What we know, what we need to know. *Theory into practice*, *39*(4), 237-247.

Rasinski, T. V., Reutzel, D., Chard, D., & Linan-Thompson, S. (2011). Reading fluency. *Handbook of reading research*, *4*, 286-319.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology*, *62*(8), 1457-1506.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization1. *Journal of applied behavior analysis*, *10*(2), 349-367.

Testing Center for Technological & Vocational Education of Taiwan (TCTVET). (2014). Retrieved from http://www.tcte.edu.tw/en/

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education.
Japan: Asia-Pacific Society for Computers in Education

Assisting Tools for Selecting Proper Semantic Meaning by Disambiguation of the Interference of the First Language

Nattapol KRITSUTHIKUL[a,c,d,*], Shinobu HASEGAWA[b], Cholwich NATTEE[a], & Thepchai SUPNITHI[d]
[a]*Information, Computer, and Communication Technology (ICT), SIIT, TU, Thailand*
[b]*Center for Graduate Education Initiative (CGEI), JAIST, Japan*
[c]*School of Information Science (IS), JAIST, Japan*
[d]*National Electronic and Computer Technology Center (NECTEC), Thailand*
*nattapol.kritsuthikul@gmail.com

Abstract: In this work, we proposed an assisting tool to EFL Thai students to find the incorrect concept word usage in writing. To prevent selecting incorrect term effect by translation from native language, a list of commonly confusing words along with a method to score words in co-occurrence are exploited. The proposed method can indicate the miss-using word in terms of semantic and suggest the possibly correct one with detail and reason.

Keywords: English as Foreign Language (EFL), Writing Skill, Natural Language Processing (NLP), Semantic, Meaning, n-gram, search engine

1. Introduction

Most of EFL students are influenced of translating their native language to English language when they try to communicate in English. However, semantic of words in languages is apparently not equal in terms of scope, sense, usage, etc. (Speaks, 2014), and translation often applies incorrectly due to the lack of a clear understanding of the word meanings. While selecting the word in translation, students often select the word with a board meaning or the frequently seen word confused by its polysemy as they assume the word has a sense of explanation in the same form.

The issues of the translated words from native language to English are greatly noticeable in the work of writing. EFL students use incorrect word to express their content because they do not know the word that represents their concept. The incorrect issues can be categorized into four types: (1) using word with boarder meaning (hypernym), (2) using word with excessive specific meaning (hyponym), (3) using frequently seen word with similar but incorrect meaning or usage (disjoint similar concept) and (4) using direct word-to-word translation in the proverb or grammatical pattern (ignoring correlated concept). These issues are originated with the interference of their native language because of translation and the original language not containing the concepts.

From observation, case (1), using the word with boarder or too general sense, has been found the most. Unfortunately, common students clearly know the concept that they want to mention, but they do not know the equivalent translated word in English thus they select the words with general concept which are in their knowledge. Moreover, the other cases can be happened in their writing from time to time based on students' limitation of English knowledge. For example, the concept "sandal" is wanted to be expressed. For case (1), students who do not know the word may select the word "shoe" because the word is the only concept they have for footwear. Some students may use the word "slipper" as case (3) because they do not know the difference among those concepts, which are the indoor and outdoor usage purpose, since the detail may be missing from a bilingual dictionary. Furthermore, there is a case that the concept "sandal" is expressed as a compound word in their native language, such as Thai for 'รองเท้า (/rongtao/ - shoe - [noun]) - แตะ'(/tae/ - sound made by flip-flop shoe - [noun but it is polysemy to common word for verb as 'to touch']), hence they may invent new words, for instance "*touch shoe" or "*tae shoe", to represent the concept because they lack the English word for the concept or they believe there is no such concept in English.

These issues become more commonly found in EFL writing works because for them, these concepts are very confusing in terms of conceptual ambiguity from the divergence of their native language. Moreover, the translation method in expressing English for EFL students cannot be prevented since English language is greatly different from their native, and it interrupts and limits the way of their recognition. Hence, reducing the interference of the native language (L1 or the first language, henceforth) is the key to improve their English expressing. There are the words that often found as confusing words in translation provided by published dictionaries and guideline for translation by veteran translators. This information is a good hint to assist on disambiguation for those words.

In this work, we aim to develop an assisting tool to help EFL students in Thailand on selecting a proper English word. Since the case of using the hypernym is the most common and foremost issue, we aim to handle this issue first. By comparing the English written output from students with widely-published data, the likeness of pattern and wording in the written work is matched. This will result as if consulting student individual writing sentence with the large corpus to check the co-occurrence of the given words to evaluate the selection of accompanying words. By employing a list of confusable words in translation from Thai to English and WordNet, the tool is designed to reduce the interference of Thai native language on selecting the English words with commonly found cases, using hypernym and hyponym. We expect that this tool will improve their English writing skill and help them to gain more understanding of semantic concepts of English words.

2. System Architecture and Prototype System

The system is designed to find the inappropriate English words in terms of semantic meaning in the writing work of EFL student in Thailand, and to suggest a list of the better words for manual selection. The system architecture is illustrated in Figure 1.

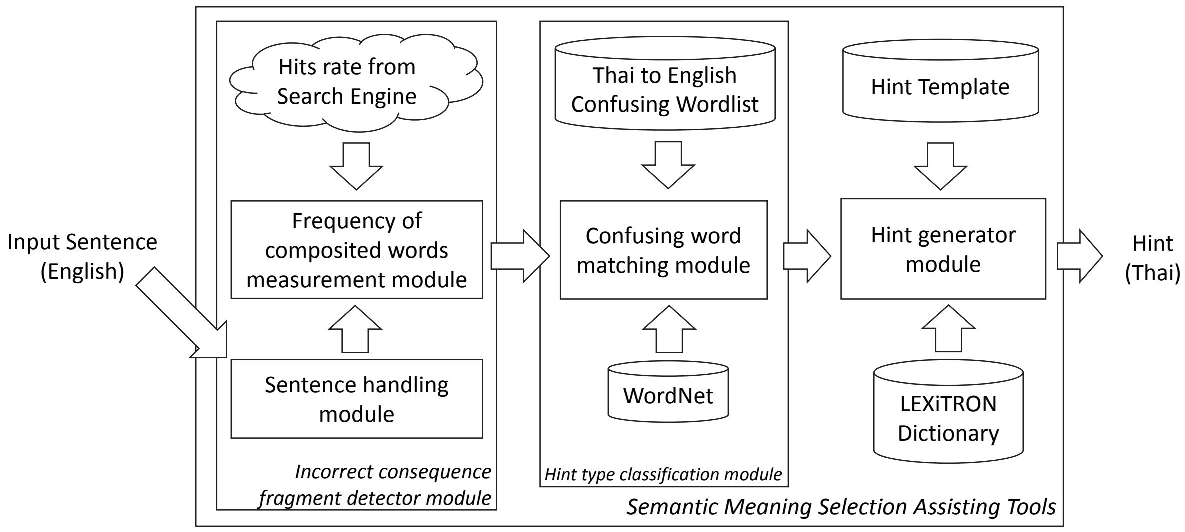


Figure 1. System architecture of the proposed system

2.1 Sentence handling module

This module gets an input as a digital English sentence made by EFL Thai students. To find an incorrectly semantically used English word among sentence, words are chunked with the *n*-gram model. The system divides the sentence into a group of consequent fragments.

For example, an input sentence from a student is "I write a picture". It will be assigned in *n*-gram model as given in Figure 2.

***n*-gram**	**consequent fragments**			
	1st-word	**2nd-word**	**3rd-word**	**4th-word**
4	I	write	a	picture
3 #1	I	write	a	
3 #2		write	a	picture
2 #1	I	write		
2 #2		write	a	
2 #3			a	picture

Figure 2. An *n*-gram model of the sentence "I write a picture"

2.2 Frequency of composited words measurement module

To acknowledge with the words with an inappropriate semantic meaning within the sentence, Hits rate (Number of search result) from search engine (e.g. Bing (Microsoft Bing, 2009), Google (Google, 1998), and so on) is employed as a concordance for measurement. Since we believe that the less frequency the words are used in co-occurrence, the more chance they are incorrectly composed with the wrong semantic meaning together. The rate is assigned to every roll. As a result from Figure 2, the hits rates of their responding consequent fragments are exemplified in Figure 3.

***n*-gram**	**consequent fragments**				**Hits rate from Bing**	**Hits rate from Google**
	1st-word	**2nd-word**	**3rd-word**	**4th-word**		
4	I	write	a	picture	**7**	**43,400**
3 #1	I	write	a		8,420,000	190,000,000
3 #2		write	a	picture	5,850,000	**295,000**
2 #1	I	write			18,400,000	10,900,000
2 #2		write	a		57,700,000	278,000,000
2 #3			a	picture	53,400,000	48,400,000

Figure 3. Hits rates of each gram by using Bing and Google

From Figure 3, according to extremely low hits rate from both Bing and Google for *n*-gram where $n = 4$, it is assumed that *n*-gram where $n = 4$ is incorrect. Moreover, *n*-gram where $n = 3$ starting from 'write' (3 #2 roll) also gets a remarkably low once comparing to other hit rates.

2.3 Confusing word matching module

Once the consequent fragments are found with low Hits rate, each fragment is examined through the Confusing Wordlist provided from published dictionaries and a guideline from veteran translators. Within the Confusing Wordlist, words ambiguous to each other are given together in a list format with their POS. An example of Confusing Wordlist is shown in Figure 4.

1. PEOPLE@N	CITIZEN@N	POPULATION@N	NATIVE@N	INHABITANT@N
2. ABOVE@PREP	OVER @PREP	HIGHER@ADJ		
3. WRITE@V	DRAW@V		PAINT@V	
4. PREVENT@V	PROTECT@V			
5. ADVICE@V	INTRODUCE@V	SUGGEST@V	GUIDE@V	
6. SANDAL@N	SLIPPERS@N	FLIPFLOP@N	FLIP-FLOP@N	THONGS@N
ESPADRILLE@N	MULE@N			
7. TALK@V	SPEAK@V	SAY@V	TELL@V	CONVERSE@V

Figure 4. An example of Confusing Wordlist

With words given in Confusing Wordlist, an example from Figure 3 is found with the word in given in line#3 from Figure 4 therefore the system attempts to replace the found word with the given alternative words in the list and re-do the Frequency of composited words measurement module With replacing word 'write' with word 'draw', we gain the result demonstrated in Figure 5.

***n*-gram**	**consequent fragments**				**Hits rate from Bing**	**Hits rate from Google**
	1st-word	**2nd-word**	**3rd-word**	**4th-word**		
4	I	draw	a	picture	**156,000**	**2,080,000**
3 #1	I	draw	a		1,110,000	34,100,000
3 #2		draw	a	picture	5,190,000	**39,900,000**
2 #1	I	draw			4,870,000	3,190,000
2 #2		draw	a		32,200,000	8,900,000
2 #3			a	picture	53,400,000	48,400,000

Figure 5. Hits rates of each gram after replacing the confusing word 'write' with word 'draw'

From comparing Figure 3 and 5, we found that the Hits rate of roll 3 #2 from Google is boosted about 135 times while the whole sentence "I draw a picture" obtains much higher hits rate ratio than the sentence "I write a picture".

2.4 Hint generation module

To give a suggestion with reason, WordNet (Princeton University, 2010) is exploited to this work to show a relation between written word and correct word. There are three cases.

- If the written word is a **hypernym** by WordNet of the word returning better Hits rate, the template which mentions the word in use is "*too general term*" will be shown.
- If the written word is a **hyponym** by WordNet of the word returning better Hits rate, the template which mentions the word in use is "*too specific term*" will be shown.
- If both words are **not related** within WordNet, only the suggested words are given as a possible better word based on Confusing Wordlist.

Moreover, the suggested word will be given which Thai definition and Thai translation provided by digital Thai-English bilingual dictionary, LEXiTRON (NECTEC LEXiTRON, 1995) to give more details for learners.

3. Experiment

The objective of this experiment was to evaluate the usability of the prototype. Research subject is fifteen volunteered Thai students studying in Grade 8 from a provincial boarding school in Chonburi, Thailand. Each student was assigned to compose ten sentences in both Thai and English. Specific words from Confusing Wordlist shown in Figure 6 must once be used in the each of written English sentences. The total of those 150 sentences were given to the system and returned results given in Figure 7.

able / above / accept / accord / accordance / according to / advice / affect / agree / bring / capable / citizen / come / converse / draw / effect / enable / equal / equivalence / espadrille / exact / except / flip-flop / flipflop / go / going to / gonna / good / guide / higher / in accordance with accurate / inhabitant / introduce / lay / lie / mule / native / over / people / population / prevent / protect / raise / rise / said / same same / sandal / say / slippers / speak / suggest / take / talk / tell / thongs / well / write

Figure 6. Specific words from Confusing Wordlist assigned to students

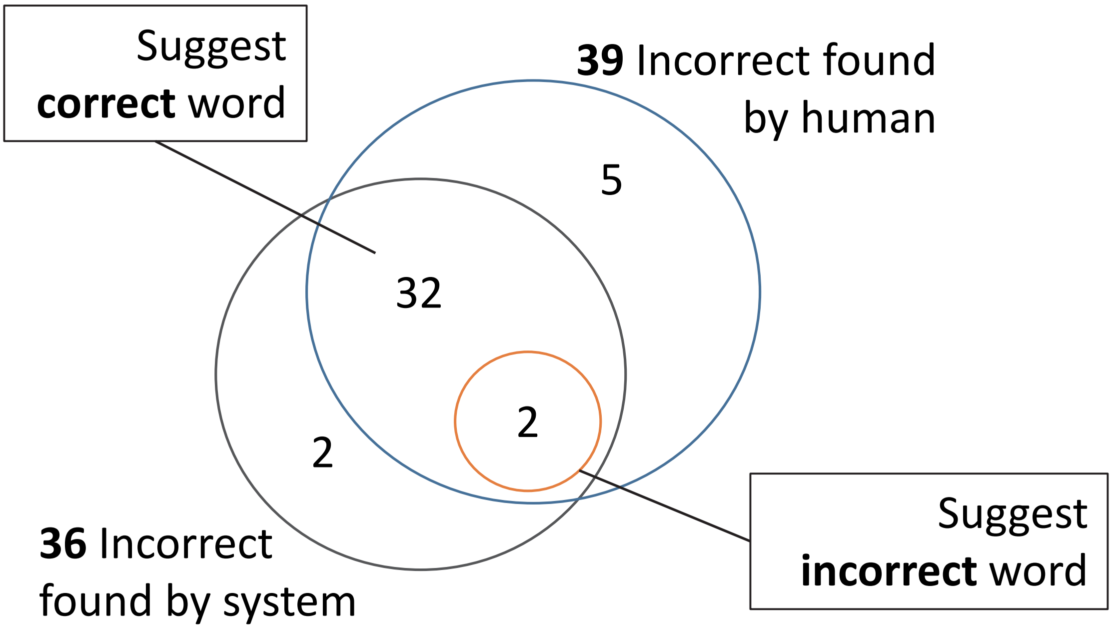


Figure 7. Results of System comparing with human

4. Discussion

From results given in Figure 7, the system returned 36 found incorrect words from 150 sentences. From those 36 found words, there were two words that are apparently correct but the system marked as incorrect. These findings were caused by the lower Hits rate of the correct one. In details, the found one is “I speak English well”. Since the n-gram chunked the sentence into gram based fragments, the immediate word ‘I speak’ was focused while the word ‘speak’ is in the Confusing Wordlist along with 'say, talk, etc.' as shown in Figure 4 line 7. Hench, the system decided to try 'I say' which gives much higher Hits ratio. This issue was generated from considering low gram number. The less n-gram word is searched for Hits, the more Hits ratio will be returned. Therefore, bi-gram word should be avoided and high number of n-gram should take higher priority.

Moreover, there are five incorrect cases that the system cannot find. This issue comes from the insufficiency of the words in the Confusing wordlist. Since the Confusing wordlist was only gathered with publicly provided data, there are other possible confusing words. It is best to gather more and more words to cover the list. Furthermore, type of confusing should be categorized to increase the scope and detail of the confusion.

Last, there are two of the suggested hints led to incorrect word selection. Since the system provides the list of possible words, students, who lack knowledge of the words though there are additional details given by bilingual dictionary, cannot select the proper word among them. It is possible that the details given by the bilingual dictionary are insufficient for low-proficiency student to understand. Thus, to facilitate more details, another detail related to the word in Thai should be added. Furthermore, an image of the word will be greatly helpful to exemplify the visual instance of the word mentioning an entity while video of motion is an explicit example of acting semantic.

5. Conclusion and Future work

This paper presents an assisting tool to help EFL Thai learners to select proper words by their semantic that they intend to. This work applies the Hits ratio from search engine to consider the words in use as statistical concordance. The low Hits result is used to find an inappropriate words using in co-occurrence. By employing confusing wordlist, the found incorrect words in consequence are given with the reason why the words are incorrect. The result of the system is to find the incorrect word among writing work with the suggest words and reason of the confusion. From testing with 150 written sentences by Grade 8 Thai students, 34 from 39 incorrect words were found while 32 from 34 found results were suggested accurately with words containing appropriate concepts to the context.

To improve the system, we plan on adding more confusing words based on Thai to English translation to increase coverage. Types of confusing will be categorized to scope and give better reason for word incorrectly used in writing. Additional information such as an image of entity and a short video of motion will be attached to suggested hint for learner to clearly understand the implicit concept of word.

Acknowledgements

This research was conducted under a grant in the SIIT-JAIST-NECTEC Dual Doctoral Degree Program.

References

Speaks, Jeff, "Theories of Meaning", The Stanford Encyclopedia of Philosophy (Fall 2014 Edition), Edward N. Zalta (ed.), forthcoming. (April 23, 2014) Retrieved http://plato.stanford.edu/archives/fall2014/entries/meaning/.

Microsoft Bing (June 1, 2009). Retrieved http://www.bing.com/

Google (September 4, 1998). Retrieved http://www.google.com/

Princeton University "About WordNet." WordNet. (2010). Retrieved http://wordnet.princeton.edu/

NECTEC LEXiTRON “LEXiTRON: an online Thai-English Electronic Dictionary.” (1995). Retrieved http:://lexitron.nectec.or.th/

Liu, C.-C. et al. (eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education.
Japan: Asia-Pacific Society for Computers in Education.

Effects of Students Using Smartphones to Receive Different Amount of L1 Support for Listening Comprehension and Vocabulary Recall

Gwo-Jen HWANG[a], Yi-Hsuan HSIEH[a], Ching-Kun HSU[b*]
[a] *Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan*
[b] *Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan*
[*] ckhsu@ntnu.edu.tw

Abstract: The study aims to explore whether different amount of the first language (Chinese) information will lead to different effectiveness on high school students' listening comprehension and vocabulary recall. Furthermore, the study also discusses the relationship between different difficult degrees of videos with two kinds given amount of subtitles/ captions and students' learning performance in English listening comprehension and vocabulary recall. One class was given the full first language subtitled movie with no foreign language captions while the other class was given the partial first language subtitled movie with full foreign language captions. The findings indicated that although the students received only the assistant of the partial first language subtitles, learners' understanding of the listening comprehension will not be affected and their performance in the vocabulary recall test statistically better than those who use the full first language.

Keywords: subtitle, caption, listening comprehension, vocabulary recall

1. Research Background and Purposes

English seems to be the most important communication tool for the global world. Many linguists would actually call it a universal language or a lingua franca. From the year 2005, Ministry of Education in Taiwan stipulates that every elementary school should start English education from the third grade. Yet English education is not confined to schools only, as living in a world surrounded by technology and multimedia nowadays, people have far more accesses to learn a foreign language (FL) than before. It is likely to use technology and multimedia to assist language learning. While there has long been interested in the relationship between multimedia and language learning, those studies have extensively investigated in the field of reading (Lysenko & Abrami, 2014) and vocabulary acquisition (Emine & Gülcan, 2012). However, the use of multimedia or mobile devices for listening comprehension is relatively unexplored; in particular how the effects of the interference of different amount of the first language (L1) information on the listening comprehension and vocabulary recall.

It goes without saying that if people want to be able to communicate with others, they should somehow comprehend what other people are talking about instantly. It can't deny that listening is an essential skill in understanding a language (Chung, 1999; Liu, Chen, & Chang, 2009). From the year 2013, every high school graduate has to participate in the TELC which is referred to as Test of English Listening Comprehension. Therefore, it is urgent to provide high school students with more opportunities of foreign langauge listening practice. Besides TELC, the TOEIC, Test of English for International Communication, also includes 100 questions for the listening comprehension part.

However, in their own country, there is rare opportunity for the students to touch foreign languages after English class. Fortunately, with the advance of technology, there are abundant methods and materials available to support students in learning a foreign language. In daily life, many convenient computer-assisted or online learning systems are also handy for training English listening proficiency (Chapelle, 2009; Liu, Liu, & Hwang, 2011).

Of all the rich learning resources, videos are the most common and popular way to promote English listening comprehension. As a result, many researches regarding EFL learning often choose videos as learning materials instead of audios or texts (Chapple & Curtis, 2000; Vanderplank, 2010). The reason for the popularity of choosing videos as learning materials is partly because there are subtitles or captions on the videos. Chun and Plass (1997) have proposed that videos with captions are useful in learning second language reading. Recently, Hsu, Hwang, Chang and Chang (2013) also indicated that veiling the easy foreign vocabulary and showing only the difficult words in the captions can promote elementary school students' listening comprehension.

Before proceeding with the following exploration on listening comprehension or vocabulary recall, the study needs to define two important concepts, "subtitles" and "captions". According to Longman dictionary of contemporary English, "subtitles" means words printed over a film in a foreign language to translate what is being said, such as an English film with Chinese subtitles. It is contrasted with the concept "captions", which deals with "words printed above or below a TV program, film, etc., to say what it is about or to give further information to help viewers who are deaf or hard of hearing to follow the dialogue, and it's usually provided in the same language." The above mentioned was as defined in related studies (Hsu, Hwang, Chang, & Chang, 2013; Danan, 2004) as well.

The study aims to explore whether different amount of L1 (Chinese) information will lead to different effectiveness on high school students' listening comprehension and vocabulary recall. This study hypothesize that less L1 (Chinese) given information will not affect the understanding toward students' listening comprehension. In the meantime, it is expected that less L1 (Chinese) interference can even be beneficial to students' vocabulary recall. L1 interference or so-called "cross-linguistic and language transfer" refers to the influence of native language structures on students' performance and development in the target language (Hashim, 1999). Furthermore, this study also discusses the relationship between different difficult degrees of videos with two kinds given amount of subtitles/ captions and students' learning performance in English listening comprehension and vocabulary recall. To evaluate the proposed hypotheses, an experiment has been conducted to investigate the following research questions:

(1) Whether less L1 (Chinsee) given information will affect the understanding of the listening comprehension?
(2) Can less L1 (Chinese) interference be beneficial to students' vocabulary recall?
(3) Will different difficult degrees of videos with two kinds of subtitles/ captions bring about diverse learning performance in English listening comprehension and vocabulary recall?

2. Method

2.1.Participants

The participants included two classes of 11^{th} graders of a private high school in Taipei city. There were a total of 86 students who volunteered to take part in the experiment. The average age of the subjects was 16. It's worth mentioning that the school is an all girls' school, so there will not be gender difference in the case. All the participants learned

English as a foreign language and were all taught by the same instructor who had taught the English course for over 7 years.

2.2. Research Design

In this study, a quasi-experiment was conducted in a high school regular English course. The participants included two classes of 11th graders and there were a total of 86 students. Each student was equipped with one smartphone and a pair of earphone. One class was assigned to be the experimental group and the other was the control group. The experimental group, including 44 students, used the smartphones and earphones to watch the filtered L1 (Chinese) subtitle video with its corresponding captions (English) individually, while the control group with 42 students used the smartphones and earphones to watch the video with full L1 (Chinese) subtitles and with no (English) captions.

The videos for the two groups are the same content with different amount of the first language subtitles and foreign language captions. Figure 1 shows the different amount of the first language subtitles and foreign language captions for the control group (right) and the experimental group (left). In addition, there are two kinds of different difficult filtering degrees of the videos being used in the experimental group. Video one (V1) displays all English words (foreign language) and partial Chinese (the first language) translations which hidden the translation of 220 words, while video 2 (V2) presents all English words (foreign language) and partial Chinese translations (the first language) which hidden the translation of the highest frequently used 1000 words. In the control group, both two videos will have full first language (Chinese) subtitles and with no (English) captions.

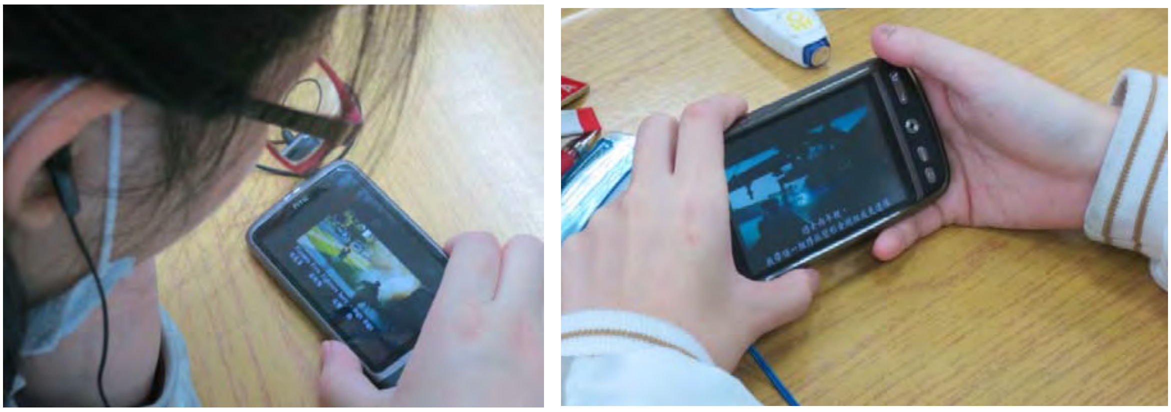


Figure 1. Students use the smartphones to watch the videos individually. The one to the left is the experimental group (partial L1 subtitles with full captions) and the one to the right is the control group (full L1 subtitles with no captions).

All the videos are 15-minutes long. The students in either group can operate the function of the player in the smartphone to "play", "pause" or "replay" the video according to their own watching requirements at will in the given time (20 minutes). After watching each video for 20 minutes, both groups of students took a listening comprehension test and a vocabulary recall test to assess their understanding in listening comprehension and the remembrance of the vocabulary. The experiment was conducted for a month, as shown in Figure 2.

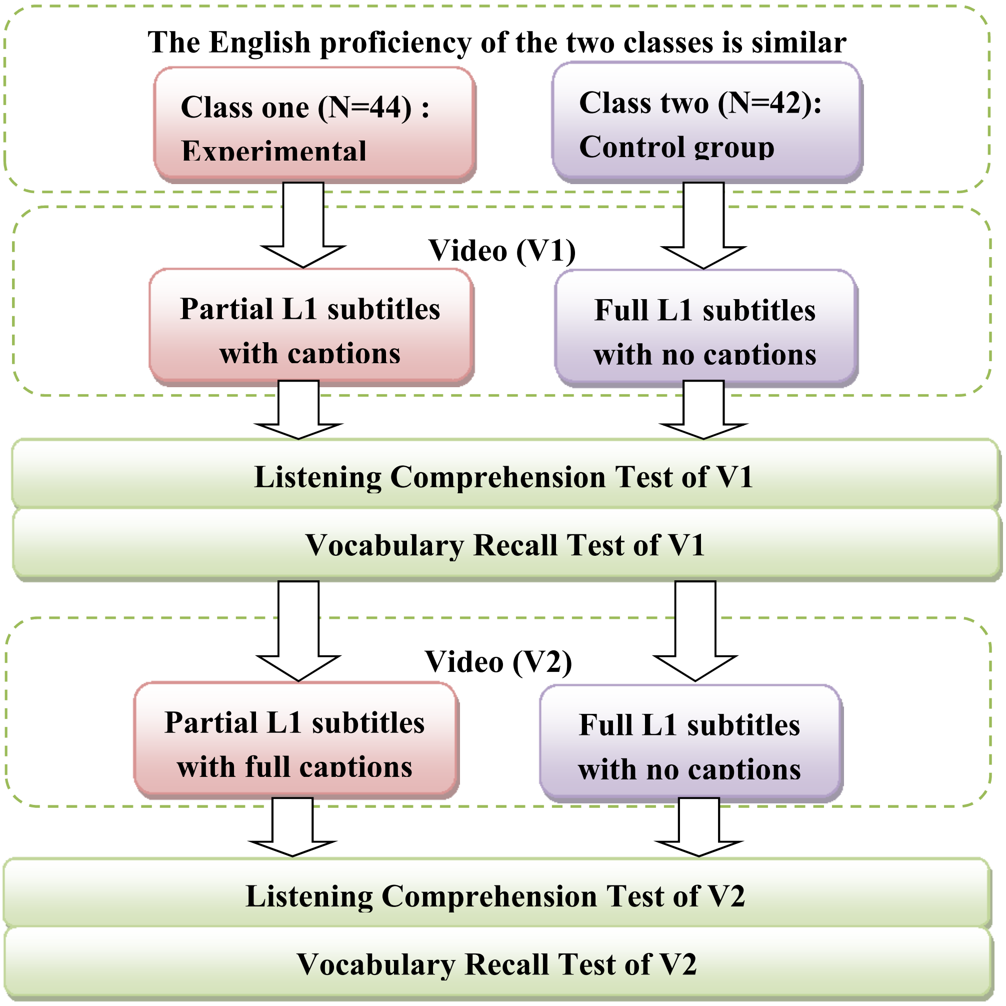


Figure 2. The procedure of the experimental treatments.

2.3. Research tools

The instruments utilized in the present study were the listening comprehension and the vocabulary recall test. The listening comprehension test and the vocabulary recall test are developed by two experts in the English teaching field. One is an experienced professor in a university and the other is an English oral training instructor who is also a native speaker. The listening comprehension test consists of 5 multiple choice questions, each account for 20 points, with a perfect score of 100. The vocabulary recall test consists of two parts. The first part is the dictation test, it consists of 5 questions, and each accounts for 8 points. The second part is the vocabulary cloze test, it consists of 6 questions, and each accounts for 10 points. The two parts add up to be a perfect score of 100.

3. Research Results

3.1. Analysis of listening comprehension test

One of the objectives of the study was to examine whether to reduce the amount of the L1 (Chinese) information will affect students' understanding of the listening comprehension.

By employing Independent t-Test on the listening comprehension test scores of the two groups, the results of this study showed that the mean values and standard deviations of the listening comprehension test scores were 39.05 and 16.20 for the control group, and 39.09 and 13.09 for the experimental group, as shown in Table1. The result showed that there was no significant difference between the two groups in the listening comprehension test (*t=0.01; p>.05*). As a result, it indicated that reducing the amount of the L1 (Chinese) information will not affect students' understanding toward the listening comprehension.

Table 1. Independent t-Test Analysis results of listening comprehension test

Group	N	Mean	SD	*t*
Experimental group	44	39.09	13.09	*0.01*
Control group	42	39.05	16.20	

**p<.05*

3.2. Analysis of vocabulary recall test

One of the objectives of the study was to examine whether to reduce the amount of the L1 (Chinese) information will affect students' recall of vocabulary. Table 2 showed the Independent t-Test result of the vocabulary recall test of the two groups. The means and the standard deviations of the control group were 31.17 and 13.87, while they were 39.20 and 14.04 for the experimental group. According to the results (t= 2.67, p < .05), it is found that there was a significant difference between the two groups; that is to say, less L1 (Chinese) interference and full foreign language caption will be beneficial to students' vocabulary recall. Students who watched the video with less amounts of L1 (Chinese) subtitles but full foreign language caption exhibited better learning performance in vocabulary recall.

Table 2. Independent t-Test Analysis results of vocabulary recall test

Group	N	Mean	SD	*t*
Experimental group	44	39.20	14.04	2.67^{*}
Control group	42	31.17	13.87	

**p<.05*

3.3. The variation between difficult degrees of videos

Another objective of this study was to examine the relationship between different difficult degrees of videos with two kinds of subtitles/ captions and students' learning performance in English listening comprehension and vocabulary recall. The study employs two different difficult degrees of videos in the experiment. The degrees of the videos were assessed by two experienced university professors in the language learning field. Video 1 (V1) is considered the harder one, and video 2 (V2) is thought to be the easier one to the subjects in this study.

From the research results, it was found that students in the experimental group which were provided with filtered L1 (Chinese) subtitles along with corresponding FL (English) captions performed better in V1's vocabulary recall test compared with the control group, as shown in Table 3. However, the results also showed that speaking of the V1's listening comprehension test, the control group statistically performed better than the experimental group. The results inferred that because the difficult piece (V1) has left more L1 (Chinese) subtitles with FL (English) captions words in the video which led students of the experimental group to divert attention from only listening to seeing the vocabulary provided. Therefore, those L1 subtitles and FL captions provided in the V1 improved the performance of the experimental group in V1's vocabulary recall test while because those experimental group students' attention was distracted by the vocabulary provided in the video, the test scores in V1's listening comprehension part was consequently lower than the control group’s test result.

Table 3. ANOVA analysis of vocabulary recall of different difficult degrees of videos.

Group Treatment	Video Category	N	Mean	SD	F	Pair Comparison
Experimental (E)	Hidden the translations of 220 sight words (V1)	44	43.14	19.28	*5.99**	EV1 > CV1*
	Hidden the translations of 1000 HFU words (V2)	44	35.27	15.20		
Control (C)	All L1 (V1)	42	27.29	16.16		
	All L1 (V2)	42	35.05	18.42		

**p<.o5*

The research results was just in line with the theory of the split-attention effect in multimedia learning (Mayer, 1998). Many previous studies have also proved this so-called redundancy effect that integration of the redundant information with essential information imposes a cognitive load interfering the learning process. Under this circumstance, the redundant sources of information is beneficial for learning. (Bobis, Sweller and Cooper, 1993; Chandler and Sweller, 1991; Kalyuga, Chandler and Sweller, 1998; Sweller and Chandler, 1994).

The reason why there were no significance appeared between V1’s and V2’s vocabulary recall in the experimental group was somehow because no matter in V1 or V2, there were FL (English) captions embedded in both videos, and thus there were no significant difference between V1 and V2’s vocabulary recall experiment results. Conversely, because the study provided less L1 information with less FL captions in the video (V2), the experimental group students could put full attention to the listening and thus the students in the experimental group performed better in V2’s listening comprehension test. Accordingly, it is inferred that fewer captions may be better for students to practice listening proficiency.

On the other hand, it was because students in the control group who were provided with full L1 (Chinese) subtitles that they hold a higher level of understanding to the plots in V1’s listening comprehension test, although V1 was viewed as a harder one in two videos. However, since the study provide only L1 (Chinese) subtitles with no FL (English) information to subjects of the control group, it was hard for them to do the vocabulary recall, and hence the scores in the V1’s vocabulary recall test was lower than those in the V2’s.

Table 4. ANOVA analysis of listening comprehension of different difficult degrees of videos.

Group Treatment	Video Category	N	Mean	SD	F	Pair Comparison
Experimental (E)	Hidden the translations of 220 sight words (V1)	44	28.18	20.83	*26.08**	EV2 > EV1* EV2 > CV2*
	Hidden the translations of 1000HFU words (V2)	44	50.00	18.04		
Control (C)	All L1 (V1)	42	54.76	19.78		CV1 > CV2*
	All L1 (V2)	42	23.33	21.15		CV1 > EV1*

**p<.o5*

4. Conclusions

This study examined two groups of students with different given amount of L1 (Chinese) subtitles and FL (English) captions. The control group was given full L1 subtitled movie with no FL captions while the experimental group was given filtered L1 subtitled movie with the corresponding FL captions. The findings indicated that although the experimental group received only partial L1 subtitles will not affect learners' understanding of the listening comprehension. Moreover, the findings further showed that with only filtered L1 subtitles, the experimental group students' performed statistically better than those in the control group with full L1 subtitles in the vocabulary recall test. While most of the multimedia assisted learning studies laid huge stress on using as much multimedia tools as one can during the teaching process, one of the major contribution of this study was that the results showed although it was partly useful with the help of the multimedia such as L1 subtitles, yet it could also be a distraction toward students' learning process. Therefore, this study recommended that to avoid distraction from the multimedia distraction, there should be only the needed amount of the multimedia provided. It was not that as much multimedia assisted as better in teaching. For example, when it comes to course design, the instructors should choose appropriate and not excessive amount of multimedia according to the curriculum objectives. Moreover, with a view to achieving the teaching goals properly, the technology-assisted multimedia or tools should take students' learning level into consideration.

This study found that if students need immediate aid to understand the content, they do not need to be provided with full L1 language. That is to say, the videos or movie series now available on TV showing L1 subtitle may provide opportunities for students to touch foreign language conditions and practice listening, but can actually offer limited training opportunities for learners to recognize vocabulary. To sum up the conclusions in this study and previous studies (Hwang, Chang, & Chang, 2013 ; Hsu , 2014 ; Hsu, Hwang, & Chang, 2014), if students want to train their listening proficiency and they have better prior proficiency of that target language, they can use no captions as well as subtitles, or only use the caption as well as subtitle of harder words. If students want to train their listening proficiency but they have poor prior proficiency of the target language, they can use caption which providing full or partial foreign language words and partial subtitles which are the translation of the harder words. If the students also want to recognize or memorize vocabulary during using the video to learn foreign language, they can use captions with full foreign language words and leaving only essential L1 information while only providing L1 subtitle without foreign language captions will not affect the vocabulary recall very well, but just support the immediate understanding to the content.

Accordingly, one of the major contributions of this study is to propose that the importance of foreign language (FL) captions is much higher than the significance of L1 language subtitles. However, previous studies have found that videos with whole foreign language but without any L1's help will bring a risk of cognitive failure understanding. Still, this study found that full L1 (Chinese) subtitles without any FL(English) captions neither do any good to language learner's vocabulary recall, but mainly instant comprehension of content. As to how much amount of the L1 information should be left to assist the learners, in fact, this study suggests future research combine with instant or online dictionary at hand or to provide scaffolding based on individual differences.

Still, it might be difficult to claim that the findings in this paper are significant since the number of subjects is not large and the experimental period is relatively short. For that reason, it was worth conducting extended studies with a large number of participants and a longer period of time in the future. Also, it is worth mentioning that the English listening and vocabulary proficiency can't be achieved or promoted overnight. Still, it is believed that this study could provide a descriptive basis for the following and additional research as long as there is a continuing need for and adequate theoretical basis for the practical application of language teaching and learning.

Acknowledgements

This study is supported in part by the Ministry of Technology and Science in Taiwan under contract numbers: NSC 102-2511-S-003-055-MY2, MOST 103-2628-S-003-003-MY2, and NSC 102-2511-S-011 -007 -MY3.

References

Bobis, J., Sweller, J. and Cooper, M. (1993). Cognitive load e€ects in a primary school geometry task. *Learning and Instruction, 3,* 1-21.

Borras, Isabel; Lafayette, Robert C.(1994). Effects of Multimedia Courseware Subtitling on the Speaking Performance of College Students of French. *Modern Language Journal*, *78*(1),61-75.

Chandler, P. and Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8,* 293-332.

Chapelle, C. A. (2009). The relationship between second language acquisition theory and computer-assisted language learning. *The Modern Language Journal*, *93*, 731-753.

Chapple, L., & Curtis, A. (2000). Content-based instruction in Hong Kong: Student responses to film. *System*, *28*(3), 419–433.

Ching-Kun Hsu (2014). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices . *ReCALL*, Available on CJO 2014 doi:10.1017/S0958344014000214

Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology, 1* (1), 1-35.

Chung, J. (1999). The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study. *Foreign Language Annals*, *32*(3), 295-308.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta, 49*(1), 67-77.

Emine Türk & Gülcan Erçetin. (2012) Effects of interactive versus simultaneous display of multimedia glosses on FL reading comprehensionand incidental vocabulary learning. *Computer Assisted Language Learning, 27*(1), 1-25.

Hashim, A. (1999).Crosslinguistic influence in the written English of Malay undergraduates. *Journal of Modern Languages,12*(1), 59-76.

Hsu, C.-K., & Chang, C.-K. (2010). Effects of Automatic Hidden Caption Classification on a Content-based Computer-Assisted Language Learning System for Foreign Language Listening. *ICCE2010*, Putrajaya, Malaysia.

Hsu, C.-K., Hwang, G.-J., Chang, Y.-T. and Chang, C.-K. (2013). Effects of Video Caption Modes on English Listening Comprehension and Vocabulary Acquisition Using Handheld Devices. *Educational Technology & Society*, *16*(1): 403–414.

Hsu, C.-K., Hwang, G.-J., & Chang, C.-K. (2014). An Automatic Caption Filtering and Partial Hiding Approach to Improving the English Listening Comprehension of EFL Students. *Educational Technology & Society, 17* (2), 270–283.

Kalyuga, S., Chandler, P. and Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40,* 1-17.

Liu, P. L., Chen, C. J., & Chang, Y. J. (2009). Effects of a computer-assisted concept mapping learning
strategy on EFL college students' English reading comprehension. *Computers & Education*, *54*, 436-445.

Liu, G. Z., Liu, Z. H., & Hwang, G. J. (2011). Developing multi-dimensional evaluation criteria for English learning websites with university students and professors. *Computers & Education*, *56*(1), 65-79.

Lysenko, Larysa V.; Abrami, Philip C. (2014). Promoting reading comprehension with the use of technology. *COMPUTERS & EDUCATION*, *75*, 162-172.

Sweller, J. and Chandler, P. (1994). Why some material is di cult to learn. *Cognition and Instruction*, *12*, 185-233.

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, *43* (1), 1–37.

Liu, C.-C. et al. (eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education.

Virtual English village: A task-based English learning platform in Second Life

Yu-Ju LAN[a]*, Hsiao-Hsuan WEI[b] & Ya-Li CHIU[b]

[a] *National Taiwan Normal University, Taiwan*
[b] *Wanfu Elementary School, Taipei, Taiwan*

*yujulan@gmail.com

Abstract: The purposes of the current study are to develop a virtual English learning village for elementary EFL learners in Taiwan to allow them to learn English beyond the regular school schedule. Four learning contexts were developed and used by 117 EFL learners from an elementary school in Taipei City for around 5 months. Learners' motivation of learning English via executing language tasks in this virtual village were evaluated through questionnaire administration. The results reveal learners' positive perception of the English learning mode as proposed in this study. Thus, a task-based English platform in virtual worlds can be a potential solution to the problems of the lack of authentic contexts nowadays in the traditional EFL settings, and consequently is able to provide elementary-school EFL learners with an immersive environment for building authentic experiences.

Keywords: Virtual worlds, task-based learning, English as a foreign language (EFL), English village

1. Introduction

In the digital era, English is regarded as an international language to which IT is widely applied in various fields around the world. In non-English-speaking countries, such as Taiwan, English proficiency is highly valued as a requisite feature of participating competitively in the international community (Lan, Sung, & Chang, December 2009). Traditionally, Taiwanese students often learn and practice English in static and old classroom activities, instead of learning from genuine or real world simulations of daily life. More and more teachers or scholars believe that a scenario environment is a pivotal component necessary for the acquisition of language (Lan, 2014). A meaningful language must be learned in conjunction with society, culture and personally relevant life experiences (Hedberg & Alexander, 1994 ; Krashen, 1981).

From this ideological frame of references, 12 elementary schools in Taipei City are selected and supported by the Taipei City Government to establish active and kinesthetic learning environment which is called *English Village*, hoping that children can begin to dream of a life with a broader and more global perspective. The 12 chosen schools designed the English Villages based on the consideration following specific pedagogical purposes of their own schools. However, similar topics such as airport, local culture, restaurant, post office, hospital, etc. could be found in different Villages owned by different elementary schools. Typically, each English Village is not only used by the students of the elementary school with that Village, but it is also visited by the students from all the elementary schools located nearby in the same educational district due to the limited amount of government budget to support every school to own its own English Village. One obvious problem has been encountered since the Village is established: the lack of opportunity for students to preview the skills needed in experiencing the Village before visiting the setting or reviewing what they have learned or experienced in the setting.

The purposes of this study were to develop and evaluate an authentic learning environment in virtual worlds (Second Life, SL) which aim at providing all the students from the participating or other elementary schools with an immersive environment to pre- and review the materials embedded in the Village they visit in the real world. The methods are briefly described below, followed by the results and the conclusion.

2. Methods

2.1 Participants

117 students from an elementary school having an English Village (Wanfu English Sky Castle, WESC) in Taipei City, with three learning classrooms (the airport, the world style setting, and the setting with the local culture theme), participated in this study. They came to computer lab, logged in the virtual Village to preview or review what they would or had learned in WESC both before and after they visited the setting.

2.2 Research Design

A qualitative research design was adopted in the study. Students' perception of English learning in SL was administered and analysed to determine what the participating elementary school learners thought about learning English in such a virtual village.

2.3 Instruments

2.3.1 Usability Questionnaire

The questionnaire used in this study was revised from what was developed by Lan (2014). 22 items of 4-point Likert Scale belonging to 4 dimensions (pragmatic, ease to use, and ease to learn, and satisfaction) are included in the questionnaire to determine the participants' perceptions of English learning in the virtual village.

2.3.2 Virtual English Village in SL

Four learning contexts aimed at providing students with learning scenarios identical to what they experienced in the real English Village (WESC) were developed in Second Life. The contexts include (1) an international airport for learning check-in and passport control, (2) a ring toss game for learning to play the game in a traditional night market in Taiwan, (3) a food court in a night market for learning the names of foods and sentences used for ordering foods in a night market, and (4) a restaurant for learning to order food in a restaurant. Figures 1 and 2 show the real and virtual contexts in WESC and SL used in this study, respectively.

(1)

(2)

(3)  (4) 

Figure 1. The Real Contexts in WESC Used in this Study: (1) Airport, (2) Ring Toss, (3) Food Court in the Night Market, and (4) Restaurant.

(5)  (6) 

(7)  (8) 

Figure 2. The Virtual Contexts in SL Used in this Study: (1) Airport, (2) Ring Toss, (3) Food Court in the Night Market, and (4) Restaurant.

2.3 Procedure

The study lasted for around 5 months. Each topic was learned for one month, plus another month for trainings on SL operation and questionnaire administration. At the beginning, all the students received the training on using SL, such as moving their avatars, clicking virtual objects, and answering questions in SL. Then all the participating students went to the computer lab once a week to learn the vocabulary words and sentences needed in experiencing the real English Village in their school. Three-stage activities were designed for guiding students to learn the target materials: firstly, learning vocabulary words by clicking the objects in SL; secondly, learning sentences by listening to the demonstration given by the non-player character (NPC); and thirdly, learning to communicate with the NPC via choosing correct responses. All the learning activities in each stage are self-directed. Besides, students will receive feedback from the system while practicing the conversation. They will be also given rewards from the system for having completed the learning targets. After all the learning stage is completed, a usability questionnaire was administered to all the participants for understanding their perception of the learning mode in SL.

3. Results

Table 1 lists the descriptive results of the usability questionnaire. Based on the data listed in Table 1, it is found that all the dimensions of the questionnaire got pretty high scores from students' responses (above 3 out of 4) and very low standard deviation (less than .01).

Table 1: The descriptive results of the usability questionnaire.

Dimensions	Scores (N=117)	
	Mean	SD
Pragmatic	3.31	.03
Ease to use	3.39	.05
Ease to learn	3.35	.07
Satisfaction	3.35	.07
Total	3.32	.09

In addition to the Mean and SD obtained from all the respondents as listed in Table 1, it is also found that for each item of the questionnaire, near 90 percent of the participants expressed that they extremely agree or agree with it that learning English in the virtual contexts in SL benefited their English learning.

4. Discussion and Conclusion

The primary mission of the English Villages is *environmental stewardship*, *curriculum integration*, and *a sampling of global cultures*. Using varied subject matters and pedagogies, the English Village aims to develop a variety of lesson formats that can be adapted to the needs of diverse learning styles of elementary school children in keeping with contemporary theories of multiple types of intelligence. By incorporating virtual contexts in virtual worlds into the physical English Village, learning is enhanced by the self-directed and task-based learning platform without special or temporal limitations.

The results of this study are in line with many of the previous TBA researches in SL contexts (e.g., Ellis, 2003; Deutschmann et al., 2009; Grant, 2008; Jee, 2011; Lan, 2014; Nunan, 2004). SL combines network connectivity and virtual reality, providing users with a virtual yet "real" environment in which they can interact socially with others via resident avatars (James, 2005). According to Burdea (1993), a virtual environment consists of three essential features: immersion, imagination, and interaction. Cooke-Plagwitz (2008) suggested that an environment such as SL can inspire language learners to reach learning goals without losing interest or motivation.

Based on the positive results obtained from the current study, it is suggested that virtual learning contexts should be included in regular EFL syllabus for providing elementary school EFL learners with an authentic and immersive English learning which has higher flexibility, lower costs and fewer limitations.

Acknowledgements

The authors would like to thank the Ministry of Science and Technology, Taiwan, R.O.C., for financially supporting this research under Grant Nos. NSC 101-2511-S-003-031-MY3 and MOST 103-2628-S-003-002-MY3. They are also thankful that this research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C., the "International Research-Intensive Center of Excellence Program" of NTNU and the Ministry of Science and Technology, Taiwan, R.O.C. under Grant No. NSC 103-2911-I-003-301.

References

Burdea, G. (1993, April). *Virtual reality systems and applications*. Electro'93 International Conference, Short Course, Edison, NJ.
Cooke-Plagwitz, J. (2008). New directions in CALL: An objective introduction to Second Life. *CALICO Journal*, *25*(3), 547-557.
Deutschmann, M., Panichi, L., & Molka-Danielsen, J. (2009). Designing oral participation in Second Life – a comparative study of two language proficiency courses. *ReCALL, 21*(2), 206-226.
Ellis, R. (2003). *Task-based language teaching and learning*. Oxford: Oxford University Press.
Grant, S. (2008). Immersive multi-user virtual environments: A new platform for foreign language teaching and learning. Conference proceedings of *the 17th Biennial Conference of the Asian Studies Association of Australia*, July 1-3, 2008, Melbourne, Australia.
Hedberg, J. & Alexander, S. (1994). Virtual reality in education: Defining researchable issues. *Educational Media International*, *31*(4), 214–220.
James, H. (2005). *Finding meaning in the second half of life: How to finally, really grow up.* NY: Gotham Books.
Jee, M. (2011). ESL Students' interaction in Second Life: Unofficial task periods activities. *Korean Journal of Applied Linguistics*, *27*(4), 55-80.
Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
Lan, Y. J. (2014). Does Second Life improve Mandarin learning by overseas Chinese students? *Language Learning & Technology*, *18*(2), 36–56.
Lan, Y. J., Sung, Y. T., & Chang, K. E. (December 2009). Let us read together: Development and evaluation of a computer assisted reciprocal early English reading system. *Computers & Education*, *53*(4), 1188-1198.
Nunan, D. (2004). *Task-based language teaching*. UK: Cambridge University Press.

Liu, C.-C. et al. (eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education.

Model for Supporting Cognitive and Metacognitive Strategies in Technology Enhanced Language Learning

Katrin SAKS*, Äli LEIJEN
University of Tartu, Estonia
*katrin.saks@ut.ee

Abstract: Good language skills bridge the gaps in global communication. Even though English is being considered the world language, advancing the English proficiency is the primary task in the countries where it is not the mother tongue. The learners' language proficiency has been found to be correlated with their use of language learning strategies (LLS) (Hsiao & Oxford, 2002; Benson, 2011). That is why measuring the learners' use of LLS is considered one of the most widely-spread methods of estimating the efficiency of language studies. The skills and habits of strategy use are very individual and usually develop over many years. However, as shown in several studies (e.g Brunstein & Glaser, 2011; Kondo *et al*, 2012; Nash-Ditzel, 2010) it is possible to support learners' language studies leading them to use more effective learning strategies. The aim of the current study was to conduct a literature review on different interventions that have been used to support using effective learning strategies in the context of technology-enhanced learning and drawing on the results of this review to design a model and concrete assignments to support learners' effective LLS use towards greater language proficiency and self-regulation. The most promising interventions supported both cognitive and metacognitive activities and often utilized the form of prompts. Based on these principles a model was designed for scaffolding language learning assignments in a blended learning environment. Support for learners is provided with prompts that aim to guide learners to more efficient and conscious use of LLSs.

Keywords: language learning strategies, language proficiency, self-regulation, intervention

1. Introduction

The widespread use of learning technologies in different educational settings has generated the growing need for students to self-regulate their learning activities (Bannert & Reimann, 2012), assuring independence and learner autonomy when directing their studies. When in traditional classroom students may stay passive knowledge receivers, learning in technology-enhanced environments presumes learner autonomy, abilities to self-analyze and self-regulate their learning activities. Many studies have identified a significant positive correlation between academic achievement and self-regulated learning ability (Dabbagh & Kitsantas, 2005; Schunk & Zimmerman, 1994) which indicates that good learners typically have good self-regulating learning abilities. However, students rarely demonstrate adequate skills of self-regulation which in turn hinders to achieve satisfactory academic outcomes (Lee *et al*, 2010; Kiewra, 2002). Similar deficit of self-regulated learning skills has been reported in the studies of language learning (Benson, 2011). The learner's use of language learning strategies is considered one way of assessing the efficiency of his language studies (Hsiao & Oxford, 2002) as strategy use is connected with language proficiency and have much potential for enhancing learning, learner autonomy, independence and self-regulation (Wong, 2011).

Self-regulated learning skills are of crucial importance to be academically successful, at the same time they seem to be complicated learning skills to acquire. Therefore, the most efficient ways to support learners' cognitive and metacognitive learning strategies have to be found and instructed. As there is no single understanding of an efficient way of scaffolding self-regulation, the focus of the current study is (1) to find out the most effective supports and conditions for scaffolding self-regulation reported in empirical studies, and (2) on the basis of them to develop a model with specific assignments to assist self-regulated language learning mechanisms in a blended English language course (English for Specific Purposes). When developing the model we drew on the theoretical frameworks of LLSs by R. Oxford (1990) and self-regulation by P. Pintrich (2000). Subsequently, the theoretical frameworks we proceeded from, will be introduced.

2. Models of learning and learning strategies

Many constructivist learning theories and models emphasize the role that self-regulation plays in the learning process. The theory of strategic learning (Weinstein, 1994) is focused on students as active, self-determined individuals who process information and construct knowledge. The model has placed the learner at its core, and surrounds him with three interactive components that explain successful learning: skill, will, and self-regulation. *Skill* refers to the actions or thinking processes which are related to recognition of key concepts and processes, and how meanings are constructed. *Will* indicates individual learning attitudes, acceptance of new information, will to concentrate and make efforts, and anxiety toward his own learning performance. *Self-regulation* describes the learner's ability to manage his personal learning process, especially how to plan, monitor, focus on and evaluate his own learning. Categorized by the above three components, the learning strategies refer to any *thoughts*, *behaviours*, *beliefs* or *emotions* that facilitate the acquisition, understanding or later transfer of new knowledge and skills (Tsai, 2009). This general framework coincides with the approach to contemporary language learning supporting the principles of communicative language learning and metacognition. In the following subsection, an overview of the most dominating classification of LLSs is given.

2.1 Language learning strategies

Different classifications of language learning strategies have been produced by many researchers (Rubin, 1975; Stern, 1975; O'Malley & Chamot, 1990) that gave their input to the six-strategy taxonomy of Rebecca Oxford, designed in 1990. Strategies are the learner's toolkit for active, conscious, purposeful and attentive learning, and they pave the way towards greater proficiency, learner autonomy and self-regulation (Hsiao & Oxford 2002). According to Oxford, LLSs have the features of contributing to the main goal, allowing learners to become more self-directed, being problem-oriented, including specific actions taken by the learner, involving many aspects of the learner, not just cognitive, supporting learning both directly and indirectly, being not always observable, being conscious, possible to be taught, flexible, and influenced by a variety of factors (1989). The system of LLSs developed by Rebecca Oxford is believed to be more comprehensive and detailed than earlier classification models by her predecessors (Jones, 1998). Oxford divided the LLS into two main categories: direct and indirect ones (Table 1). *Direct strategies* including memory, cognitive and compensation subgroups, directly involve the target language, such as reviewing and practising. *Indirect strategies* - metacognitive, affective and social ones, provide indirect support for language learning, such as planning, cooperating and seeking opportunities.

Table 1: Strategy groups based on R. Oxford (1990).

Direct strategies: directly involve the target language, such as reviewing and practising	**Memory** strategies: aid in entering information into long-term memory and retrieving information when needed for communication	* Creating mental images
		* Applying images and sounds
		* Reviewing well
	Cognitive strategies: used for forming and revising internal mental modes and receiving	* Practising
		* Receiving and sending messages

	and producing messages in the target language	* Analysing and reasoning
		* Creating structure for input and output
	Compensation strategies: needed to overcome any gaps in knowledge of the language	* Guessing intelligently
		* Overcoming limitations in speaking and writing
Indirect strategies: provide indirect support for language learning, such as planning, cooperating and seeking opportunities	**Metacognitive** strategies: help learners exercise executive control in planning, arranging, focusing, and evaluating their own learning process	* Centring your learning
		* Arranging and planning your learning
		* Evaluating your learning
	Affective strategies: enable learners to control feelings, motivation and attitudes related to language learning	* Lowering your anxiety
		* Encouraging yourself
		* Taking your emotional temperature
	Social strategies: facilitate interaction with others, often in a discourse situation	* Asking questions
		* Cooperating with others
		* Empathizing with others

Oxford considered both cognitive and metacognitive strategies necessary for efficient language learning and so they are both comprised in her framework (1990). In the context of language learning, the role of metacognition is frequently emphasized but it is not defined clearly enough. To be able to assess the learners' use of metacognitive strategies and self-regulation in general, it is important to understand the construct, its components and their interaction.

2.2 Self-regulated learning

There are several theories of self-regulated learning and numerous definitions which are important to understand the issues in this context. One of the initial, fundamental definitions comes from Bandura (1986), who incorporating it into his social cognitive theory of human behaviour, viewed *self-regulation* as the process of influencing the external environment by engaging in the functions of self-observation, self-judgment, and self-reaction. Drawing on his works, Zimmerman (1986) defined *self-regulated learning* as the process where students activate and sustain cognitions and behaviours systematically oriented toward the attainment of their learning goals. Winne (1996) accents the metacognitive perspective defining *self-regulated learning* as a metacognitively-governed behaviour where learners regulate their use of cognitive tactics and strategies. Another distinction between models of self-regulation is the postulated influence of the situation on self-regulation behaviour. Boekaerts (1997) defines *self-regulated learning* as a complex interaction between (meta)cognitive and motivational regulation. In her model that consisted of six components she differentiated both regulation systems in relation to three levels (goals, knowledge, and cognitive strategies).

The theories agree that self-regulated learning is an active and constructive process whereby students regulate different cognitive, metacognitive, motivational, volitional and behavioural processes during their learning (Winters *et al*, 2008). The numerous models of SRL that propose different constructs and conceptualizations share some general assumptions and features. Subsequently, Pintrich's framework based on Zimmerman's cyclical three-phase model and four assumptions will be explained.

2.2.1 Pintrich's conceptual framework for self-regulated learning

Pintrich's general framework for theory and research is based on four assumptions: active, constructive assumption; potential for control assumption; goal, criterion or standard assumption; and finally, mediators between personal and contextual characteristics and actual achievement or performance (2000). Drawing from these assumptions he defined SRL as an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features in the environment (Pintrich, 2000). His framework is a complete and comprehensive model that enables to describe a very complex concept of self-regulated learning. In his model Pintrich described the regulation of a learning process in four areas: cognition, motivation and affect, behaviour and context. In these areas he distinguished four phases: forethought and planning, monitoring, control and reflection. Regulation is the keyword which covers all phases and areas. Although Pintrich's framework is very elaborate and describes the system of strategies in detail, his instrument (Motivated Strategies for Learning Questionnaire) for measuring learners' strategy use which draws on his framework does not provide satisfactory factor structure or model fit indices (Davenport, 2003; Dunn *et al*, 2011; Hamilton &Akhter, 2009; Saks *et al*, 2014). That is why researchers are still looking for more reliable theories and designing more fit measuring instruments to assess learners' self-regulation (see e.g. Toering *et al*, 2012).

3. Supporting learners' self-regulated learning strategies

Acquiring sufficient self-regulated skills and this way reassuring better academic achievements (Dabbagh & Kitsantas, 2005; Schunk & Zimmerman, 1994) cannot be considered self-evident. These skills have to be instructed and supported throughout the whole learning process. This is a challenge for instructional designers and teachers to develop and apply effective strategies and encourage learners to develop their self-regulated learning skills in the learning process. The aim of the current study was to find out the most effective supports and conditions for scaffolding self-regulation reported in previous empirical studies. The following is an overview of the studies describing the most efficient interventions.

There are several effective ways to support learners self-regulation, starting with designing SRL assisted mechanisms in personalised e-learning systems (Chen, 2009) to persuasive mobile textings (Goh *et al*, 2012) and elaborated training programs (Bannert & Reimann, 2012). Based on researches attempting to support self-regulated learning Bannert identified three principles for effective intervention: first, instruction on self-regulated learning must be integrated with the domain-specific instruction being embedded in the subject matter; second, the application conditions and the usefulness of taught self-regulated learning strategies must be explained to students. Otherwise, students may feel disturbed and interrupted, and will not use them. To avoid this it is recommended to model and explain how these conditions support their learning. And third, it is important that sufficient training time is provided in order to internalize and automatize the self-regulated learning strategies and skills. (Bannert & Reimann, 2012). SRL can be supported following different principles. Hannafin distinguishes four types of scaffolds: *conceptual scaffolding* consists of aids that guide students' understanding of content. It guides learners regarding what to consider (Hannafin *et al*, 1999). *Metacognitive scaffolding* supports the underlying processes associated with individual learning management. It guides students' ways of thinking and reflecting on their task (e.g., training and prompts for self-monitoring and reflection). *Procedural scaffolding* shows how to utilize available resources and tools orienting to system features and functions. *Strategic scaffolding* involves alternative approaches to learning activity supporting analyzing, planning, strategy and tactical decisions (Hannafin *et al,* 1999).

Earlier researches have provided evidence that the most efficient support for learner's self-regulation is combined metacognitive scaffolding. Berthold (2007) used the combination of prompting and writing learning protocols for self-evaluation and feedback. Based on the results of several content and self-report tests and content analysis of learning protocols he reported that using prompts stimulated the elicitation of cognitive and metacognitive learning strategies. Also, academic results were better in the groups who received cognitive or the combination of cognitive and metacognitive prompts. The author concluded that cognitive and metacognitive strategies are not independent of each other but complementary. Metacognitive strategies control and regulate cognitive ones (Berthold *et al*, 2007). These findings are supported by Brunstein and Glaser (2011) who also prompted using cognitive (writing) strategies with self-regulated ones. Assessing the learners' writing skills (story quality, plan, text revisions, writing knowledge) and self-efficacy they concluded that an intervention that combined the instruction of writing strategies with self-regulation skills exerted a strong, coherent, and sustainable influence on procedural (planning and revising), declarative (knowledge), and self-related (self-efficacy) aspects of writing promoting novice writers' compositional achievements. Lee and his colleagues presented a study which examined the effects of two scaffolding strategies on learners' comprehension and self-regulation (2010). They combined generative learning strategy prompts and metacognitive feedback. Based on the results of knowledge and self-report SRL questionnaire (MSQL) they summarised that generative learning strategy prompts with metacognitive feedback improved learners' self-regulation and use of generative strategies and, accordingly, their learning performance. In contrast, generative learning strategy prompts without metacognitive feedback improved only learners' use of generative strategies (Lee *et al*, 2010). Similar effective metacognitive scaffolds have also been reported by Kramarski and Michalsky (2009; 2010) and Kramarski and Gutman (2006) who used IMPROVE self-questioning model.

Metacognitive scaffolding enables to foster several self-reported aspects of SRL, including self-monitoring, strategy use and interest (Winters *et al*, 2008) whereas the best results are achieved in the combination of cognitive and metacognitive support. With this knowledge we start designing our model for supporting learners' self-regulated learning skills.

4. Model development and an example of supporting cognitive and metacognitive strategies in an ESP course

Taking the theoretical frameworks of LLSs by Oxford (1990) and self-regulated learning by Pintrich (2000) as a basis, we developed a model for supporting learners' cognitive and metacognitive learning strategies for the blended English for Specific Purposes (ESP) course. Pintrich has distinguished four areas for SRL (cognition, motivation and affect, behaviour, and context) (Pintrich, 2000). In in this study we will focus on the area of cognition and behaviour throughout all four phases. Cognitive and metacognitive strategies were supported concurrently as the evidence has shown that combined scaffolding of the two gives the best results in the support of self-regulation as well as content knowledge (Berthold *et al*, 2007; Brunstein & Glaser, 2011; Lee *et al*, 2010).

For this course, four specific language learning assignments were created to support the development of students' metacognitive and cognitive LLSs (Table 3). The assignments were specially designed to take maximum advantage of the affordances of the digital learning environment (Moodle). Special attention was paid to students' active use of language when solving problems connected with real-life situations in the tourism industry (Tasks 3 and 4). The four assignments were accompanied by other tasks which are carried out in the class in the course of regular studies: reading and analysing texts, summarizing, comparing and contrasting etc. Students' interaction and communication are encouraged throughout the whole learning process, as well as in preparatory and follow-up phases. All learning activities are reflected orally in the classroom as well as in written form in students' learning diaries.

The first assignment designed for the intervention is compiling a learning plan. It starts with oral discussion in the class where students are prompted to think on the goals they could have when starting the course, also their needs considering their level of language skills, the cognitive strategies they are used to employing when learning a language, the ways of assessing and giving feedback that could be most beneficial for them. If this is a new activity for students it is important to encourage them to open up and express their doubts and expectations. It is also important to explain why it is necessary to set goals and plan their activities beforehand. As a follow-up activity, the students, following the prompts, write their answers to the digital learning plan which will be the first submission of their learning journals. Learning plan is not a finished document. Students are encouraged to return to it any time they feel that they could change or complete it. It is important to explain to the students that learning plan is an open document and their entries can be modified according to their needs, interests and level of development. The digital learning journal serves as a diary where students record their thoughts and impressions of the learning process. The journal, which is visible only for the learner himself and the teacher, gives information about the student's progress, problems he may face as well as his self-monitoring skills. This is also the place where the teacher can give feedback to students' progress and answer his questions.

The second assignment is writing an essay. The reason why this assignment was included in intervention was its focus on supporting reading, writing and compensation as cognitive strategies in addition to metacognitive ones. It starts with the class discussion again where the students are explained the assignment and interest towards the task is aroused. Subsequently, the prompts are used to activate students' prior content knowledge and metacognitive knowledge. This is followed by setting goals, making plans for writing and time planning. The strategy use which is prompted in the class activities is basically metacognitive. The cognitive strategies students need for writing are prompted in the digital learning environment in the form of questions and study-tips. These are accompanied by metacognitive prompting for monitoring and self-evaluating. Figure 1 illustrates a sample of metacognitive prompts on planning.

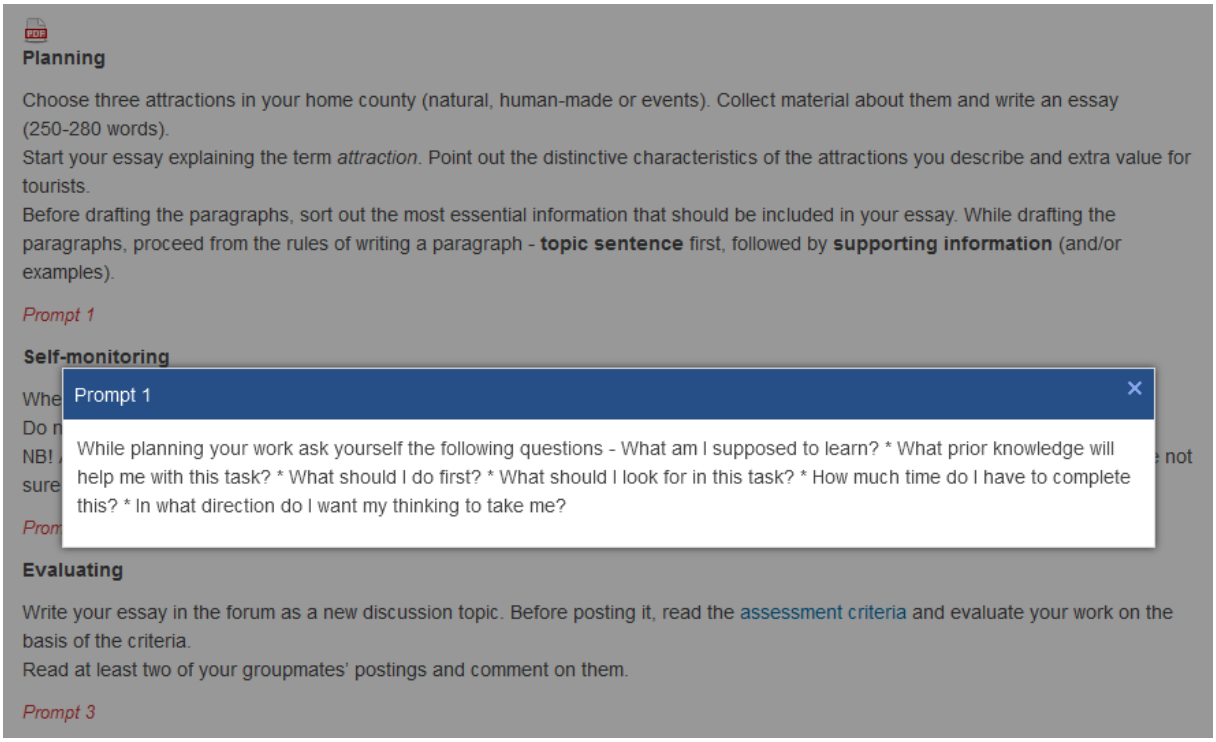


Figure 1. A sample screen capture of metacognitive prompting for Task 2.

As a follow-up activitiy in the class, the discussion on the whole writing process will be encouraged. Students are asked to share the problems they faced when writing, regulation processes they took up to overcome the problems and they are also asked to self-evaluate their activity throughout the whole process.

The third and fourth assignments are both pairworks and follow the same structure – collecting information about a destination or certain tourism enterprises, compiling the summary and making a presentation on the results. The main value of this task is its possibility to connect the real-life situation with language learning. The tasks are set this way that learners have to solve authentic problems using authentic materials and be able to justify their decisions. Similarly to the previous task, they also start with the class discussion to arouse interest and activate prior content and metacognitive knowledge. Since these tasks are pairworks and students set goals, plan their activities and time schedule, and divide the tasks together, it is important to apply social and active language use strategies. Also, the independent work in the digital learning environment which follows the preparatory work in the class, demands from students employing various cognitive and metacognitive strategies which are prompted throughout the tasks. The oral presentations in the class are followed by reflection of the whole process. Students are encouraged to share their impressions and self-evaluate their activity as an individual and as a pair.

When developing the model special attention was paid to providing the assignments with appropriate prompting, cognitive as well as metacognitive. Earlier studies have proven that the frequency of use of certain self-regulatory processes are consistently associated with learning gains, (Winters *et al*, 2008), therefore similar structure and similar logic of prompting has been used throughout the course. The model and the efficiency of the developed scaffolding will be tested within further researches.

Table 3: Model for supporting cognitive and metacognitive strategies in a blended ESP course.

Level 1 – F2F work in the classroom		Strategies supported with prompting: Metacognitive strategies		Strategies supported with prompting: Cognitive strategies
TASK 1 **Learning plan -** discussion based on questions given	oral prompting	setting goals for the course, planning the activities and strategies	oral prompting	oral expression, discussing, justifying opinion
TASK 2 **Essay -** instructions, discussion, making a plan	oral prompting	setting goals for the task, planning the activities, time planning	oral prompting	oral expression, discussing, justifying opinion
TASK 3 **Evaluating destination** (pair work) - instructions, discussion, making a plan	oral prompting	setting goals for the task, dividing tasks, planning the activities, time planning	oral prompting	oral expression, discussing, justifying opinion, making a presentation
TASK 4 **Comparing three tourism enterprises** (pair work) - instructions, discussion, making a plan	oral prompting	setting the goals for the task, dividing tasks, planning the activities, time planning	oral prompting	oral expression, discussing, justifying opinion, making a presentation

Level 2 - independent work in the web-based environment		Strategies supported with prompting: Metacognitive strategies		Strategies supported with prompting: Cognitive strategies
TASK 1 **Learning plan -** writing the learning plan answering the questions given	written prompting	setting goals for the course, planning the activities and strategies		writing str., compensation str.
TASK 2 **Essay -** working with resource materials, writing an essay, uploading it to the web, commenting on others' essays	written prompting	activating prior knowledge, time planning, monitoring his activity, evaluating the outcome of his activity	written prompting	searching information, analysing, critical evaluation, reading and writing str.
TASK 3 **Evaluating destination** (pair work) - working with resource materials, discussion, making a presentation, commenting on others' works	written prompting	activating prior knowledge, time planning, monitoring his activity, evaluating the outcome of his activity	written prompting	searching information, analysing, critical evaluation, reading and writing str.
TASK 4 **Comparing three tourism enterprises** (pair work) - working with resources, discussion, making a presentation, commenting on others' works	written prompting	activating prior knowledge, time planning, monitoring his activity, evaluating the outcome of his activity	written prompting	searching information, analysing, critical evaluation, reading and writing str.

Conclusion

The evidence has shown that for supporting students' language studies and proficiency and for enhancing their self-regulated learning skills, the use of their cognitive and metacognitive learning strategies has to be scaffolded (Brunstein & Glaser, 2011; Kondo *et al*, 2012; Nash-Ditzel, 2010; Wong, 2011). Based on earlier studies (Berthold *et al*, 2007; Brunstein & Glaser, 2011; Lee *et al*, 2010; Kramarski & Michalsky, 2009, 2010) we may confirm that learner's self-regulation can be supported, and the best results are provided by metacognitive scaffolds which are also the most frequently used type of scaffolding. Former studies have also proved that self-regulation is most efficiently enhanced by the combination of different strategy scaffolding, basically cognitive and metacognitive. The main utilized form of strategy support is prompting.

On the basis of the results of the analysis, a model and concrete assignments were developed to support learners' cognitive and metacognitive strategies. For intervention four specific language learning assignments were designed and accompanied with prompts. Prompting scaffolds strategy instruction in the classroom as well as in the digital learning environment. While designing the model it was considered necessary to provide various phases of the tasks with prompts of cognitive as well as metacognitive strategies. Students are supported to monitor and self-evaluate their learning activities and self-efficacy with constant discussions in the class. The model and the efficiency of the developed scaffolding will be tested within further researches.

References

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall.

Bannert, M., Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*, 193-211.

Benson, P. (2011). Teaching and researching autonomy. Great Britain: Pearson Education Ltd.

Berthold, K., Nückles, M., Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*, 564-577.

Boekaerts, M. (1997). Self-Regulated Learning: A new Concept embraced by Researchers, Policy Makers, Educators, Teachers, and Students. *Learning and Instruction, 7*, 161–186.

Brunstein, J. C., Glaser, C. (2011). Testing a Path-Analytic Mediation Model of How Self-Regulated Writing Strategies Improve Fourth Graders' Composition Skills: A Randomized Controlled Trial. *Journal of Educational Psychology, 103*(4), 922-938.

Chen, C.-M. (2009). Personalized E-learning system with self-regulated learning assisted mechanisms for promoting learning performance. *Expert Systems with Applications, 36*, 8816–8829.

Dabbagh, N., & Kitsantas, A. (2005). Using web-based pedagogical tools as scaffolds for self-regulated learning. *Instructional Science, 33*(5-6), 513-540.

Davenport, M. (2003). Modeling motivation and learning strategy use in the classroom: An assessment of the factorial, structural, and predictive validity of the Motivated Strategies for Learning Questionnaire (Doctoral Dissertation, Auburn University, 1999). *Dissertation Abstracts International, 64*, 273.

Dunn, K. E., Lo, W.-J., Mulvenon, S. W., Sutcliffe, R. (2011). Revisiting the Motivated Strategies for Learning Questionnaire: A theoretical and statistical reevaluation of the metacognitive self-regulation and effort regulation subscales. *Educational and Psychological Measurement, 72(2)*, 312-331.

Goh, T.-T., Seet, B.-C., Chen, N.-S. (2012). The impact of persuasive SMS on students' self-regulated learning. *British Journal of Educational Technology, 43*(4), 624–640.

Hamilton, R., & Akhter, S. (2009). Construct validity of the Motivated Strategies for Learning Questionnaire. *Psychological Reports, 104*(3), 711-722.

Hannafin, M., Land, S., Oliver, K. (1999). Open learning environments: foundation, methods, and models. In C. M. Reigeluth (Ed.), Instructional-design theories and models: A new paradigm of instructional theory, Vol. II (pp. 115–140). Mahwah, NJ, US: Lawrence Erlbaum.

Hsiao, T.-Y., Oxford, R. (2002). Comparing theories of language learning strategies: a confirmatory factor analysis. *The Modern Language Journal, 86*(3), 368–383.

Jones, S. (1998). Learning styles and learning strategies: towards learner independence. *Forum for Modern Language Studies, 34*, 115–129.

Kiewra, K. A. (2002). How Classroom Teachers Can Help Students Learn and Teach Them How to Learn. *Theory into Practice, 41*(2), 71-80.

Kondo, M., Ishikawa, Y., Smith, C., Sakamoto, K., Shimomura H., Wada, N. (2012). Mobile Assisted Language Learning in university EFL courses in Japan: developing attitudes and skills for self-regulated learning. *ReCALL, 24*(2), 169-187.

Kramarski, B., Gutman, M. (2006). How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning 22*, 24–33.

Kramarski, B., Michalsky, T. (2009). Investigating Preservice Teachers' Professional Growth in Self-Regulated Learning Environments. *Journal of Educational Psychology, 101*(1), 161-175.

Kramarski, B., Michalsky, T. (2010). Preparing preservice teachers for self-regulated learning in the context of technological pedagogical content knowledge. *Learning and Instruction*, *20*, 434-447.

Lee, H. W., Lim, K. Y., Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Educational Technology Research and Development, 58,* 629–648.

Nash-Ditzel, S. (2010). Metacognitive Reading Strategies Can Improve Self-Regulation. *Journal of College Reading and Learning*, *40*(2), 45-63.

O'Malley, J. M., Chamot, A. U. (1990). Learning strategies in second language acquisition. Cambridge: Cambridge University Press.

Oxford, R., Crookall, D. (1989). Research on language learning strategies: methods, findings, and instructional issues. *The Modern Language Journal*, *73*, 404–419.

Oxford, R. (1990). Language learning strategies: what every teacher should know. New York: Newbury House/Harper and Row.

Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In Boekaerts, M., Pintrich, P. R. & Zeidner, M. (Eds.) Handbook of Self-Regulation (pp. 531-566). Academic Press, Elsevier.

Rubin, J. (1975). What the "good language learner" can teach us. *TESOL Quarterly*, *9*, 41–51.

Saks, K., Leijen, Ä., Edovald, T., Õun, K. (in press). Cross-cultural adaptation and psychometric properties of the Estonian version of MSLQ. Procedia – Social and Behavioral Sciences.

Schunk, D.H., Zimmerman, B.J. (1994). Self-regulation of learning and performance: Issues and educational applications. Hillsdale, NJ: Erlbaum

Stern, H. H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review*, *31*, 304–318.

Toering, T., Elferink-Gemser, M. T., Jonker L, Heuvelen, M. J. G., Visscher, C. (2012). Measuring self-regulation in a learning context: Reliability and validity of the Self-Regulation of Learning Self-Report Scale (SRL-SRS). *International Journal of Sport and Exercise Psychology*, *10*(1), 24-38.

Tsai, M.-J. (2009). The Model of Strategic e-Learning: Understanding and Evaluating Student e-Learning from Metacognitive Perspectives. *Educational Technology & Society*, *12* (1), 34–48.

Weinstein, C. E. (1994). Strategic learning/strategic teaching: flip sides of a coin. In Pintrich, P. R., Brown, D.R. & Weinstein C. E. (Eds.) Student Motivation, Cognition, and Learning: Essays in Honor of Wilbert J. McKeachie (pp. 257-273). Hillsdale, NJ: Lawrence Erlbaum.

Winters, F. I., Greene, J. A., Costich, C. M. (2008). Self-Regulation of Learning within Computer-based Learning Environments: A Critical Analysis. *Educational Psychology Review*, *20*, 429-444.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and individual differences*, *8*, 327-353.

Wong, M. S-L. (2011). Language learning strategy use: a study of pre-service teachers in Malaysia. Retrieved from http://www.eric.ed.gov/PDFS/ED521415.pdf on 13.06.2013.

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses. *Contemporary Educational Psychology*, *11*, 307–313.

Liu, C.-C. et al. (eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education.

The Impacts of Using Interactive E-book on the Learning Effectiveness of English blank-filling cloze

Gwo-Jen HWANG[a], Yi-Hsuan HSIEH[a], Ching-Jung HSUEH[a], Ching-Kun HSU[b*]
[a]*Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan*
[b]*Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan*
*ckhsu@ntnu.edu.tw

Abstract: *This study used interactive e-book in English instruction and explored its benefits as well as negatives of 9th grader students' learning achievements in the English blank-filling cloze test. A questionnaire is used to investigate students' "perceived usefulness" and "perceived ease of use" toward using the interactive e-book in learning English. To find out the effects of using interactive E-book on English blank-filling cloze test, a four-lesson experiment was conducted in the English course in a junior high school. The experimental results show that although the students using interactive e-book did not outperform the ones receiving conventional instruction in learning English, the statistical feedbacks in terms of "perceived usefulness" and "perceived ease of use" on the questionnaire showed that the students highly appreciated and accepted the interactive e-book learning activity.*

Keywords: e-book, blank-filling cloze test, learning achievement, perceived usefulness, perceived ease of use

1. Background and objectives

1.1 Introduction

In the exam-oriented society, English is crucial for Taiwanese students. It is not only the first foreign language they learned in elementary school but also the fundamental key to successfully entering their ideal schools. What your English skill level is seems to be a significant qualification that everyone has to concern for studying or job-applying. With the state-of-the-art technology development, learners' learning habits change. The prosperity of the Internet and the trend of online application alter the traditional way people acquire the information thoroughly. Moreover, the new-generated knowledge revolution in reading and publishing shows that reading is not just reading, and books are not just books. Thus, the paper-texted books have been transformed into digital materials. To cater readers' multiple tastes, digitalizing the reading materials has been taken for granted.

People born from the mid- 1990s to 2000s are called “Z generation,” or “the Internet Generation.” Most of them grow up in a technology-surrounded environment and they have plenty of chances to obtain abundant multimedia information. Among them, electronic picture books, a kind of literature in digital format edited for children to read, are included. (De Jong & Bus, 2002; Korat, 2009) Some scholars pointed out using educational electronic book as a tool can support children’s literacy. Learning from a familiar electronic book makes students close to their own learning styles.

However, the e-books not only have digital characteristics to make students more attentive, but also have various interactive ways to promote readers’ reading effectiveness (De Jong & Bus, 2003, 2004). On the other hand, some scholars pointed out that we should encourage teachers in the language field use computer technology combining teaching strategies to improve the negative reading effectiveness (Adam & Wild, 1997).

1.2 Purpose of this study

As for the academic subject, English, the level for senior high is much higher than it for junior, especially in the numbers of recognized words. The learning gaps make freshmen who just attend senior high frustrated. In addition, lots of students claim that they can hardly find the key points to answer the English blank-filling close test.

The researchers for this project are both high school English teachers. According to the curriculum plan by The Ministry of Education in Taiwan, junior high school students need to know 1200 basic words and 2000 advanced words. Furthermore, 7000 recognized words for senior high school students. Thus, we decided to teach students the words formation rules. In other words, we show them the concept for distinguishing prefixes, roots words and suffixes. We are looking forward to helping students memorize new words by decomposing English words and lightening their learning burdens.

The teaching material adopted from the senior high English textbook, New Fareast version by Professor Chen Chunyin. In accordance with the schedule, we focus on Book 1 Lesson 7 (Into Aesop’s World). Meanwhile, we make some extension to teach students the general rules to distinguish the morphological features and make introductions about some common slang about animals. Within the teaching process, we make use of the interactive characteristic of E-books to promote students’ learning motivation and intensify their learning effect. What we do makes students pay less effort but achieve more. Most important of all, it can lower the cognitive loads and mental efforts.

The English blank-filling cloze test is considered tougher than any other part in the English test. To get high score, students have to recognize the words, the morphological features and idioms. The coherence between semantics and syntax is important as well. Consequently, in this study, an experiment has been conducted on the “the impact of using interactive e-book on English blank-filling close test." To evaluate the effectiveness of the proposed approach, the following research questions are investigated.

(1) Does the interactive e-book learning have negative or positive impact on the learning achievements of the students?

(2) Do the students who learn with the interactive e-book highly perceived usefulness and perceived ease of use?

2. Literature review

2.1 e-books

By means of the specialty of the computer multimedia, e-books including learning materials like texts, pictures, voices, videos and animations give readers a whole new reading experience. The so called e-books don't necessarily have the form of "books", but using all types of screen devices to combine specific software reading the digital contexts. They are in place of the traditional paper reading in the past.

The main purpose of the e-book topic aims to make students utilize this e-book to review after school on their own. Moreover, researchers look forward to making all students understand the learning contents by themselves. Diminishing the uneven educational development like "twin-peak" phenomenon and shortening the urban-rural gaps of the students' learning achievements are the researchers' top priorities.

Basically, the e-books are divided into two categories. One is for Reading Only and the other is for Reading and Playing. The former one focus on pronunciation of the whole article, but lacks of interactive games. On the contrary, Reading and Playing provides interactive games, multimedia effects and all kinds of animations. By clicking the hyperlink, readers have lots of chances to practice reading skills. (De Jong & Bus, 2003) This study belongs to the category of Reading and Playing. Inclusive of animations and multimedia effects, the e-book combining learning materials and interesting games makes students impressed while learning and motivates their learning interests.

2.2 Interactive multimedia teaching

Based on the instructional quality of the interaction, Misanchuk and Schwier (1992) identified three levels of interaction, namely reactive, proactive, and mutual interactions. (1) A reactive interaction: It is a response to a given question. Using the repetitious practices help learners to construct basic concepts. (2) Proactive interaction: According to learners' learning experiences, the system makes examples to help them interpret the concepts and present the learning materials in an understandable way. Learners can adopt the hierarchical structure as the most appropriate approach and enjoy the freedom to read the content wholly or partially. (3) Mutual interaction: In a mutual interaction environment, the learner and system are mutually adaptive in reactions with each other. Based on the data that the learners imported, the system will classify the thinking patterns of the learners and arrange suitable learning environment for learners to choose from under different learning situations. Besides, the system will show the consequence of different options. Through these constructive advices, learners read in an appropriate environment.

This study is in the form of reactive interaction. Researchers use teaching games and multimedia effects within the e-book like familiar topics and pictures to help learners practice repeatedly and acquire the concept constructively.

In sum, the main purpose of this study is to help learners to learn through the e-book anywhere, anytime without the assistance of others. Not limited to school time or any learning field. As long as the learners want to learn, he can download the e-book that he needs immediately. In this way, students' learning will be neither interrupted nor limited. Researchers hope that every student can study independently and be comprehensive about the learning materials through the e-book. Most important of all, improving the "twin-peak" phenomenon and shortening the urban-rural gaps of the students' learning achievements can solve the uneven distribution of the teaching resources and human resources problem.

3. Experiment design

3.1 Participants

The subjects included two classes of 9th graders of a private high school in New Taipei City, Taiwan, R.O.C. A total of 109 students voluntarily participated in the study. One class was assigned to be the experimental group and the other was the control group. There are 54 students in the experimental group, who use interactive e-books to learn, while there are 55 students in the control group, who learn by the traditional way of the lecture and the worksheets. The average age of the subjects was 15. The experiment was conducted in the Advanced Placement English course. All of the students were taught by the same instructor who had taught the English course for over 10 years.

3.2. Instruments

The learning achievement test was developed by two experienced teachers. The questions in the test sheets were adapted from the teaching material in class. The pre-test scores were collected from the blank- filling cloze test in the weekly test student already took before the learning activity. The pre-test aimed to evaluate students' prior knowledge toward the blank-filling cloze test. The study collected 6 times of the scores. Every test consisted of five multiple choice questions, and each question accounts for 1 point, with a total of 6 times of the scores of 30. The post-test was 30 questions, and each question accounts for 1 point, with a total of the scores of 30. The post-test aimed to evaluate the students' learning achievement in the learning activity.

The questionnaire used to evaluate students' "perceived usefulness" and "perceived ease of use" was adapted from Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, Judy C. R. (2010) using a six-point Likert rating scheme. The questionnaire included 9 items, such as "I think it is helpful to use this kind of learning mode when learning something new."; "I don't need to spend too much time or energy during the learning process"

3.3. Experimental Procedures

Before the experiment, the two groups of the students took a pre-test, which was a regular part of the test every week. During the learning activity, the students in the experimental group learned with the interactive e-books, while the students in the control group learned with the conventional way of lecture and the worksheets. It was worth mentioning that the way with which the control group learned was just the way they used to do. That is, the students in the control group reflected the way students' learned with conventional paper-based instruction before. The whole teaching procedure lasted for 4 periods of the class in a total of 200 minutes.

After the learning activity, students took the post-test and completed the questionnaire of the "perceived usefulness" and "perceived ease of use" to evaluate students' learning achievement and their perceptions toward the "perceived usefulness" and "perceived ease of use".

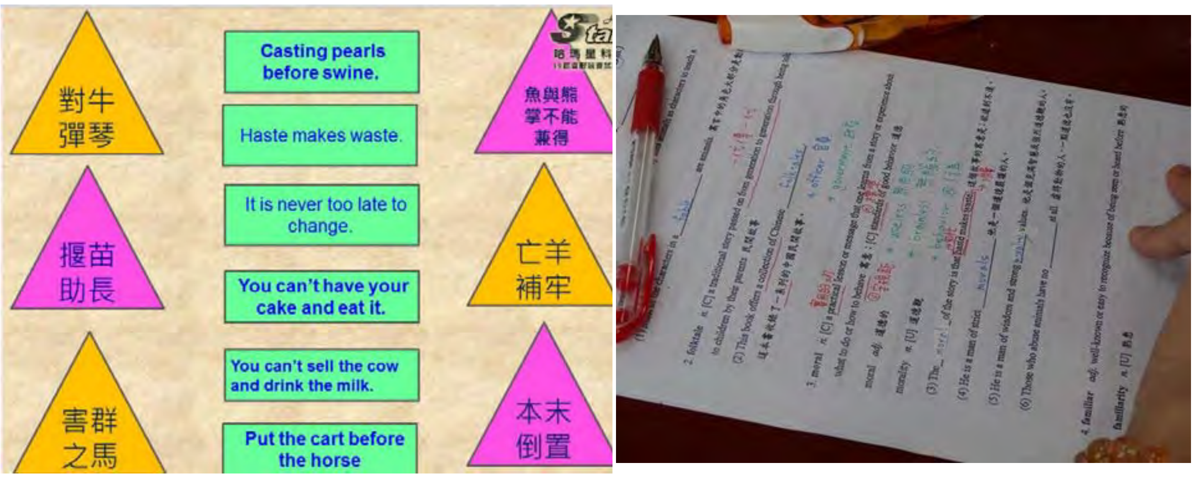


Figure 1. (Left) A page of the interactive e-books for the experimental group;
(Right) A page of worksheets for the control group.

4. Research results

4.1. Analysis of learning achievements

The purpose of this study was to examine the impact of using interactive e-book on

English blank-filling cloze test. The pretest scores of the two classes are not significant ($t = 1.20; p>.05$). After the independent t-test on the post-test ($t = 0.87; p> .05$), the results still showed no significant differences in test scores between two groups. Consequently, this study implied that the junior high school students using e-book to instruct foreign language could not make sure to outperform those using paper-based instruction when learning the blank-filling cloze.

Table 1. t-test of the pretest and post-test

Group	N	Pre-test			Post-test		
		Mean	SD	*t*	Mean	SD	*t*
Experimental group	54	16.15	6.12	*1.20*	11.78	7.16	*0.87*
Control group	55	14.53	7.90		10.58	7.24	

4.2. Analysis of "perceived usefulness" and "perceived ease of use"

In order to understand students' "perceived usefulness" and "perceived ease of use" toward interactive e-books, the study utilized a questionnaire to evaluate. As shown in Table 2, the average perception of the "perceived usefulness" of the experimental group was 4.28 out of a Likert 6-point scale, it meant students hold a positive attitude towards the "perceived usefulness"; and the average "perceived ease of use" of the experimental group was as high as 4.21 out of a Likert 6-point scale, showing the students generally agreed that interactive e-book is easy to operate. To sum up, the study results suggest that students' hold a positive attitude toward learning with interactive e-books. Although the learning effectiveness of using e-books in practicing blank-filling cloze was not significantly better than the learning effectiveness of paper-based instruction, the students were willing to accept new technology in language learning. Therefore, future learning will not be limited in paper-based instruction.

Table 2. Descriptive data of "perceived usefulness" and "perceived ease of use"

Category	N	Mean	SD
perceived usefulness	54	4.28	0.74
perceived ease of use	54	4.21	0.85

5. Discussions and conclusions

This study integrates the interactive e-books into English teaching, hoping to assist students' skills in answering questions in the English blank-filling cloze test. It expected that the effectiveness of learning English blank-filling cloze could be not decreased due to the interference of electronic media. The experimental results of this study found that the learning achievement of e-book instruction was as good as paper-based instruction. There was no significant difference between the e-book instruction and paper-based instruction. Furthermore, the students' feedback from the questionnaire shows that interactive e-book really enhance students' interest in learning English. The students in the experimental group highly perceived usefulness and ease-of-use on average. Through the operation of e-books and educational games, students look forward to each page of the e-books and anticipate the surprise the books bring to them. During the experimental process, the researchers find that students, even the low-achievers who were easily doze off in class, actively participate in classroom activities with the help of the e-books.

The pictures and games embedded in the e-books succeeded attracting the attention of students, allowing them to reveal a pleasant smile in the learning process, and hence the atmosphere in the classroom is happy and harmonious. Integrating technology into teaching create a win-win situation for both teacher and students. Although due to limited funds and manpower, students cannot hold and operate a Tablet PC individually to use e-books this time, yet this interactive e-books system truly presents the learner-centered, teacher-assisted blueprint. Students think jointly the content of e-books, discussing with each other, and giving immediate feedback. All these improve the learning outcomes and the classroom atmosphere. After school, students can also download e-books to do the review or preview from the cloud space.

Although the subjects were all from high school Advanced Placement classes, those students still have to prepare the Comprehensive Assessment Program for Junior High School Student. These subjects may therefore like candles burning at both ends; they cannot very well prepare the high school curriculum. Thus, the learning achievement research results are not significant. However, feedback from the questionnaire which shows that considering the "perceived usefulness" and "perceived ease of use", the experimental group highly approved and accepted interactive e-books. Nevertheless, the main limitation of this study is the time limitation. Because graduation ceremony and educational examinations are around the corner, the experiment cannot last long. The study recommends that for the future study, the experiment time may require a month at least, the effectiveness will be more obvious.

Acknowledgements

This study is supported in part by the Ministry of Technology and Science in Taiwan under contract numbers: NSC 102-2511-S-003-055-MY2, MOST 103-2628-S-003-003-MY2, and NSC 102-2511-S-011 -007 -MY3.

6. References

Adam, N. and Wild, M. (1997). Applying CD-ROM interactive storybooks to learning to read. *Journal of Computer Assisted Learning, 13*, 119-132.

De Jong, M.T., & Bus, A.G. (2002). Quality of book-reading matters for emergent readers: An experiment with the same book in a regular or electronic format. *Journal of Education Psychology, 94*(4),145-155.

De Jong, M.T., & Bus, A.G. (2003) How well suited are electronic books to supporting literacy?. *Journal of Early Childhood Literacy, 3*(2), 147-164.

De Jong, M.T., & Bus, A.G. (2004). The efficacy of electronic books in fostering kindergarten children's emergent story understanding. *Reading Research Quarterly, 39*(4), 378-393.

Korat, O. (2009). Reading electronic books as a support for vocabulary, story comprehension and word reading in kindergarten and first grade. *Computers and Education, 55,* 24-31.

Korat, O., & Shamir, A. (2012). Direct and indirect teaching: Using e-books for supporting vocabulary, word reading, and story comprehension for young children. *Journal of Educational Computing Research, 46*(2), 135-152.

Kynaslahti, H. (2003) In Search of Elements of Mobility in the Context of Education. *Mobile Learning* (eds. H. Kynaslahti & P. Seppala) , 41- 48.

Misanchuk, Earl R.; Schwier, Richard A. (1992) Discusses audit trails (i.e., records of learner responses) in interactive multimedia and hypermedia instruction. *Journal of Educational Multimedia and Hypermedia*, v1 n3 , 355-72.

Pinkwart, N. & Hoope, H. U. & Milrad, M. & Perez. J. (2003). Educational scenarios for cooperative use of Personal Digital Assistants. *Journal of Computer Assisted Learning, 19*, 383-391.

Correlation of English Test Outcome From TVE Joint College Entrance Examination of Taiwan VS. Professional English Reading Speed and Comprehension

***Hong-Fa HO*[a], *Yi-Yeh CHUNG*[a], *Chen-Ku LIN*[b], *Chen-Hsiung LIN*[c], & *SOH O-K*[d]**
[a]*Department of Electrical Engineering, National Taiwan Normal University, Taiwan (R.O.C.)*
[b]*Ph.D. Graduate Institute of Political Science, Taiwan Normal University (R.O.C.)*
[c]*Department of Industrial Education, National Taiwan Normal University (R.O.C.)*
[d]*School of Languages & Linguistics, University Kebangsaan Malaysia*
* 3rrr168@gmail.com

Abstract: Educators throughout history have tried to evaluate the different aspects of learning performances. Notice the importance of learning for understanding and time spend studying, this research explores the correlation based on performance outcome from English examination of the English as foreign language (EFL) participants to professional English reading performances on reading speed and reading comprehension, using eye chasing device for observation. 15 student volunteers with technical background participated with experiment. SPSS 21 was used for descriptive and Pearson correlation statistics, with both hypotheses accepted on positive correlation among the English testing outcomes of technical background participants, who have recently completed their Technological and Vocational Education Joint College Entrance Examination of Taiwan 2014, and professional English reading speed and comprehension. Realizing that prior knowledge helps with reading, the authors recommend the use of intensive reading practices of both contextual and graphical images while training the EFL learners. Once they build up solid foundation of broader knowledge, they will be able to read faster and comprehend better.

Keywords: English subject testing outcome (from Technological and Vocational Education Joint College Entrance Examination of Taiwan, TVE JCEET), reading speed (RS), reading comprehension (RC), eye chasing device (ECD), professional English reading (PER)

1. Introduction

1.1 Purpose and Objective

Educators throughout history have tried to evaluate the different aspects of learning performances (Alexander, 2000). Knowing the importance of learning for understanding and time spend studying, this research explores how testing outcome of the participants of English test of Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET), from vocational and technical high school, correlate with professional English reading performances on reading speed and reading comprehension, using the eye chasing device for observation. This research aims at finding possible answers on both technical and vocational students' visual reading formation. The research also attempts to offer suggestion on the way to improve EFL students' visual learning with education. To serve these purposes, the authors have come up with the following objectives:

1. To explore the differences of English subject testing outcome (from Technological and Vocational Education Joint College Entrance Examination of Taiwan) and professional English reading speed among technological and vocational students.
2. To detect the relationship between English subject testing outcome (from Technological and Vocational Education Joint College Entrance Examination of Taiwan) and reading comprehension among technological and vocational students.

2. Literature Review

2.1 Background on Reading Comprehension

A growing number of eye chasing tests have yielded a handful of findings relevant to reading experiences (Biedert, Buscher, & Dengel, 2010; Lowell & Morris, 2014; Mayer, 2010). Nevertheless, using eye chasing application to identify the implication between English testing outcome of TVE JCEET, and professional English contents' reading speed and reading comprehension have not been examined. Hence, the authors attempt to fill the gap.

Paivio (1990)'s dual coding theory helps on explaining the reading comprehension through text-diagram representations. The theory describes that the text and diagram representations are deposited in various cognitive systems due to variety of physical formats (Clark & Clark, 2010; Ho, Tsai, Wang, & Tsai, 2014; Paivio, 1990). Therefore, a combination of scanning or reading diagram and text both could be intensely used to synthesize and enhance cognitive information processing for comprehension than using text or diagram alone (Clark & Clark, 2010). The combinations are important because the information are imperative knowledge for speed of processing and understanding and they can also be expanded for broad range of advanced learning, convert into skills for processing and application, and as outcomes of performance output (Lin, 2011; Orquin & Mueller Loose, 2013). It is stated that the longer the duration time was, based on former research experimental results, the greater the amount of attention that was allocated and accessed (de Koning, Tabbers, Rikers, & Paas, 2010; Ho et al., 2014).

The more the readers know about the subject, the easier the reading understanding~~s~~ the reader gains. The authors therefore perceive that familiarity with professional English subject contents may also reflects on the timing of reading and gazing at professional English articles. Efficient readers trigger their prior knowledge for reflecting on their schemata, predicting the layouts and outcomes, generalizing and actively reading for answers, as these characteristics enhance understanding and speed (Anderson, 1994; Fitzgerald, 1995; Goodman & Goodman, 1978; Peregoy & Boyle, 2000), even when they pursuit professional English reading. This quality is ~~due~~ associated with the factors that knowledge is able to be generalized, conceptualized, and theorized to encompass various scopes, transferred to multiple applicability, and enhanced for creativity (Lin, 2011). Base on previous findings, the authors have identified that with the longer duration of reading time, the greater the attention spent on professional English reading, while the reading speed may expand, the reading comprehension may intensely extend.

2.2 Eye Chasing Studies and Reading Speed

In 1993, Hegarty and Just conducted a research study by using reading with help of visual chasing framework. In this research, they have revealed that participants of technical background regardless of their high and low mechanical abilities in their reading behaviors tended to read some clauses or sentences before viewing diagram to construct relationship. Readers' prior knowledge of subjects and strategies of reading cognitively and psychologically can help efficient readers to associate, inference, and form speed and recognition of professional reading contents and materials (Ehrlich & Johnson-Laird, 1982; Glenberg & Langston, 1992; Hegarty & Just, 1993; Just & Carpenter, 1987). Mayer (2003) on the other hand, described that "learners' cognitive integration process involves selecting appropriate aspects of words or images, building coherent visual and verbal mental models from each representation, and finally incorporating both mental models based on learners' prior knowledge to generate learning" (Mayer, 2003). As in conjunction to Myer's study, the later experiments conducted also indicates that, as revealed by heat map of eye chasing device, visual distribution of participants tend to spend more time reading the textual than the graphical records. Participants of high prior knowledge (high PK) showed longer fixation durations than low prior knowledge (low PK) participants. The high PK participants revealed more inter-scanning transitions than low PK participants between words and pictures, but diagrams as well. Therefore, it has been learned that high PK participants are more capable of integrating text and graphic information (Ho et al., 2014). With certain subjects being more intensive and complex, i.e. specialty and expert contents, than general readings, with the latter being more frequently studied, the authors thus focus on the lesser known of professional English reading. Syndicating and generalizing the findings from previous studies, the authors recognize that prior knowledge will help with reading comprehension and forming generalization of understanding to other lesser familiar professional readings, as aforementioned factors, including but not limited to prior knowledge concept may enhance understanding, could also be features influencing professional English reading speed and comprehension of this research's participants.

2.3 Technological and Vocational Education Joint College Entrance Examination of Taiwan

The Technological and Vocational Education (TVE) system is important to nurture human resources in Taiwan. Under the commission of Taiwan Ministry of Education (MOE), the nationwide TVE Joint College Entrance Examination, of various professional fields, including English subject as one of the mandatory tests, has been used for technological and vocational colleges and universities throughout Taiwan (TCTVET, 2014). Therefore, this test is a well-respected and legitimate exam. The outcomes and associated references from TVE JCEET consequently provide the relevance for this research goal.

2.4 Research Questions

To accomplish the purpose of this study, the following two questions were proposed.

1. What is the relationship between English examination outcomes from TVE JCEET, for people of English as foreign language (EFL), with technical background (TB), to professional English reading speed?
2. What is the relationship between English examination outcomes from TVE JCEET, for people of English as foreign language (EFL), with technical background (TB), to professional English reading comprehension?

2.5 Hypothesis

Based on the questions proposed, this research intends to test the following hypotheses:

1. There is a positive correlation between the participants of technical background, earning a high English testing outcome from Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET), to be able to read professional English articles faster than those with a lower score, In words, despite of reading familiar or unfamiliar professional English subject, participants with technical background, those who have received high scores on the English examination of TVE JCEET are able to read fast, those with lower score are reading slowly.
2. There is a positive correlation between the participants of technical background, earning a high English testing outcome from Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET), to have positive professional English reading comprehension. In other words, despite of reading familiar or unfamiliar professional English subject, participants with both technical and vocational background, are able to achieve positive outcome because they have done well in English of TVE JCEET.

3. Methodology

3.1 Research Framework and Process

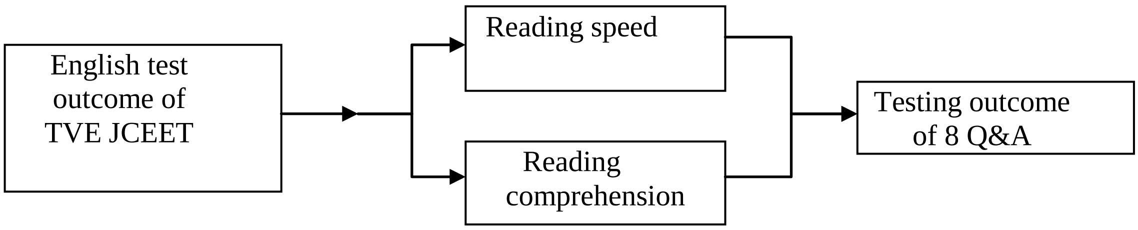


Figure 1. Eye Chasing Device Was Used for Professional English Reading Eye Movement.

3.2 Material

To conduct this research experiment: the authors have come up with 8 professional English reading articles similar to Taiwan college entry level of English proficiency, with 2 from computer engineering subject, 2 from mechanics subject, 2 from bio-medic subject, and 2 from business fields, along with 8 multiple choice questionnaires, one for each of the 8 articles, preinstalled them into laptop, for participants to read on screen display and respond through mouse clicks, to test participants' reading comprehension and speed, without participants' knowledge of contents in advance. To observe eye movements while reading, eye chasing device, head stabilizer, laptop computer were used. All participants read and answered all 8 assigned professional English readings and questionnaires, both.

3.3 Participant

To examine the English reading proficiency of those technical background participants, 15 students from the Na Fu Vocational and Technical High School of Taiwan, who completed their Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET) in 2014, was invited to the pilot test. It is with the help of eye chasing device that the eye movements, such as, region of interest (ROI) and total contact time (TCT) by participants, are made available for analysis.

3.4 Design and Procedure

Participants are told to read 8 English professional articles and answer 8 questions after finish reading each one of the 8 articles. Participants were then put to test at this experiment. One National Taiwan Normal University (NTNU) Professor and 5 National Taiwan Normal University (NTNU) research assistants proctored the experiment at Na Hu High School. 4 sets of same eye chasing devices were mounted, for prompt testing, to observe the eye movement of participants when they read and answered the questions.

3.6 Instruments

Four recently made of same EyeNTNU-180 eye chasers created by National Taiwan Normal University Electronic Department team, and sponsored by the Taiwan' Ministry of Education and the Ministry of Science and Technology, were used to monitor and record eye movements of 15 participants while reading professional English articles and responding to questionnaires and answers Q&A. Each of EyeNTNU-180 set includes a laptop computer in front of a camera to record eye movements. The laptop presented professional English articles and questionnaires. In this experiment, a sampling rate is 180Hz. To avoid errors in eye chasing measurement caused by shaking of head and inconsistency of eye movement, a head stabilizer was used to fix head position. The distance between the screen and participants were set to 60cm straight apart. The normal time duration of a fixation of this experiment was set to 80 milliseconds.

3.7 Data Collection

The recently produced EyeNTNU-180 eye chasing device was adopted to monitor eye movements, to assess time of professional English reading, and to site coordination. By identifying the parameters and ROI (Region of Interest), the authors were able to determine the reading speed and possible comprehension, when measured against the testing outcomes from TVE JCEET English subject and the 8 Questionnaires and Answers after reading the arranged professional English articles. The authors had designed a computer program to examine the data of inspected components. The ROIs were categorized by ROI-splitter software and eye movement analyzer to evaluate eye movement data to generate scan paths, total gazing time, and eye contacts. Data are presented as following:

Participant	TVE JCEET of English test outcome	RC
1	78	6
2	72	4
3	70	4
4	68	4
5	56	4
6	52	4
7	48	3
8	48	3
9	48	3
10	46	3
11	44	2
12	44	2
13	42	2
14	42	1
15	42	1

Figure 1. The outcome of English subject TVE JCEET scores, for each one of the 15 participants, and the amount of Q&A answered correctly out of 8 English professional readings, identified as reading comprehension (RC).

Professional English Reading Subject	Average Words per Minute (WPM)	Standard Deviation
Mechanical	3777.5214	10741.04128
Electrical engineering	1635.2505	1624.66644
Bio-Medic	2368.9379	3272.17114
Business	1143.5126	753.10136

Figure 2. Descriptive statistic of average words per minute (WPM)~~,~~ and Standard deviation, read by all 15 participants together for each one of the 4 professional English subjects, by combining the outcomes of 2 questions per subject together into four aforementioned professional subjects. Based on the experiment, it is noticed that technical background participants tend to read Mechanical field at 3777.5214 WPM, more than Business subject of 1143.5126. To account for the individual participant's WPM, the authors took total number of words in the article / (tct/(1000x60)), tct stands for total

eye gazing time used.

	TVE JCEET	Professional English Reading Comprehension	Mechanical	Electrical Engineering	Bio-Medic	Business
Pearson correlation	1	.865**	.377	-.105	-.340	-.372

Explain: ** p<0.001 ; * p<0.01 , N=15

Figure 3. Two tails Pearson correlation between the 4 professional English readings and TVE JCEET English testing outcome.

4. Results and Discussions

4.1 TVE JCEET English and Professional English Reading tests' outcomes, Statistic Report

Figure 1 shows the outcome of English subject TVE JCEET scores, for each one of the 15 participants, and the number of Q&As answered correctly out of the 8 English professional readings from this correlation experiment, identified as reading comprehension RC. The authors hypothesized that participants of technical background who has received higher TVE JCEET English scores would tend to comprehend better when they attend to professional English reading, taking their exam outcomes as indictors. This hypothesis is reflected in Figure 1 with higher TVE JCEET earners getting higher RC scores. Figure 2 is the descriptive statistic, (by combining the results of two questions per subjects together and blending them into 4 professional English categories) which shows the average words per minute (WPM), and the standard deviations, read by all 15 participants together for each one of the 4 subjects examined. Participants of technical background on average tend to read more WPM on Mechanical, followed by Bio-Medic, then Computer Engineering, and finally Business professional subjects. To account for the individual participant's WPM, the authors took total number of words in the article / (tct/(1000x60)), tct stands for total eye gazing time used.

Figure 3, indicates that the two tails Pearson correlation table shows significant relationship between TVE JCEET English testing outcome and RC, number of questions answered correctly after reading the 8 professional English articles, but specific subjects. The examination outcomes strengthen the explanation that testing outcome and comprehension have positive correlation as stated in hypothesis 2, regardless of professional subjects read. The descriptive statistics on standard deviation and average TVE JCEET English test outcomes, and professional English reading Q&As answered correctly by the 15 pilot test participants are listed in Figure. 4. From this experiment, the authors have identified that TVE JCEET English testing outcome has positive relationships on reading speed, and professional English reading comprehension.

Professional English reading subject	Average	Standard Deviation	N
TVE JCEET English test outcome	53.33	12.414	15
Professional English reading comprehension	3.07	1.335	15

Figure 4. Descriptive Statistic of standard deviation and average TVE JCEET English test outcomes and professional English reading Q&As answered correctly.

5. Conclusion

Performance is an important output for educational results that educators have strived to assess and invent different features of learning practices. Knowing the importance of learning for comprehension and the amount of time spend studying, this research intends to explore how testing outputs of the participants of English subject of Technological and Vocational Education Joint College Entrance Examination of Taiwan (TVE JCEET), from vocational and technical high school, link with professional English reading performances on reading speed and reading comprehension, using eye chasing device for surveillance. With Paivio (1990) dual coding theory and its supporters advocate on the benefits of reading comprehension, through both text-diagram representations, the authors combined literature review and this research experiment corresponded that reading of textual or graphical images alone would not be as effective in terms of speed and comprehension as scanning through both together. While knowledge and information processing are able to be transferred, synthesized, and created to encompass broader scopes of outputs (Ho, 2014; Lin, 2011), and with the results of this experiment, the authors therefore accept the two hypotheses proposed, because they have showed positive correlations on English subject testing outcome of TVE JCEET to professional English reading speed and comprehension.

The more and in-depth of the existing knowledge and training, the faster the one is able to read for comprehension, as the same applies for professional English subjects of familiar and unfamiliar field for those of English as foreign language practitioners. The authors would recommend of intensive reading practices of contextual and graphical images while training the EFL learners. Once they build up the solid foundation of broader knowledge, they will be able to read faster and know better.

Acknowledgements

This research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the "International Research-Intensive Center of Excellence Program" of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. NSC 103-2911-I-003-301. We would also like to thank Hao-Yuan Lin who has prepared and revised previous versions of this document.

References

Alexander, F. K. (2000). The changing face of accountability: Monitoring and assessing institutional performance in higher education. Journal of Higher Education, 411-431.

Anderson, R. C. (1994). Role of the reader's schema in comprehension, learning, and memory.

Clark, R. E., & Clark, V. P. (2010). From neo-behaviorism to neuroscience: Perspectives on the origins and future contributions of cognitive load research. *Cognitive load theory*, 203-228.

de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and instruction, 20*(2), 111-122.

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of verbal learning and verbal behavior, 21*(3), 296-306.

Fitzgerald, J. (1995). English-as-a-second-language learners' cognitive reading processes: A review of research in the United States. *Review of Educational research, 65*(2), 145-190.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of memory and language, 31*(2), 129-151.

Goodman, K. S., & Goodman, Y. M. (1978). Reading of American Children Whose Language is a Stable Rural Dialect of English or a Language Other Than English. Final Report.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language, 32*(6), 717-742.

Ho, H. N. J., Tsai, M. J., Wang, C. Y., & Tsai, C. C. (2014). PRIOR KNOWLEDGE AND ONLINE INQUIRY-BASED SCIENCE READING: EVIDENCE FROM EYE

TRACKING. *International Journal of Science and Mathematics Education, 12*(3), 525-554. doi: 10.1007/s10763-013-9489-6
Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*: Allyn & Bacon.
Lin, H.-Y. (2011). The impact of intellectual capital with knowledge management on Taiwan MICE industry (Master thesis). *National Taiwan Normal University Graduate Institute of International Human Resources Development*, 1-117.
Lowell, R., & Morris, R. K. (2014). Word length effects on novel words: Evidence from eye movements. *Attention, Perception, & Psychophysics*, *76*(1), 179-189.
Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and instruction*, *13*(2), 125-139.
Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and instruction*, *20*(2), 167-171.
Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: a review on eye movements in decision making. *Acta psychologica*, *144*(1), 190-206.
Paivio, A. (1990). *Mental representations: A dual coding approach*: Oxford University Press.
Peregoy, S. F., & Boyle, O. F. (2000). English learners reading English: What we know, what we need to know. *Theory into practice, 39*(4), 237-247.
Testing Center for Technological & Vocational Education of Taiwan (TCTVET). (2014). Retrieved from http://www.tcte.edu.tw/en/

A Tablet-based Chinese Composition Assessment System

Kat LEUNG[a], Barley MAK[b] & Howard LEUNG[c*]
[a]*Caritas Institute of Higher Education, Hong Kong*
[b]*The Chinese University of Hong Kong, Hong Kong*
[c]*City University of Hong Kong, Hong Kong*
*howard@cityu.edu.hk

Abstract: Chinese composition is an important aspect for students to learn the Chinese language. In this paper, we present our tablet-based Chinese composition assessment system. We have implemented a system to be run on a pen-based tablet device to help teachers assess student's composition. Our system is able to display student's composition work on the tablet device, allowing teachers to markup on the student's work to provide written comments and record his/her voice to provide oral comments. The recorded assessment session can be played back such that the written comments can be shown which are synchronized with the oral comments. Our system facilitates teachers to provide both written and oral comments easily which can be stored and played back. Our system can also be used by researchers to analyze teachers' assessment process by studying their feedback.

Keywords: Chinese composition, tablet, assessment, written and oral feedback

1. Introduction

Various new technologies have been proposed to help students learn different aspects of the Chinese language. Novel technology has been developed for mobile device to help teachers implement reading strategy instruction and support students' individual and co-operative reading activities in Chinese language classes (Chang, Lan, Chang, & Sung, 2010). A measure has been proposed to assess the ability of Chinese character recognition made by foreigner Chinese learners and a two-phase learning strategy has been presented for these learners to learn Chinese characters (Ho & Lin, 2010). A mobile-assisted game was adopted to let students learn the formation of Chinese characters and the social interactions have been examined to analyze how student grouping can affect the learning outcome (Wong, Hsu, Sun, & Boticki, 2013).

Many research papers focused on Chinese composition which is an important component in learning the Chinese language. Various strategies and activities have been proposed which aims to enhance the students' skill in Chinese composition. Blogging was adopted in a Chinese composition class in a high school in Taiwan (Lou, Wu, & Smith, 2010). They concluded that students' writing skills were enhanced in a positive way. In the study by (Ying, Leung, Lee, & Chow, 2014), the authors focused on identifying the role of teachers in Chinese composition marking as well as determining the difference between traditional paper-based marking and tablet-based marking. Two teachers reported the usual time taken to mark a composition in the traditional paper-based marking which was found to be similar to the time required for them to mark a composition using our proposed tablet-based composition marking system.

In terms of teaching Chinese composition, Li (2005) proposed two new methods for teaching Chinese composition and described some example lesson plans with games that aimed to increase students' interest in doing composition in class. They had shown significant increase in writing speed for junior secondary students over 13 weeks after applying their proposed method. Sim (2005) described the teaching of Chinese composition in Singapore by comparing the past and the present. So (2005) believed that information technology can bring new opportunities in teaching Chinese composition. In particular, So (2005) gave a visionary example that "the students can also watch their essays being marked from their computer terminals at any time".

Shum (2005) compared four different methods in evaluating Chinese compositions of senior secondary school students in Hong Kong: 1) teacher provided detailed feedback to the students; 2) teacher used symbolic codes to mark the mistakes; 3) the evaluation was based on peer assessment according to a checklist; and 4) the evaluation was self-conducted according to a checklist. The survey results showed that the first group, i.e., the detailed teacher evaluation, was a much more popular method perceived by the students than the second group, i.e., the teacher evaluation with symbolic codes. This indicated that students prefer detailed feedback and our system provides an easy way for teachers to give written as well as oral feedback. Lo (2006) stated that two methods are commonly used to provide formative assessment in school: 1) individual, face-to-face, oral feedback in class; and 2) assessment paper review with follow-up remedial work or activities. Lo (2006) also mentioned that one of the challenges in employing the first method is the heavy workload of the teachers for preparing the oral feedback. The sound recording feature in our system allows teachers to record their voice feedback while they assess the student's composition so that it will take less time for the teachers for preparing the oral feedback.

In this paper, we present our tablet-based Chinese composition assessment system. We have implemented a system to be run on a pen-based tablet device to help teachers assess student's composition. Our system is able to display student's composition work on the tablet device, allowing teachers to markup on the student's work to provide written comments and record his/her voice to provide oral comments. The recorded assessment session can be played back such that the written comments can be shown which are synchronized with the oral comments. Our system facilitates teachers to provide both written and oral comments easily which can be stored and played back. Our system can also be used by researchers to analyze teachers' assessment process by studying their feedback.

2. Chinese Composition Assessment System

2.1 System Overview

Our system design is illustrated in Figure 1. The student first composes a Chinese passage as instructed by the teacher. The student then submits the composition to the teacher. The teacher can use our system to open the composition and provide oral and written feedback in this assessment stage. After the teacher finishes the assessment, the marked result is sent back to the student. The student can use our system to playback the teacher's assessment to receive the oral and written feedback in this playback stage.

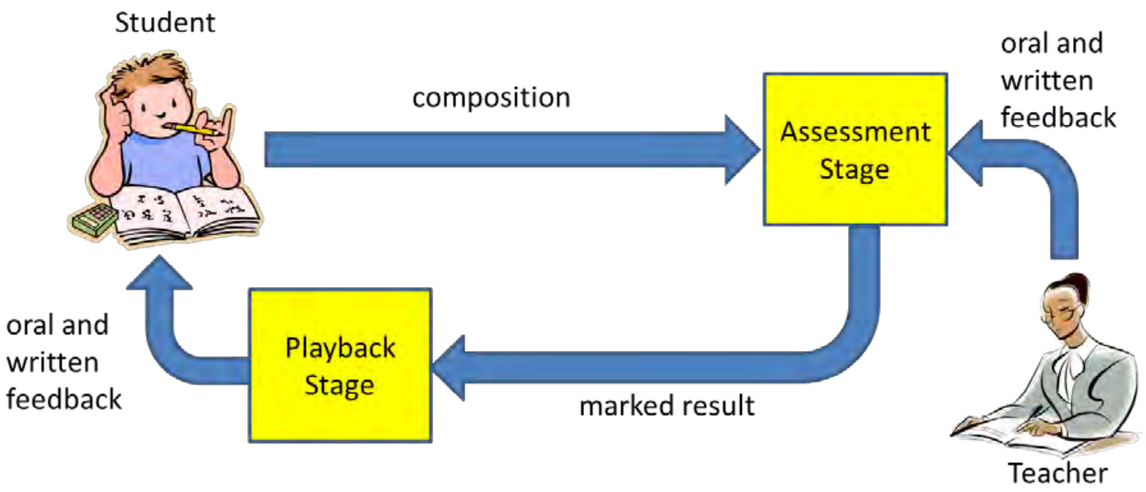


Figure 1. System Design.

The assessment stage and the playback stage will be explained in more details in the following sections.

2.2 Assessment Stage

In the Assessment Stage, the teacher opens the student's composition and reads it using our system. The teacher can markup directly on the student's work to provide written comments. The teacher can also record his/her voice to provide oral comments.

2.2.1 Opening Student's Composition Work

The student forms his/her composition on normal grid paper in the traditional way. The finished composition is scanned and stored in the commonly used JPEG format. Sometimes the composition work may consist of several pages and they will be stored with the same prefix and numbered in sequence. The teacher can run our system on a tablet and open the student's composition. The scanned composition will be displayed using our interface. Besides, since the grid paper may be in landscape or in portrait, our system can also be run in the landscape mode or in the portrait mode in the tablet to accommodate both options, and the composition work will be displayed with the correct aspect ratio, as illustrated in Figure 2 and Figure 3. We also provide a list box in our interface which can show a list of possible teachers' names and a particular teacher can choose his/her name there to identify who performs the assessment in case the students' composition get distributed among different teachers or the same student's composition is marked by more than one teacher. The assessment date is also shown and will be saved together with the teacher's feedback.

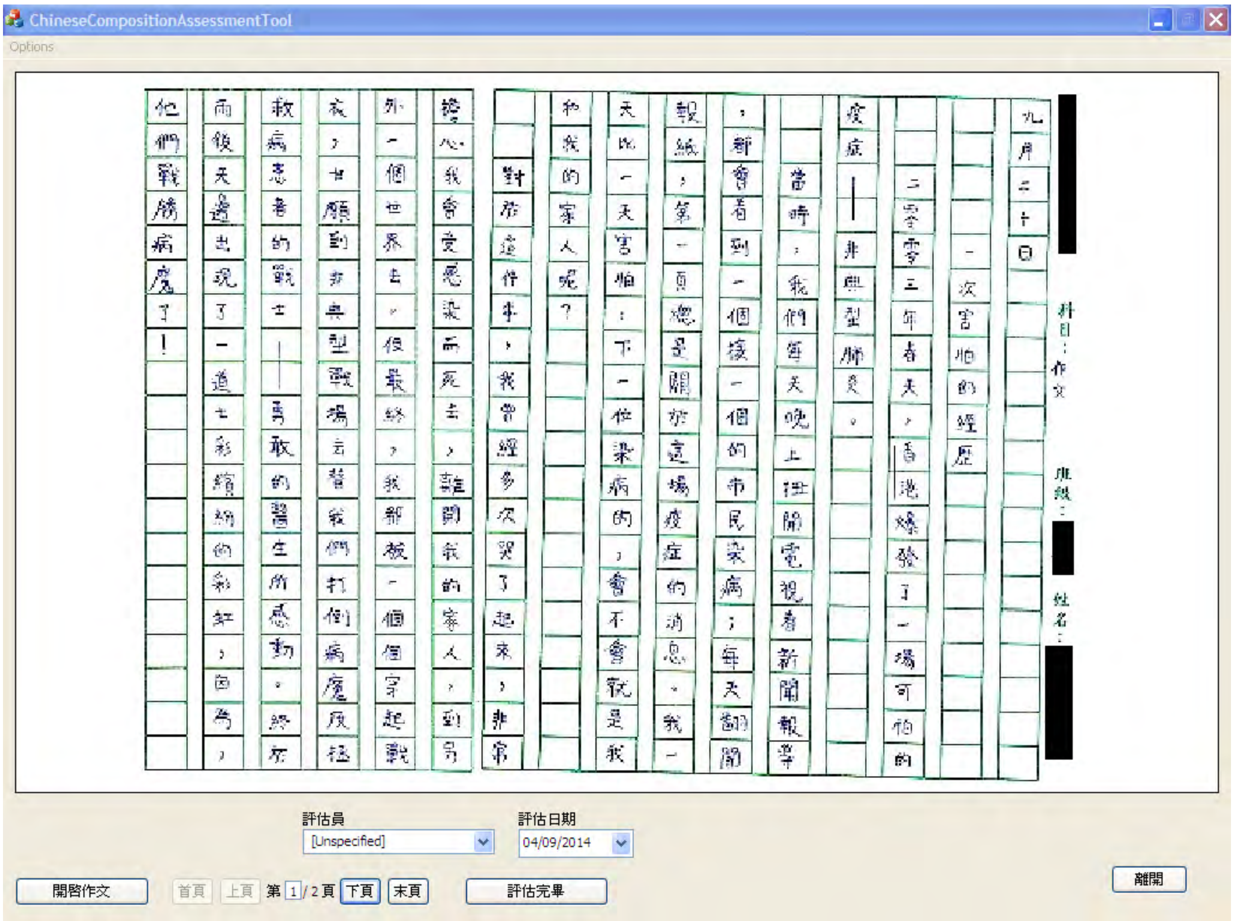


Figure 2. Display of Student's Composition Work in Landscape.

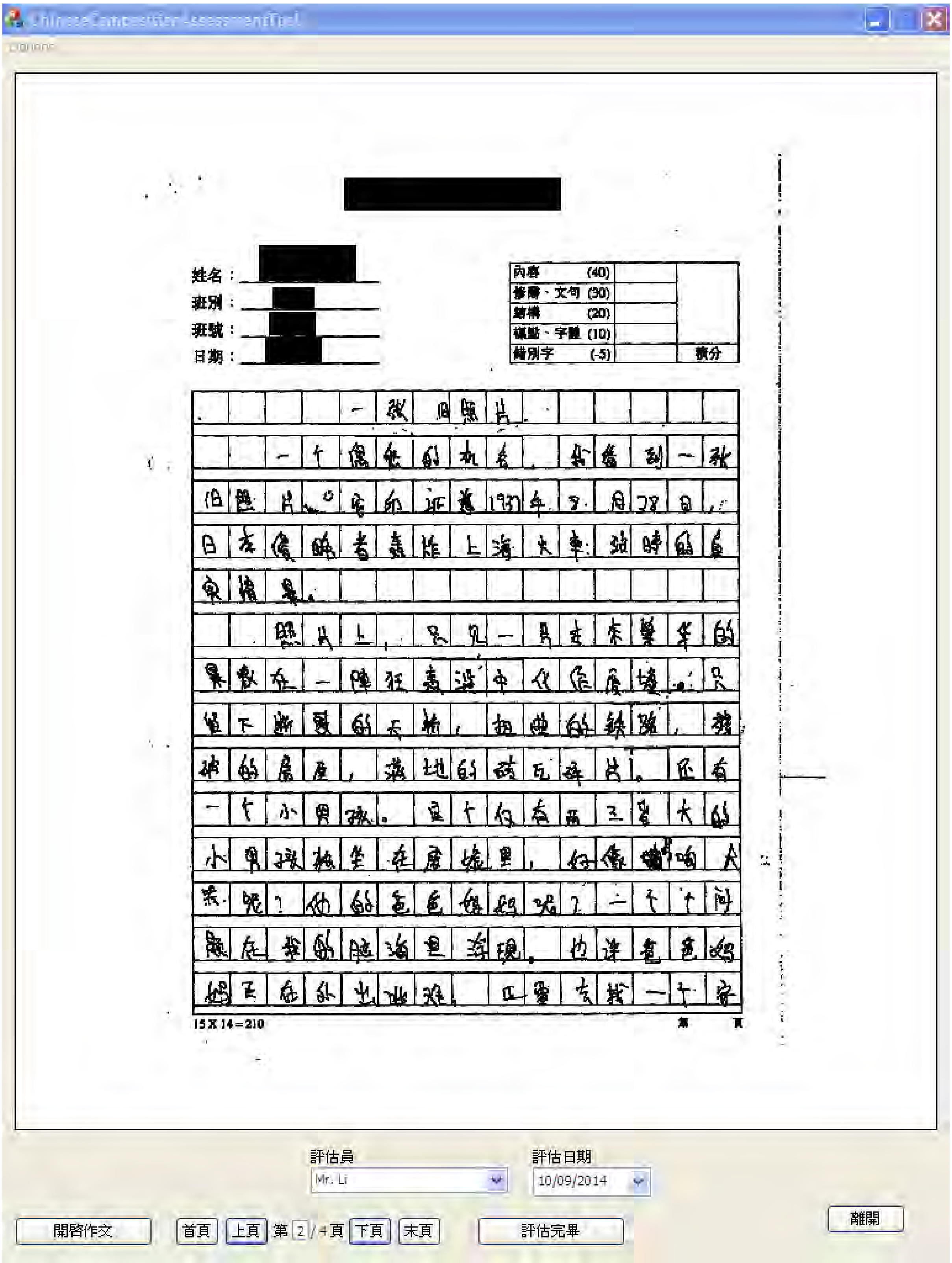


Figure 3. Display of Student's Composition Work in Portrait.

Some people may argue why we do not let the students form their composition directly in the tablet to save the trouble of paper writing and scanning. Our existing approach for scanning the paper-based composition has the benefit that existing work by the students can be scanned and ready to be used by our system for the assessment. Besides, some students may find it easier to compose their work with pen and paper in the traditional way. Nevertheless, we agree that allowing students to provide input directly to form their composition in the tablet can be a good option and we may add this feature in our future work.

2.2.2 Navigating to Different Pages

A student's composition work may span several pages. As mentioned in the previous sub-section, each page will be scanned with the same prefix in the file name which is numbered in sequence. Our system will detect automatically the total number of pages exist in the student's work after the teacher opens it with our system. In our interface, we have provided several buttons to facilitate the easy navigation of the pages. In particular, there are 4 buttons that allow the teacher to navigate to the previous page, the next page, the first page and the last page.

2.2.3 Providing Written Comments

Once the composition is displayed in our interface, the teacher can markup on the tablet using a stylus. The teacher's markup will be shown in red. With this function, the teacher can circle the characters that are written wrongly or underline some sentences whose structure is not well organized. The teacher can also write down explicit comments for the students to improve their work. This function simulates the actual scenario in which a teacher marks a student's composition work on paper with a red pen such that the teacher does not need much extra time to learn to use our system for the assessment.

Figure 4 and Figure 5 illustrate two example pages of student's composition with the teacher's markup made using our system.

2.2.4 Providing Oral Comments

In addition to written comments, our system is able to allow the teacher to record his/her voice during the assessment such that the teacher can provide oral comments. This function can be complement to the written comments. For example, the teacher can circle a character while saying which part of the character is written wrongly. The teacher can underline some sentences and voice out how they can be revised to improve the flow. The teacher can also provide general comments to the student to provide advice about how to improve the composition skill in the future. This feature can also save the teacher's time such that the teacher does not write down too much details in the written comments as the feedback can be included in the oral comments.

This voice recording feature can also be used by researchers to study the way teachers mark students' composition. Researchers can ask the teacher to voice out what they are thinking while they are assessing the composition. Under this "Think Aloud" approach, the teacher's oral comments can be saved and analyzed by the researchers together with the written comments.

2.3 Playback Stage

The assessed composition can be played back by the student. First the student can open the marked result using our system to view the markup by the teacher. The student can navigate to the previous page, next page, first page and last page using similar buttons as introduced in the previous section for the assessment stage. The student can also playback the assessment session such that the student can see when the teacher writes down the comments with the exact timing and listen to the teacher's oral feedback which is synchronized with the written comments. Our interface provides control buttons to let the student to play, pause and stop the playback session.

This feature can also help researchers to playback the teacher's oral comments and the written comments to study the assessment process made by the teacher. As mentioned previously, the researchers may ask teachers to voice out their thoughts in a "Think Aloud" approach which are stored as oral comments. The researchers may analyze different assessment strategies adopted by the teachers and try to devise novel ways for students to improve their composition skills.

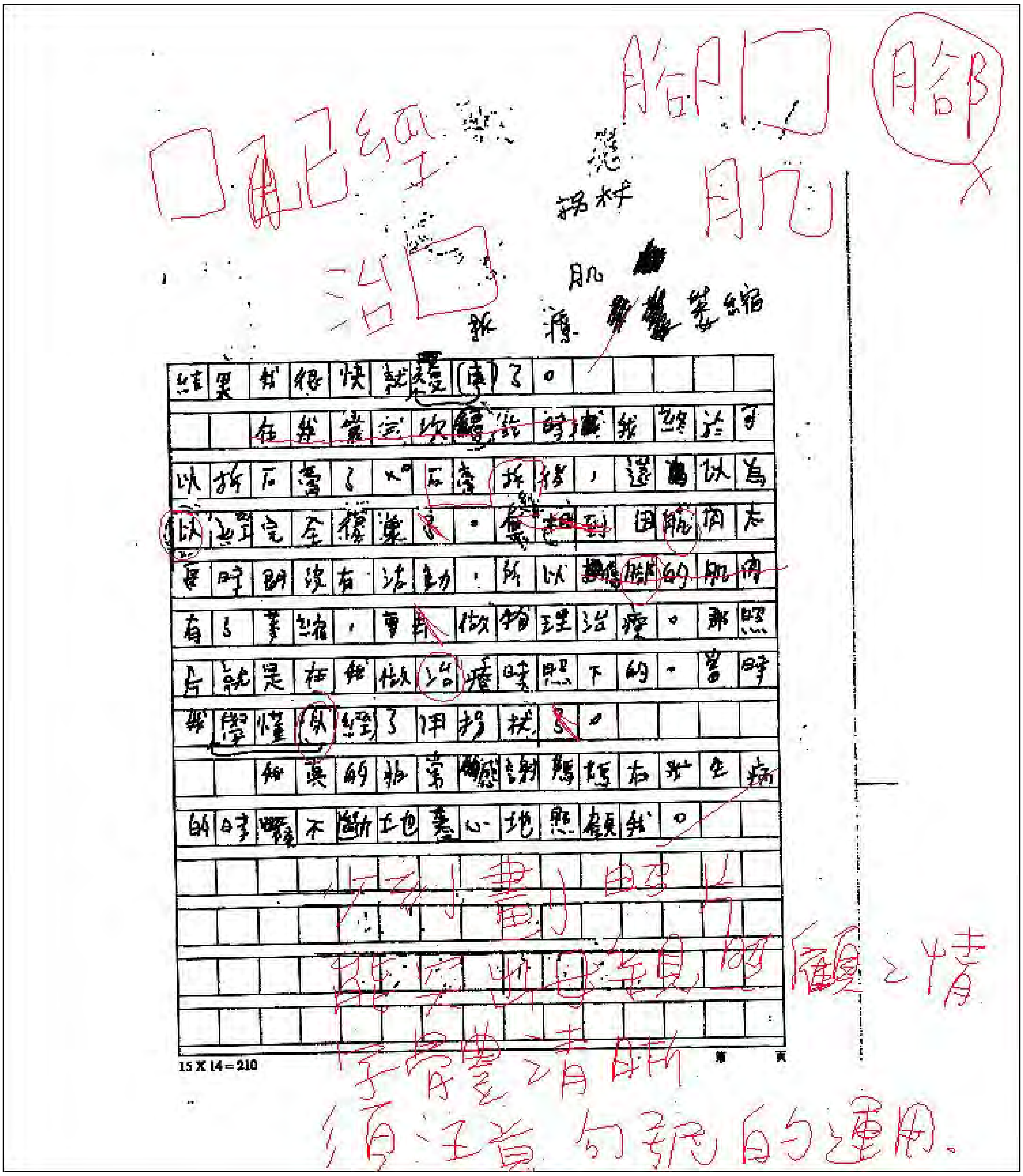


Figure 4. Example 1 of Student's Composition with Teacher's Markup.

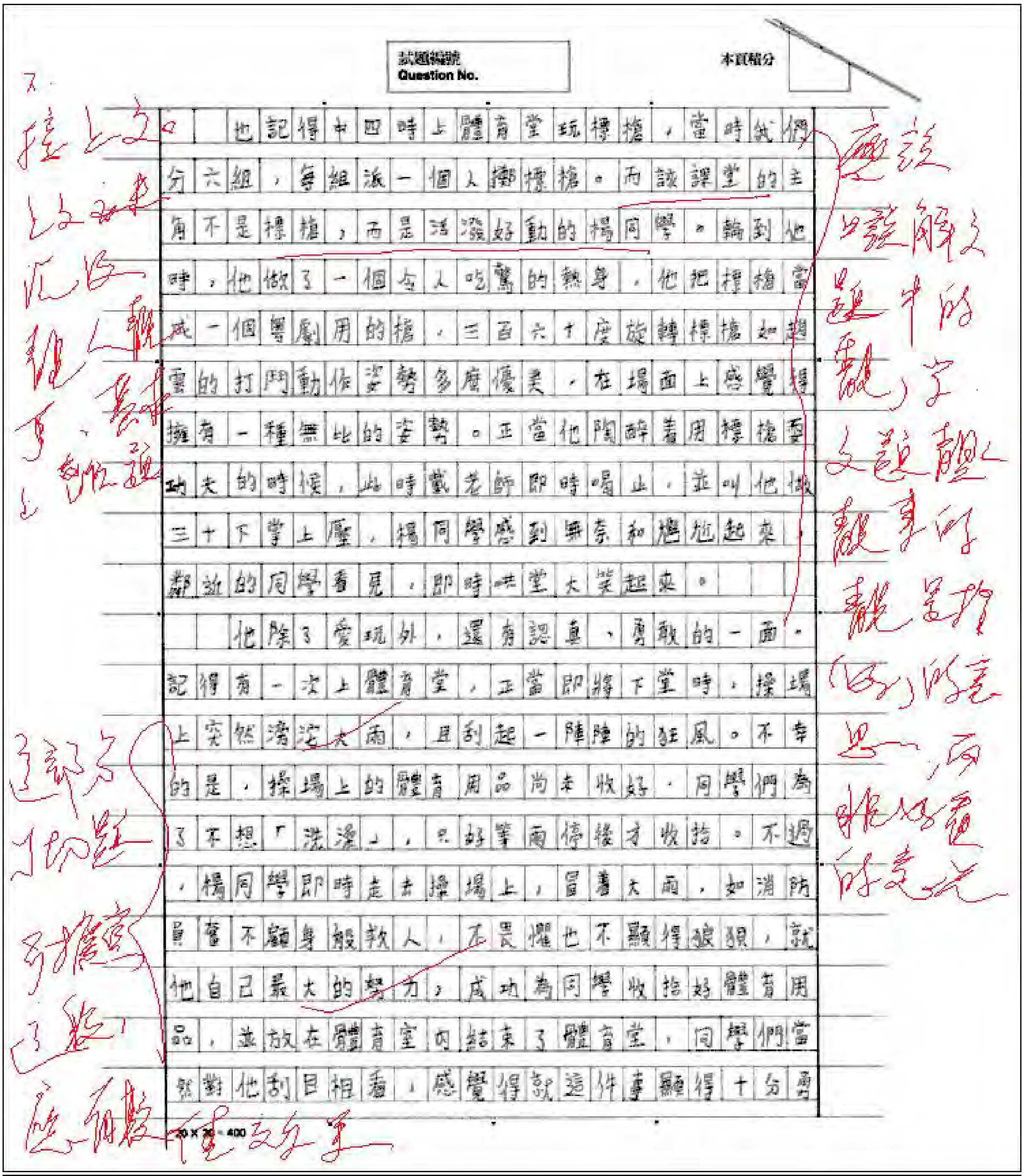

試題編號
Question No.
本頁積分
也記得中四時上體育堂玩標槍，當時我們
分六組，每組派一個人擲標槍。而該課堂的主
角不是標槍，而是活潑好動的楊同學。輪到他
時，他做了一個令人吃驚的熱身，他把標槍當
成一個粵劇用的槍，三百六十度旋轉標槍如趙
雲的打鬥動作姿勢多麼優美，在場面上感覺得
擁有一種無比的姿勢。正當他陶醉着用標槍耍
功夫的時候，此時戴老師即時喝止，並叫他做
三十下掌上壓，楊同學感到無奈和尷尬起來，
鄰近的同學看見，即時哄堂大笑起來。
他除了愛玩外，還有認真、勇敢的一面。
記得有一次上體育堂，正當即將下堂時，操場
上突然滂沱大雨，且刮起一陣陣的狂風。不幸
的是，操場上的體育用品尚未收好。同學們為
了不想「洗澡」，只好等雨停後才收拾。不過
，楊同學即時走去操場上，冒着大雨，如消防
員奮不顧身般救人，不畏懼也不顯得狼狽，就
他自己最大的努力，成功為同學收拾好體育用
品，並放在體育室內結束了體育堂，同學們當
然對他刮目相看，感覺得就這件事顯得十分勇
20 x 20 = 400

Figure 5. Example 2 of Student's Composition with Teacher's Markup.

3. Conclusions and Future Work

Chinese composition is an important aspect for students to learn the Chinese language. In this paper, we propose a tablet-based Chinese composition assessment system which can help teachers to provide written and oral feedback. We present our implementation of the system and introduce the features. The recorded assessment session can be played back for student to receive the feedback. Our proposed technology enhances students' Chinese composition skill as detailed feedback can be made available with the teacher's written comments synchronized with the oral comments. Researchers can also extract different patterns from the teachers' assessment to devise strategies to help students improve their Chinese composition skill.

As future work, there are many potential additional features that can be incorporated into our system. For example, we can implement the input function that allows students to write or type their composition directly into our system such that there is no need to scan paper work. In addition to applying our tablet-based assessment for helping teachers to provide feedback of Chinese composition, we will look into the issues of developing more advanced technologies to support peer assessment such that students can work together on their Chinese composition in a more collaborative manner. As included in the study of Shum (2005), the adoption of peer-assessment by classmates was shown to motivate students to revise their Chinese composition. We can explore some pattern recognition techniques that can be used to categorize the teachers' comments in a more automated fashion and provide some quantitative analysis.

References

Chang, K. E., Lan, Y. J., Chang, C. M. & Sung, Y. T. (2010). A mobile-device-supported strategy learning for Chinese reading comprehension. *Innovations in Education and Teaching International*, 47(1), 69-84.

Ho, H.-F. & Lin, P.-Z. (2010). Chinese character Learning Review System with Item Response Theory Analysis, *Proceedings of Research in Reading Chinese (RRC) Conference 2010*, Canada.

Li, B. J. (2005). Innovations of teaching Chinese composition in schools in mainland China. In M. S. K. Shum & L. D. Zhang (Eds), *Studies in Writing: Teaching Writing in Chinese Speaking Areas* (pp. 131-147). New York, USA: Springer.

Lo, Y. C. (2006). Practice and Challenges of School-based Formative Assessment. *Annual Conference of International Association for Educational Assessment*, Singapore.

Lou, S.-J., Wu, S.-C., Shih, R.-C. & Tseng, K.-H. (2010). Adoption of Blogging by a Chinese Language Composition Class in a Vocational High School in Taiwan. *Australasian Journal of Educational Technology*, 26(6), 898-916.

Shum, M. S. K. (2005). Effects of Four Methods of Evaluation of Chinese Composition in Hong Kong Secondary Schools - Peer Checklist Evaluation at a Closer Look. In M. S. K. Shum & L. D. Zhang (Eds), *Studies in Writing: Teaching Writing in Chinese Speaking Areas* (pp. 199-213). New York, USA: Springer.

Sim, S. H. (2005). Teaching Chinese composition in Singapore secondary school – Past, Present and Future. In M. S. K. Shum & L. D. Zhang (Eds), *Studies in Writing: Teaching Writing in Chinese Speaking Areas* (pp. 245-258). New York, USA: Springer.

Wong, L.-H., Hsu, C.-K., Sun, J., & Boticki, I. (2013). How Flexible Grouping Affects the Collaborative Patterns in a Mobile-Assisted Chinese Character Learning Game? *Educational Technology & Society*, 16 (2), 174–187.

Ying, K. K., Leung, K., Lee, R., & Chow, D. (2014). Chinese Composition Teachers' Commentary Styles and Patterns in a Tablet-Based Marking Environment. *Proceedings of ICHL 2014* (pp. 323-334). Shanghai, China: Springer.

Learning to learn collaboratively on Facebook – A pilot study

Sarah Hsueh-Jui LIU[a] & Yu-Ju LAN[b]*
[a] *Kainan University, Taiwan*
[b] *National Taiwan Normal University, Taiwan*
*yujulan@gmail.com

Abstract: This study aims to examine negotiated interaction amongst English-as-a-foreign-language undergraduates in a real time text-based environment. It looks at the interactional patterns of negotiation on Facebook between dyads of learners and native speakers of English. Overall, all dyads tended to negotiate meaning rather than form, as numerous errors (e.g., morphosyntactic errors) were not repaired during the tasks. The learners were found to employ a number of interactional strategies for communication breakdowns, e.g., clarification request and comprehension check. In addition, the results derived from the interview data reveal that all learners had a positive perception of working with others on Facebook, e.g., increasing motivation and engagement.

Keywords: Language Related Episodes, Computer-Mediated Communication, Synchronous Text-Based CMC

Introduction

Over the past two decades, the use of online chat or instant messaging in synchronous computer-mediated communication (CMC) has become a popular means to enhance second language learning (Sauro, 2012). Due to its capacity to store data files, CMC has potentially provided opportunities for researchers to examine interaction (Peterson, 2009). In light of this, much of the existing research, such as Lee (2001), Peterson (2009), and Sotillo (2005), has explored negotiated interaction in CMC environments.

The underpinning sociocultural theory highlights that social interaction between conversational interlocutors helps scaffolding to take place in cognitive development (Vygotsky, 1978). In second or foreign language (L2/FL) learning, learners interact with more competent individuals to negotiate new words or expressions in the language during a conversation, and the negotiated interaction may be able to provide the learners with linguistic input (Long, 1996). For instance, in collaborative negotiation, a series of ongoing interactional processes take place in order for learners to comprehend the unknown items or to convey intended messages, e.g., asking assistance from their interlocutor and clarifying the meaning of a word. In this way, the learners will be able produce their output relevant to the unknown linguistic items.

Results of negotiation research reveal that unknown words are more likely to be a trigger of negotiated routines than any other forms of linguistic aspects, such as grammar. Fernández-García and Martínez-Arbelaiz (2002) found that when their participants encountered unfamiliar lexical words, they tended to engage in negotiation for meaning. Such results are in line with Tudini's (2003) study, which revealed that lexical and structural difficulties were found to trigger modified negotiation. In terms of communication strategy use, Jepson (2005) claimed that his participants employed a greater number of clarification requests in both text and voice chat. Nevertheless, such results in terms of types of strategy use differ from study to study, as Lee (2001) reported that her participants employed clarification checks, requests and self-corrections more frequently than other strategies, e.g., word invention.

Another strand of research has investigated attitudes or perceptions of learning in CMC, whether learners perceive CMC learning positively, or whether use of CMC can foster learners' engagement or participation and motivation in the target language. For instance, Peterson (2010) revealed that learners working collaboratively with others in a virtual world demonstrated high levels of motivation and interest and low levels of stress when communicating with others in the virtual world.

This study builds upon previous research in traditional face-to-face (F2F) and CMC settings regarding how second or foreign language (L2/FL) learners employ communication strategies to overcome what they lack linguistically in the target language when conversing with their conversational interlocutors, e.g., Nakahama, Tyler, and van Lier (2001). Specifically, the current study aims to explore the synchronous text-based interaction taking place on Facebook between the learners and the native speakers of English and the learners' perceptions of collaborative learning in the CMC environment. There are two research questions: What are the interactional strategies used by the participants during the task in the real time text-based CMC? What is the perception of collaborative learning in the CMC environment?

Method

Participants

This pilot study involved ten participants, that is, five native speakers of English and five EFL learners, whose ages ranged between 21 and 25 years old and whose first language was Taiwanese or Mandarin Chinese. They had been learning English for more than ten years. The sample of the learners included three female and two male students who were sophomores and had majored in Tourism. Each dyad consisted of a native speaker and a learner, resulting in five dyads in total.

Instruments

Tasks

Three types of communication tasks were made available on Facebook, namely, the information gap and the decision-making activities, and a reading article; the first two tasks were designed based on the task-based approach (Willis & Willis, 2007), and the last one, drawn from the learners' textbook, was modified in this study. For example, the reading task was an attempt to generate negotiated interaction, as there were several words or expressions that were new to the learners.

Retrospective interviews

The interviews were held a few days after the paired participants had completed the given task in terms of the perception of collaborative learning in the text-based environment, such as, 'Can you describe the learning experience with your partner on Facebook?' and 'What is the most challenging part when learning collaboratively on Facebook?' The learners were interviewed individually in the chatroom on Facebook.

Data analysis

Existing taxonomies of interactional strategy use were adopted to analyze the text-based CMC discourse of the participants, such as Bower and Kawaguchi (2010), Lyster (1998), and Varonis and Gass (1985). According to the modified interaction, the interactional strategies may include confirmation and comprehension checks, clarification requests, and asking for assistance. One excerpt taken from Varonis and Gass (1985:78) is given below to illustrate that learners check each other's comprehension, in this instance, when the learners encountered the unfamiliar lexical item 'ingless.'

a. 140J: I was born in Nagasaki.
→ Do you know Nagasaki?
b. 120S: I'm from Venezuela.
UL J: Venezuela
→120S: Do you know?
c. 140S: declares her ingress
140J: Ingless
140S: Yes, if for example, if you. When you work you had an ingress.
→ you know?
d. 140S: and your family have some ingress
140J: yes ah, OK OK
→140S: more or less OK?

In the above discourse, each arrow indicates a trigger showing that the hearer does not understand part of an utterance, e.g., 'Do you know Nagasaki?' Following the trigger, Lines a. – d. indicate a series of negotiated episodes, where Line a. represents a trigger, Line b. was the indicator, Line c. was the response, and Line d. was the reaction to the response.

Procedure of data collection

All the participants took part voluntarily in the real time text-based communication. Afterwards, they were asked individually to work with their interlocutors on Facebook at the same time by using text chat in the chatroom. Each pair of participants notified the first researcher of this study regarding the time of the online discussion, and then the researcher invited each pair into the same chatroom to undertake the tasks. Each pair completed one set of tasks each time, which took them one hour approximately. This resulted in 15 hours of data, that is, 15 dyads of NS-NNS, and it took around six weeks to complete the tasks. One week after the five pairs of participants had completed the three types of tasks, the online retrospective interviews were held on Facebook, where individual learners were interviewed by using a text chat. Each interview lasted nearly 30 minutes, giving 150 minutes of interviews in total.

Results

This section discusses how the results of the pilot study answer the two research questions. Prior to this, the negotiated interaction that was found in the traditional face-to-face environments took place in the text-based CMC environment, such as this current study and some other studies (e.g., Fernández-García and Martínez-Arbelaiz, 2002). First, errors that occurred in the learners' output were categorized as grammatical errors, errors of lexis, or misspellings. It was found that the learners made the most errors in grammar (76%), which was followed by lexical items (29%), and the fewest errors were those of misspelling (5%). Examples of the grammatical errors included morphosyntactic errors (e.g., subject-verb agreement) and prepositions.

In the first research question, the negotiation discourse during the tasks tended to focus on meaning rather than on form. Hence, many grammatical errors were not corrected by the learners' interlocutors. During negotiation for meaning, a number of interactional strategies were identified, including confirmation and comprehension checks, clarification requests, asking for assistance, and others. In Table 1, the learners employed the most frequent strategy of asking for assistance (35.8%), whilst the strategy least frequently used was the confirmation check (9.4%). It should be noted that apart from one dyad of NS-NNS (Barry-Mei), the rest of the dyads tended to negotiate meaning rather than form.

Table 1: Summary of Interactional Strategy Use by the dyads

	Counts	%
Confirmation check	5	9.4
Comprehension check	10	18.9
Clarification request	12	22.6
Asking for assistance	19	35.8
Others	7	13.2
Total	53	100

Noticeably, negotiated interaction took place when the learners came across new words, whereupon they asked for help from their interlocutors to provide the meaning of the unknown words. In other words, negotiation was less likely to occur when learners made grammatical errors. In the current study, most learners were likely to ask their partners the meaning of new words in a reading task, e.g., 'What does X mean?'

In the second research question, most learners (N=4) had a positive perception of working with others on Facebook, as they believed that they gained linguistic knowledge from their superior interlocutors. Three learners commented that the interaction in the environment enhanced their motivation and engagement in meaningful communication. Specifically, two main aspects emerged from the interview data in relation to the perception of learning in the synchronous text-based environment, namely, cognitive support and affective scaffolding. Regarding cognitive support, the results reveal that four out of the five NNS participants agreed that interaction with native speakers could help them improve their linguistic knowledge in English in terms of vocabulary and grammar. For example, one female learner, Mei, recounted that when she had grammatical errors in her answers, her partner, Barry, corrected them by comparing the differences in sentence structure between English and Chinese: '*Barry is a native speaker of English speaking some Chinese… then he said the structure in English is more rigid than that in Chinese…*' Instances such as this were found in the interaction between Mei and Barry in the log files, where he both implicitly and explicitly explained grammar points to Mei, e.g., '*… lets [let's] keep the sentence in the present tense. Ill [I'll] give you a hint. Its [It's] in the ending.*'

Regarding affective scaffolding, when compared to their previous experience in face-to-face communication, all the learners stated that they felt much less pressure having a chat with native speakers in the text communication during the tasks. As another learner, Joan, stated '*All our conversation is done on Facebook; that makes me feel less embarrassed if I make some grammar mistakes… but if I talk to a native speaker in person, when making stupid mistakes, I will feel nervous…*'

Discussion and Conclusion

This pilot study was intended to examine the negotiated interaction taking place in a real time text-based environment by integrating the overarching socio-cultural theory applied in the L2/FL settings into the synchronous CMC. The preliminary results suggest that the learners made more grammatical errors than misspellings or errors of lexis. There were five types of interactional strategies used by the pairs of NS-NNS in this study: confirmation and comprehension checks, clarification requests, asking for assistance, and others (e.g., use of L1). The learners showed a greater preference for using the strategy of asking for assistance rather than any other strategies, such as clarification requests. It is important to note that there were various types of errors in the learners' messages, but these were not corrected by their superior interlocutors though one pair of participants tended to contribute to form negotiation while the others were more likely to initiate negotiation for meaning. Such results discovered in this study corroborate those in previous research. Finally, the results of the retrospective interviews suggest two essential components in terms of cognitive support and affective scaffolding in relation to the perception of collaborative learning in the text-based environment, such as corrective feedback on grammatical errors provided by an interlocutor.

This study, like any other study, has its inherent limitations. One limitation is that the small size of the sample being adopted in this study cannot provide generalizability to the wider context. Another limitation of this study, which needs to be addressed in relation to the categories of errors occurring in the learners' output, can also include other types of errors, such as word order. Likewise, the interactional strategies can be divided into further different categories, such as word coinage. In addition, one aspect of language use in the text-based communication may focus on the corrective feedback or recast provided by interlocutors and the responses of learners to the corrective feedback. However, those aspects require further investigations.

Acknowledgements

The authors would like to thank the Ministry of Science and Technology, Taiwan, R.O.C., for financially supporting this research under Grant Nos. NSC 101-2511-S-003-031-MY3, MOST 103-2628-S-003-002-MY3, and MOST 103-2511-S-424-004-. They are also thankful that this research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C., the "International Research-Intensive Center of Excellence Program" of NTNU and the Ministry of Science and Technology, Taiwan, R.O.C. under Grant No. NSC 103-2911-I-003-301.

References

Bower, J., & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTANDEM. *Language Learning & Technology, 15*(1), 41-71.

Fernández-García, M., & Martínez-Arbelaiz, A. (2002). Negotiation of meaning in non-native speaker-nonnative speaker synchronous discussions. *CALICO Journal, 19*(2), 279-294.

Jepson, K. (2005). Conversation – and negotiated interaction – in text and voice chat rooms. *Language Learning & Technology, 9*(3), 79-98.
Lee, L. (2001). Online interaction: Negotiation of meaning and strategies used among learners of Spanish. *ReCALL, 13*, 232-244.
Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition*, pp. 413-468. New York: Academic Press.
Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning, 48*(2), 183-218.
Nakahama, Y., Tyler, A., & van Lier, L. (2001). Negotiation of meaning in conversational and information gap activities: A comparative discourse analysis. *TESOL Quarterly, 3*, 377-405.
Peterson, M. (2009). Learner interaction in synchronous CMC: A sociocultural perspective. *Computer Assisted Language Learning*, *22*, 303-321.
Peterson, M. (2010). Learner participation patterns and strategy use in Second Life: An exploratory case study. *ReCALL, 22*, 273-292.
Sauro, S. (2012). L2 performance in text-chat and spoken discourse. *System, 40*, 335-348.
Sotillo, S. (2005). Corrective feedback via Instant Messenger learning activities in NS-NNS and NNS-NNS dyads. *CALICO Journal, 22*(3), 467-496.
Tudini, V. (2003). Using native speakers in chat. *Language Learning & Technology, 7*(3), 141-159.
Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6*, 71-90.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.

Exploring the Effectiveness of a Flipped Classroom Based on Control-Value Theory: A Case Study

Jiu-Tong LUO, Meng SUN, Bian WU & Xiao-Qing GU*
Department of Educational Information Technology, East China Normal University, China
*xqgu@ses.ecnu.edu.cn

Abstract: Flipped classroom is a newly emerged video-lecture-supported teaching approach that aims to improve learning outcomes and teaching effects though students' self-regulated learning after class, teachers' assistance and interaction during class. In this paper, we explored the effectiveness of a flipped-classroom Java Programming course based on control-value theory in a university of China. The specific goal of this case study was to explore students' achievement emotion in this flipped classroom setting and explore the correlation among the given factors. Adopting a mixed method, this study collected data on students' Java self-efficacy (control), motivation (value), achievement emotion, self-regulated learning ability learning outcomes as well as open questions about attitudes and suggestions to flipped classroom. And students' learning outcome was multi-dimension evaluated, including self-reports after each project, assignments, online activities participate and final exam. The findings showed that the control-value theory could explain the effectiveness of the flipped classroom well. In addition, implications were also concluded from this study.

Keywords: Flipped classroom, control-value theory, effectiveness, case study

1. Introduction

Over the past decades, the public were dissatisfied with teaching effects of traditional approaches. With the development of newly emerged information technology, video-lecture can help students learn by themselves. Flipped classroom is based on a student that arrives to class ready for the learning experience and prepared by watching the video-lectures provided by the teacher in advance (Bristol, 2014). Nowadays, more and more researchers are focusing on the different parts of flipped classroom to transform the traditional education system together with MOOC and micro-video. This study mainly explored the effectiveness of flipped classroom based on control-value theory put forward by Pekrun (2006) focusing on exploring students' achievement emotion and its antecedents and effects.

2. Background

2.1 Flipped classroom

The conception of flipped classroom can be traced back to 2008, the chemistry teachers at Colorado's Pike's Peak, veteran Woodland Park High School came to the idea to provide video records of their lessons online for those absent students to see what they missed (Tucker, 2012). And it soon became very famous for its concept of having students to learn self-paced. Bishop and Verleger (2013) provided a comprehensive survey study of flipped classroom and concluded that a) most researches aimed to explore students' perceptions and use single-group designs; b) students' attitude generally positive overall, although they tent to in-person lessons to video lectures and c) little work investigating students learning outcomes objectively.

2.2 Control-Value theory

Control-Value theory was first put forward by Pekrun. It focused on achievement emotion and aimed to analyze the antecedents and effects of emotion experience in achievement and academic setting (Pekrun, 2006). Various factors were concluded in this comprehensive framework, such as expectancies, attributions, intrinsic/extrinsic value, achievement emotions, outcome emotions, self-regulation of learning and so on (Pekrun et al., 2007). By using control-value theory, we can improve the understanding of students' motivation, learning, performance (Artino Jr, 2012) and engagement (Buff, 2014).

The control dimension for the students learning a course factor often refers to their expectancies, attributions and confidence, such as self-efficacy (Pekrun et al., 2007). Self-efficacy was defined as people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performance (Bandura, 1986; Askar and Davenport, 2009).

The value dimension of the course is both instinct and extrinsic (Pekrun et al., 2008). Motivation can explain the students' judgments of the course value for their own. ARCS model was a comprehensive motivation model includes attention, relevance, confidence and satisfaction for teachers develop qualified material for students in online or blended learning environment (Keller, 2010). And another factor can influence the students perceiving the value of courses is their technology acceptance. TAM is a framework measuring users' perceived and intention of using technology (Davis et al., 1989).

Achievement emotions are also intimately involved in virtually every aspect of the teaching and learning process and, therefore, an understanding of the nature of emotions within the school context is essential (Schutz and Lanehart, 2002). In the control-value theory, achievement emotion is defined as emotion tied directly to achievement activities or achievement outcomes, includes enjoyment, joy, relaxation, anger, frustration, boredom, shame, hope, pride, relief, anxiety, sadness, disappointment, hopeless and so on (Peruk, 2007).

Self-regulated learning is an important aspect of students' academic success (Effeney et al., 2013; Zimmerman, 2002). According to Zimmerman (2002), self-regulated learning is actions directed at acquiring information or skill that involve agency, purpose, and instrumentality self-perceptions by a learner. And it was cited as the effects of achievement emotion (Peruk et al., 2007).

2.3 The research framework

According to the control-value theory, we put forward a framework for this study showed as following (see Figure 1). The aim of this study was mainly focused on exploring achievement emotion and testing the framework in a flipped classroom as a pre study for the further exploration of effectiveness of this newly emerged teaching approach in the coming semester. As mentioned above, the control and value dimension had several factors, and this study only considered students Java self-efficacy as the control factor, while course motivation as the value factor.

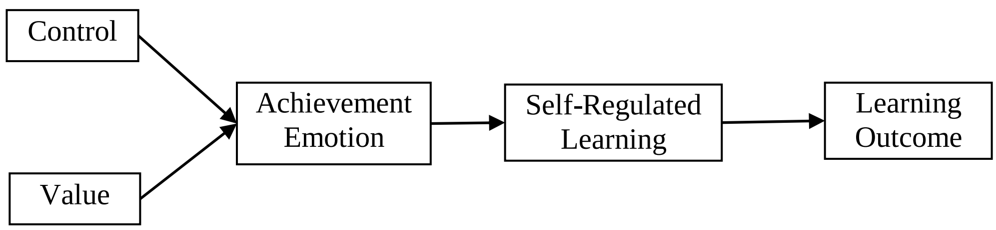


Figure 1. The Research Framework Based on Control-Value Theory.

3. Methodology

3.1 The study

The research reported here was taken in Shanghai, China that aimed to facilitate students' learning and enhance teachers' teaching by changing their roles during and after classes. There were totally 21 students (6 Male, 15 Female) including 9 sophomore and 12 junior students majoring Educational Technology in a normal university in Shanghai participated in this study.

3.2 *The pedagogy*

This research was taken in a flipped classroom setting. The students were asked to learn Java programming course by self-regulated learning out of class. Students were divided into several groups to fulfill three projects assigned to them, with each group containing 2-3 person. During this course, each student was asked to finish 3 projects with their teammates. 4 weeks were given to finish the first project which contains 3 sub-tasks, and 5 weeks for the next with 4 sub-tasks and 9 weeks for the last one with 8 sub-tasks.

The teacher provided learning materials in advance on the learning platform (Sakai). The videos were limited up to 8 minutes, known as micro-video podcast. It was reported that this kind of micro-video could contribute to the students' learning interests and attention last longer in self-regulated learning environment. Besides, several other types of materials, such as pdf, doc as well as hyperlinks of Java API etc. were also provided for the students to facilitate their self-regulated learning. Students were encouraged to participate online activities such as posting, and discussing with each other. Students also needed to submit a self-report of performance after they finished each project.

The learning outcome of the participated students were multi-dimension evaluated, including self-reports after each project, assignments, online activities participation as well as final exam. The calculation of final learning outcome followed the given formula below:

Total Learning outcome=10% * 3 Projects + 10% (online activities and self-report) + 60% final exam score.

3.3 *The methods*

Adopting a mixed method, this research used instruments which contained 12 items measuring students' Java self-efficacy adapted from Askar and Davenport (2009), 16 items measuring motivation adapted from TAM (Davis et al., 1989; Saadé and Bahli, 2005) and IMMS based on ACRS motivation model (Keller, 2010), 21 items measuring the achievement emotions adapted from the Achievement Emotions Questionnaire (AEQ) (Pekrun et al., 2011), 19 items measuring self-regulated learning ability adapted from Barnard et al. (2009) as well as open questions about attitudes and suggestions to flipped classroom. Collected data including the factors list above in this flipped classroom setting presented on a seven-point Likert scale, from strongly agree to strongly disagree, as well as demographic variables. All of the items in these questionnaires were adopted from existing scales in English, and then translated to Chinese for the participants, so the reliability of the questionnaires were guaranteed. Both quantitative and qualitative methods were adopted in this study. Descriptive and inferential statistics were used to analyze the structured questionnaires., while content analysis was used to coding the open questions.

4. Results

Only 19 students answered the Java self-efficacy and motivation scales effectively; 20 of them answered the achievement emotion and self-regulated learning scales effectively and 18 students answered all of the scales. The reliability of the questionnaires on students Java self-efficacy, motivation, achievement emotion, and self-regulated learning were good, Cornbash's α were 0.93, 0.93, 0.95, 0.74 respectively.

4.1 *The result of self-efficacy*

In this research, the Java self-efficacy (control dimension) scale was divided into three levels from basic understanding of Java Programming to confidently finishing a comprehensive Java project gradually, and each level contains 4 questions. The Min, Max, Mean and SD of each item were listed in the table 1 below. The total cornbash's α was 0.93, and 0.75, 0.70, 0.91 for level 1, 2, and 3 respectively.

Table 1: The result of self-efficacy scales (N=19).

Construct	Item	Min	Max	Mean	SD
Level 1	L1-01	3	6	4.21	0.85
	L1-02	3	7	4.63	1.16
	L1-03	3	6	4.74	0.81
	L1-04	3	6	4.84	0.96
Level 2	L2-01	2	6	4.11	1.15
	L2-02	2	5	4.11	1.05
	L2-03	3	6	4.21	0.79
	L2-04	2	6	4.37	1.12
Level 3	L3-01	2	5	3.74	0.93
	L3-02	3	7	5.11	1.10
	L3-03	1	6	4.37	1.16
	L3-04	2	7	4.32	1.20

4.2 The result of motivation

The course motivation scale (value dimension) contained five parts, such as: perceived ease of use, perceived of useful, intention to use, attention and relevance. The Min, Max, Mean, and SD of each item were listed in the table 2 below. The total cornbash's α was 0.93, and 0.87, 0.83, 0.74, 0.55, 0.84 for the five different parts mentioned above respectively.

Table 2: The result of motivation scales (N=19).

Construct	Item	Min	Max	Mean	SD	Construct	Item	Min	Max	Mean	SD
Perceived Ease of Use	PEU1	2	5	3.74	1.10	Attention	ATT1	3	7	4.84	1.01
	PEU2	2	6	4.58	1.35		ATT2	4	7	5.74	0.94
	PEU3	1	6	4.53	1.43		ATT3	2	7	4.63	1.34
Perceived of Useful	PU1	2	6	4.32	1.29		ATT4*	3	7	4.68	1.34
	PU2	2	6	4.32	1.11	Relevance	REL1	1	6	4.53	1.39
	PU3	2	6	4.16	1.17		REL2	3	7	5.21	0.92
Intention to Use	IU1	2	7	5.00	1.41		REL3	3	7	5.47	1.02
	IU2	2	7	4.68	1.34		REL4	3	7	4.95	1.08

*The items were assessed conversely.

4.3 The result of achievement emotion

The total cornbash's α of achievement emotion was 0.95, and 0.87, 0.34, 0.71, 0.84, 0.96, 0.85, 0.68 for enjoyment, hope, pride, hopeless, anxiety, boredom, and angry respectively. There were no cornbash's α of shame and relief parts, because there was only one item for both of this two parts. The Min, Max, Mean, and SD of each item were listed in the table 3 below.

Table 3: The result of achievement emotion scales.

Construct	Item	Min	Max	Mean	SD	Construct	Item	Min	Max	Mean	SD
Enjoyment	EM1	3	7	5.40	1.31	Anxiety*	ANX1	2	7	5.05	1.54
	EM2	2	7	4.75	1.16		ANX2	2	7	5.20	1.40
	EM3	3	7	5.20	1.15	Shame*	SM1	2	4	2.85	0.81
	EM4	4	7	5.50	1.05	Relief	RF1	1	6	4.00	1.41
Hope	HP1	1	7	4.05	1.40	Boredom*	BD1	3	7	5.70	1.03
	HP2	3	7	4.55	1.16		BD2	3	7	5.50	1.28
Pride	PD1	3	7	5.40	1.23		BD3	3	7	5.60	1.19
	PD2	3	7	5.05	1.05	Angry*	AG1	3	7	5.70	1.38
	PD3	3	6	4.55	1.15		AG2	3	7	5.20	1.47

Hopeless*	HPL1	3	7	5.70	1.30		AG3	4	7	5.50	1.05
	HPL2	2	7	5.10	1.48	N=20, Cronbach's Alpha=0.95					

*The items were assessed conversely.

4.4 The result of self-regulated learning

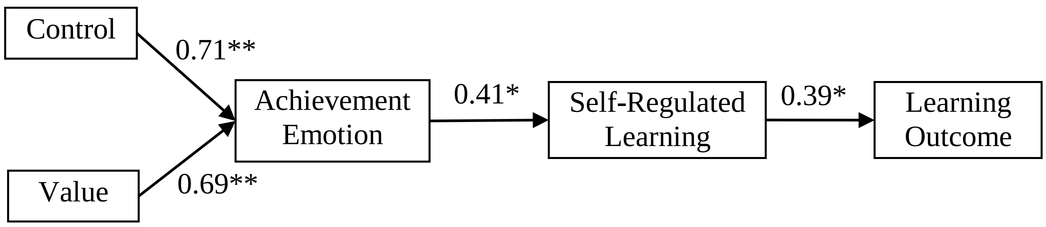
The total cornbash's α of achievement emption was 0.74, and 0.86, 0.73, 0.51, 0.58, 0.34, 0.63 for goal setting, environment structure, help seeking, time manage, task strategy, and self-evaluation respectively. The Min, Max, Mean, and SD of each item were listed in the table 4 below.

Table 4: The result of self-regulated learning scales.

Construct	Item	Min	Max	Mean	SD	Construct	Item	Min	Max	Mean	SD
Goal Setting	GS1	2	6	4.10	1.41	Time Manage	TM1	3	7	4.95	1.19
	GS2	1	6	3.55	1.40		TM2	2	5	3.60	1.05
	GS3	2	7	3.90	1.41		TM3	2	6	3.40	1.23
	GS4	2	6	4.15	1.23	Task Strategy	TS1	1	6	3.50	1.43
Environment structure	ES1	2	7	4.95	1.15		TS2	3	6	4.45	0.76
	ES2	4	7	5.15	0.99		TS3	2	6	4.30	1.13
	ES3	1	7	4.75	1.45	Self-Evaluation	SE1	2	6	3.85	0.99
	ES4	3	7	5.20	1.24		SE2	1	5	3.50	1.10
Help Seeking	HS1	2	7	5.25	1.29		SE3	1	5	4.20	1.11
	HS2	4	7	5.40	0.68	N=20, Cronbach's Alpha=0.74					

4.5 The result based on control-value theory

The correlation between students' Java self-efficacy (control) with their achievement emotion in flipped classroom setting was 0.77 (N=19, p<0.01, one-tailed); the correlation between motivation and achievement was 0.69 (N=19, p<0.01, one-tailed); the correlation between achievement emotion and self-regulated learning was 0.41 (N=20, p<0.05, one-tailed); the correlation between self-regulated learning with students' multi-dimension evaluated learning outcome is 0.39 (N=18, p<0.05, one-tailed) (see Figure 2).



** Correlation is significant at the 0.01 level;
* Correlation is significant at the 0.05 level.

Figure 2. The Results Based on Control-Value Theory.

5. Discussion and implications

As can be found in the data above, the students' achievement emotion in this study tend to be positive, for the Mean of each item was above the average (except 1 item). The results of other scales were also good and the correlations based on the framework were significant. When the students were asked whether they satisfied with their performance in this course or not, most of them still answered not too bad or general. It implicated that students could do even better and have potential of improvement in the flipped classroom. This made us confident to continue using this strategy in the coming semester, and develop more scaffoldings to support their learning.

Although the total reliability of the scales were good, but the reliability of some sub-constructs were not as good as the total because of the limited participates in this study. Actually, there were over

50 students participated in this course at the beginning, but most of them dropped out for the course pressure, limited time et al. This also showed that the students still had some resistance to this innovative teaching approach.

6. Conclusion

This study explored the effectiveness of a flipped classroom based on control-value theory. The study found that students' achievement emotion could be well explained using the framework of control-value theory and each the correlations between them were also significant.

The limitations of this study are listed as following, a) the sample size is small (only 21 participants); b) the alpha values of some scales were low, even though they were adopted from previous researches. Despite the limitations, this study severed as the pre study of the ongoing study of flipped classroom, and concluded many useful implications to the study will be take in the next semester to get a better knowledge of students' development based on control-value theory by using this innovative teaching strategy.

References

Artino Jr, A. R., Holmboe, E. S., & Durning, S. J. (2012). Control-value theory: Using achievement emotions to improve understanding of motivation, learning, and performance in medical education: AMEE Guide No. 64. *Medical teacher, 34*(3), e148-e160.

Askar, P., & Davenport, D. (2009). An Investigation of Factors Related to Self-Efficacy for Java Programming among Engineering Students. *The Turkish Online Journal of Educational Technology, 8*(1), 26-32.

Bandura, A. (1986). Social Foundations of Thought and Action. Prentice Hall, Englewood Cliffs, NJ.

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education, 12*(1), 1-6.

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. *In ASEE National Conference Proceedings*, Atlanta, GA.

Bristol, T. J. (2014). Flipping the classroom. *Teaching and Learning in Nursing, 9*(1), 43-46.

Buff, A. (2014). Enjoyment of learning and its personal antecedents: Testing the change–change assumption of the control-value theory of achievement emotions. *Learning and Individual Differences, 31*, 21-29.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science, 35*(8), 982-1003.

Effeney, G., Carroll, A., & Bahr, N. (2013). Self-Regulated Learning: Key strategies and their sources in a sample of adolescent males. *Australian Journal of Educational & Developmental Psychology, 13*, 58-74.

Keller, J. M. (2010). Motivational design for learning and performance. New York, NY: Springer.

Kong, S. C. (2014). Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: An experience of practicing flipped classroom strategy. *Computers & Education, 78*(5), 160-173.

Pekrun, R. (2000). A social cognitive, control-value theory of achievement emotions. In J. Heckhausen (Eds.), *Motivational psychology of human development*. Oxford, UK: Elsevier Science,

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341.

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. Bibliothek der Universität Konstanz.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36-48.

Saadé, R., & Bahli, B. (2005). The impact of cognitive absorption on perceived usefulness and perceived ease of use in on-line learning: an extension of the technology acceptance model. *Information & Management, 42*(2), 317-327.

Schutz, P. A., & Lanehart, S. L. (2002). Introduction: Emotions in education. *Educational Psychologist, 37*(2), 67-68.

Tucker, B. (2012). The flipped classroom. *Education Next, 12*(1), 82-83.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2): 64-70.

Analysis on Students' Acceptance of Digital Reading in Ubiquitous Cooperative Inquiry-based Learning Environment

Jing-Ya CHEN[a], Jing LENG[b], Xiao-Juan XU[c] & Xiao-Qing Gu[d*]
East China Normal University, Shanghai, China

Abstract: Due to the advanced ubiquitous technology, learning can happen almost anywhere and at any time. Most students can learn by using mobile devices they want. Digital reading is an important part of students' learning which can increase their learning interesting and support them learning. And ubiquitous cooperative inquiry-based learning demonstrates improvement of student engagements and expectations for personal success by learning and sharing experiences in small groups. More and more researchers become interested in examining the potential benefits of digital reading supported by ubiquitous technology. This paper tries to investigate students' use of digital reading in a classroom context while they engage in collaborative inquiry-based learning. Studies were conducted to examine whether individuals' reading motivation will affect they acceptance of digital reading and test scores of the subject. Results show that firstly, reading motivation correlates highly with students' acceptance of digital reading; secondly, if students have higher degree of acceptance of digital reading, their digital reading behaviors show more active. Teachers need to pay more attention to students' reading motivation to improve their digital reading performance and learning outcome.
Keywords: Digital reading, ubiquitous technology, user acceptance, reading motivation

1. Introduction

With the popularity of wireless communication and ubiquitous technologies, ubiquitous and mobile learning has become more and more important and helpful in education (Chu, Hwang, Tsai &Tseng, 2010). Ubiquitous learning is often regarded as "learning anywhere and at anytime "(Hwang& Tsai, 2011).Mobile devices (e.g. smart phones, laptops or iPad) are emerged to assist students to learn in a ubiquitous learning environment. Supported by these ubiquitous technologies, students can acquire and communicate knowledge in anywhere, and at anytime. Reading is a necessary part in every discipline. The paradigm of reading, in particular for young people, is increasingly using digital reading rather than paper (Mangen, Walgermo, & Br ønnick, 2013). Digital reading provides students a lot of functions (e.g. highlight, hyperlinks, notes) to assist them to read. If students can use these digital reading tools effectively, it might improve their memory and learning outcome. Motivation can affect their performance in different learning areas, including reading (Wigfield & Guthrie, 1995). And the level of user acceptance of information technology will also affect their using behaviors (Davis, 1989). There exists a large body of researches on how to use these technologies to enhance teaching and learning in different disciplines, but few are focused on the impact factors of students' acceptance of digital reading in ubiquitous cooperative inquiry-based learning environment.

Two studies were conducted in this paper. First, a study was conducted to investigate students' usage of the network and social software usage (N=1327). Based on the first study, 102 students in Grade 7 were invited to participate in the second study, to further investigate their use of digital reading in a classroom context while they engage in collaborative inquiry-based learning. In this study, we designed a collaborative inquiry-based learning activity. PDF tool and MOODLE platform were used in this study to support students' reading and learning. Students were encouraged to use reading tools (e.g. highlight, notes) in learning activities, for better memorizing and understanding. Analyses were conducted to examine whether individuals' reading motivation will affect their acceptance of digital reading and test scores of the subject.

2. Literature review

2.1 Ubiquitous learning

Yahya et al. (2010) proposed that Ubiquitous learning(U-learning) is a learning paradigm which takes place in a ubiquitous computing environment that enables learning the right thing at the right place and time in the right way, to make it easier to understanding the concept of u-learning. Ubiquitous computing technology in u-learning constructs a ubiquitous learning environment to enable anyone to learn at anyplace at any time. In recent years, a variety of computing and communication technologies have been developed, such as wireless communication equipment, smart mobile phones, PDAs (Personal Digital Assistant), which are being used in our daily life (Sakamura et al., 2005; Friedewald et al., 2011; Yahya et al. 2010) .For example, smartphones have been proved possible to serve as u-learning devices by researchers (Shin et al., 2011; Chen et al., 2009). A student equipped with a mobile device can connect to any other devices, and access the network by using wireless communication technologies (Uemukai et al., 2004). So creating an effective ubiquitous learning environment can increase students' engagement in learning.

2.2 Reading motivation

It's well known that children's motivation can affect their performance in different learning areas like reading (Wigfield & Guthrie, 1995). Previous researches suggest that motivation is very important to reading engagements (Wigfield etc., 2004), which can indicate children's behaviors, including choice of which activities to do, hold on at these activities, and their level of effort expended (Wigfield, 1997; Guthrie & Wigfield, 1999). Reading motivation represents students' personal goals, values and beliefs with regard to the topic, processes, and outcomes of reading (Wigfield & Guthrie, 2000).It can influence children's reading skills (Morgan& Fuchs, 2007). So reading motivation gives students powers to read and learn. Other studies use strategies to stimulate students' motivation in order to improve their reading in different subjects, such as science (Guthrie& Wigfield, 1999). In a digital world, students' learning behavior might be totally different when they are in traditional learning. Vogel, Kennedy, and Kwok (2009) suggested that students' motivation plays a significant role in engaging and sustaining students to use mobile devices for learning purposes. So some features of digital reading devices might influence students' reading motivation, and then affect their attitudes and behaviors of reading. But few studies focus on the relationship between reading motivation and digital reading behaviors. Therefore, this study attempt to find whether reading motivation related to degree of digital reading acceptance and reading behaviors in a digital reading environment.

2.3 User acceptance of digital reading

With the development of information technology and network, a lot of digital services support people reading and learning. Although almost all the children today are regard as "digital natives" and familiar with information technology, their acceptance and usage pattern of technology are different (Kennedy, 2010).Therefore, user acceptance of information technology might affect students' learning in digital learning environment. Prior researches have done some study about user acceptance of information technology, mainly focusing on assessing the design and application of systems, such as on-line learning systems (Saadé& Bahli, 2005). According to literature review, many studies analyses user acceptance level and factors by using IS success, Task- technology fit and User Acceptance of Information Technology (TAM) model .TAM is the most widely used in the study about user acceptance by using different technologies, including perceived usefulness, perceived ease of use, attitude toward using, and actual usage behavior (Davis, 1993).Davis (1989) also suggested that both usefulness and ease of use had important correlation with usage behaviors, and usefulness had a significantly greater correlation with usage behavior than did ease of use. Bennett (2008) noted that some potential factors might affect user acceptance of information technology like Socio-economic status, sex and specialized disciplines etc. Jones and Healing (2010) claimed that students' Initiative decided the students' participation in technology. Davis (1992) once used motivation theory to understand user acceptance and usage of new technology. So this study put user acceptance into a specific reading situation, and analyzes the correlation between reading motivation and user acceptance of digital reading.

2.4 Reading behavior

As to students' reading behavior, previous studies have investigated how students read in traditional learning. Morrow, Rand, & Smith (1995) suggested that read aloud behaviors in upper elementary grades can improve story reading. With the change in the traditional learning environment, the usage of ubiquitous technology support students' learning process, and change their learning behaviors. Past researches have focused on user behaviors in Web environments (Liu, 2005; Nicholas, et al., 2008).Liu (2005) claimed that digital readers are likely to develop the screen-based reading behavior which is characterized by more time spent on browsing and scanning, keyword spotting, one-time reading, non-linear reading, and reading more selectively. Until recently, little attention has been drawn to analyze reading behaviors when in a specific discipline situation. This study researches reading behaviors in the digital environment from a different perspective by observing students' reading behaviors (e.g. highlight, make notes and using navigation) in digital reading process and analyzing the relationship between students' acceptance of digital reading.

Recent researches have mentioned that motivation, user acceptance, behaviors and outcomes have some correlation, but not take it a specific learning situation into consideration. Therefore, this paper tries to examine whether individuals' reading motivation will affect their acceptance of digital reading and test scores of the subject in a ubiquitous collaborative inquiry-based learning environment.

3. Research question

In this study, two main research questions were addressed regarding students' use of digital reading in a classroom context while they engage in collaborative inquiry-based learning:

What are the relationships between students' reading motivation, digital reading acceptance and test scores?

For students with high degree of acceptance of digital reading, do they also show more active digital reading behaviors?

4. Design and Method

4.1 Contexts

In these years, Shanghai has been focusing on the IT construction in basic education and supplying personalization and ubiquitous quality education to learners from the aspect of policy and practice. The experiment school we choose is one of the earliest experimental bases for Information technology education in Chang Ning district in Shanghai. Chinese, Mathematics, English, Geography and other disciplines are involved into ICT support teaching research projects.

The reason why this study chooses geography as the experimental subject is that, this subject has its own website and laboratory room to support students studying in a ubiquitous learning environment; furthermore, in geography class reading and remembering a lot of information in maps and pictures are required. It has a high demand for color and multiple media. Therefore, geography materials are more suitable for the student to carry on digital reading.

4.2 Participants

Two studies were conducted to investigate usage of network in middle school students and understand students' acceptance of digital reading while they engage in collaborative inquiry-based learning. The participants in two studies are all from one middle school in Chang Ning district in Shanghai.

In the first study, there were 1327 student participants. A survey is designed to analyze the usage of network and social software (eg. Wechat).The second study tries to understand how digital reading can be integrated into a collaborative inquiry-based learning. In the second study, 102 students in grade 7 were involved. They were between 14 and 15 years old (49.1% female, 50.9% male).The experiment was conducted during their geography study.

4.3 Instruments

In the first study, a questionnaire is designed to investigate students' network usage. This questionnaire has 16 items which includes respondents' basic information, usage of network and social software.

In the second study, we use the reading motivation scale based on the Motivations for Reading Questionnaire (MRQ) which designed by Wigfield (1996) to measure students' reading motivation. The MRQ contains 54 items and assesses 11 possible dimensions of reading motivations, including reading efficacy, reading challenge etc. It can be used with children in late elementary school and middle school.

On measuring students' acceptance of digital reading, this study uses a model of technology acceptance to examine students' feeling about using mobile devices to read and learn in a cooperative geography learning course. This questionnaire is adapted from the technology acceptance model (TAM) which has been widely used to study user acceptance of new computer technologies.

By observing the students' digital reading behaviors in geography class, we designed a digital reading behaviors questionnaire (6 items) to analyze the correlation between degree of acceptance of digital reading and digital reading behaviors.

4.4 Design

This study integrated reading process into in a ubiquitous Jigsaw cooperative learning environment. During four weeks inquiry−based learning activities, whether there is a significant correlation between students' reading motivation and digital reading acceptance and test scores are investigated.

Our study choose'1.4 Hu-Ning-Hang district' of the seventh grade geography course as research content. The four-week inquiry theme was "the design and production of thematic maps in Hu-Ning-Hang district in Shanghai". The detail of learning activity design as shown in Table1.

Table 1: Learning activity design of Hu-Ning-Hang district

the stages of Inquiry Learning	**Students' activities**
Stage 1a: expert views on one aspect of background research	Students were separated into four expert groups which read different digital materials. During the reading phase, students should utilize the annotation services to make annotations (such as highlighting and underlining) on emphatic text and then answered the inquiry questions in MOODLE.
Stage 1b: Background research on all aspects of Hu-Ning-Hang district in shanghai	Reorganized new teams which each team was made up by four different experts and every team discuss the questions in MOODLE and made a presentation.
Stage 2: Design proposal preparations	Read digital materials the teacher offered, each team chose what kind of thematic map they wanted to make. And find relevant information online, copy useful information and upload the final version to the MOODLE.
Stage 3: Peer review of design proposals for the thematic map in Hu-Ning-Hang district in shanghai.	Draw the thematic map on computer or ipad according to the design proposal, and then upload the final thematic map to MOODLE.
Stage 4: Summary and reflection	Discussed and evaluated other teams' works in the

	MOODLE and class.

The following questions were posted by the researcher in four stages to support the online discourse in a ubiquitous environment:

Stage 1a: Expert views on one aspect of background research

Question: Imagine you are a geologist, how would you introduce related information to someone who wants to know the advantages and disadvantages about geographical location, natural conditions, economic production, tourism and cultural aspects in Hu-Ning-Hang district in shanghai? Give your opinion.

Stage 1b: Background research on all aspects of Hu-Ning-Hang district in shanghai

Imagine you are one member of the Group of Experts on of China Geography, and you have been asked by the superior to design a thematic map in Hu-Ning-Hang district in shanghai:

Question 1: After reading some case studies about thematic map, what processes or stages do you think would be involved in developing a thematic map? Make a decision about what kind of thematic map your group would design.

Question 2: What information do you need? How do you decide which information is needed? Why is the information important?

Stage 2: Design proposal preparations

The superior has listened to the presentations of the expert groups and he would like all design groups to answer a few further questions during the preparations of the design proposal.

Question 1: What kind of materials are you intending to put into your thematic map?

Question 2: How to use your thematic map and what are the advantages of it compared with general map?

Stage 3: Peer review: review the designed proposals for the thematic map in Hu-Ning-Hang district in shanghai.

Question: after reviewing other groups' proposal, can you tell why do they design their thematic map like this? Based on the aspects of knowledge acquired in stage 1 and 2, would you elaborate factors that influence their design proposal?

Stage 4: Summary and reflection

Question: Now you have completed design of a thematic map in Hu-Ning-Hang district in shanghai. Based on your insights and knowledge gained over the four stages, can you discuss for some thematic map design theories specific to Shanghai for next year's students?

4.5 Data resources and method

In the first study, 1327 questionnaires were collected from all the students in this experiment school. In the second study, the reading motivation scale was distributed before the experiment, degree of acceptance of digital reading questionnaire and digital reading behaviors questionnaire were distributed after the four- week courses. 102 questionnaires were collected, of which 91 are valid. A correlation analysis was used to investigate the correlation with students' reading motivation, degree of acceptance of digital reading and digital reading behaviors. All analyses were conducted by using SPSS version 19.0.

5. Analysis and Results

5.1 Usage of network and ubiquitous technology

In order to know the students' network and ubiquitous usage, a survey were conducted in this experiment school. The following section describes some important results.

In terms of the network usage, over half students (53.05%) prefer using smartphone to search information online. And only 95 out of 1327 students do not use internet. In the question of "Do you really like the Internet", almost 72% students very like internet, and most of them (65.56%) have ability of self-control, knowing when and how to use internet. In addition, when students meet difficulties in study, they (55.09%) first choose to search solutions online. Thus the data can imply that internet has become a large part of the students' life and they were familiar to use these internet devices with their own preferences.

As to students' attitudes about smartphones, 75.51% participants have a smartphone and the trend is still increasing. 91.18% participants think smartphone can be used as a learning tool. Smartphone can be regarded as a kind of ubiquitous technology tools which supports students to learning at anytime, anywhere. From the results, the participants' attitudes about smartphone is positive and they (43.48%) think smartphone don't affect their learning in a bad way.

In terms of the social software usage results, although over half of the participants (60.74%) still use phone or message to contact with family and friends, 39% prefers to use some social software such as QQ, Email, Wechat etc. And 945 participants (71.21%) have kept online in Wechat all day.

5.2 The relationship between reading motivation, learning outcome and students' acceptance of digital reading

Seen from table 2, participants' reading motivation and the reading scores of geography lesson were positively correlated, and reached statistical significance (r=．269，P＜0．01)in a digital reading environment. These results supported the comment that reading motivation affects directly the reading scores (Wigfield &Guthrie, 1995).Therefore, when students study in a digital learning environment, the one who has stronger reading motivation will has a better reading score.

Table 2：The relationship between reading motivation, learning outcome and students' acceptance of digital reading (N=91)

		reading motivation	Geography scores	Acceptance of digital reading
reading motivation	Pearson Correlation	1	.269**	.285**
Geography scores		.269**	1	
Acceptance of digital reading		.285**		1

**. Correlation is significant at the 0.01 level (2-tailed).

The result also indicated positive correlations between participants' reading motivation and degree of acceptance of digital reading reached statistical significance(r=．285，P＜0．01). So when students' demands for reading are higher, they show more positive attitudes for digital reading and more willingness to accept this new reading method. Among the sub-items of reading motivation (not shown), reading curiosity and reading efficiency are correlated with degree of acceptance of digital reading, but not very strong. Thus, if students have more desire and curiosity to read, they would prefer to try fresh reading methods. Therefore, it means that teachers should take students' reading motivation into consideration while using ubiquitous technology, in order to increase students' acceptance of digital environment. In these results, one thing should be considered is that although the co-relations are statistically significant, its effect sizes (r square) are not that high, which led to this result may be due to test questions in geography might not very relevant with digital reading contents. It will be improved in the further research.

5.3 The correlation between degree of acceptance of digital reading and digital reading behaviors

Students' acceptance of digital reading show high correlation with participants' reading behaviors which reached statistical significance(r=．706，P＜0．01) as shown in table 3. In a cooperative inquiry-based learning environment, if a student has a high degree of acceptance when he read digital materials, it means his digital reading behaviors will be more active. Results from table 4 indicated that all the sub-items of digital reading behaviors correlated highly with students' degree of acceptance of digital reading. This study focuses on students' digital reading behaviors including using navigation, highlight, clicking hyperlinks, taking notes, searching tools, and reread the notes or highlights they have made, which can support students reading and understanding. Thus, the higher degree of acceptance of digital reading students have, the better they can adapt to digital reading. They will use digital reading tools more actively.

Table 3: The correlation between degree of acceptance of digital reading and digital reading behavior (N=91)

		Acceptance of digital reading	Digital reading behavior
Acceptance of digital reading	Pearson Correlation	1	.706**
Digital reading behavior		.706**	1

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4: The correlation between degree of acceptance of digital reading and subitems of digital reading behavior (N=91)

		Using navigation	highlighting	Clicking hyperlinks	Take notes	Search tool	Reread the highlights/tags
Acceptance of digital reading	Pearson Correlation	.564**	.557**	.588**	.577**	.604**	.506**

**. Correlation is significant at the 0.01 level (2-tailed).

6. Conclusion

This paper mainly presents two studies. From the first study, the results indicate that almost all middle school students have passion on network learning and have ability to acquire knowledge by using ubiquitous technology (e.g., iPad, iPhone); furthermore, "digital natives" would like connect other people by using social software (e.g. QQ, Wechat), they have a high level of using mobile devices; thirdly, as to ubiquitous learning, 72% participants think mobile phone can be a learning tool, their attitudes are positive. Overall, if a good guide can be provided to students about how to use ubiquitous technologies to support learning, it might help them using network better.

In the second study, participants completed digital reading and cooperative inquiry-based learning activities in a geography course. This study focuses on investigating students' reading motivation, acceptance of digital reading and reading behaviors in digital learning environment. Results show that there's a high correlation among reading motivation, students' acceptance of digital reading and learning outcome. And if students have higher degree of acceptance of digital reading, their digital reading behaviors show more active. So when teaching in a digital learning environment, teachers can use some strategies to increase students reading motivation and students' acceptance of digital reading, which might affect their learning behaviors and improve their digital reading performance and learning outcome.

Certainly, this paper still has some limitations that need to be considered. In the second study, limitations stem from its scope, particularly the size and composition of the sample population. And in the data analyses process, a few of incomplete data were deleted, it might effect on the results of correlation analysis. But we try to remain the maximum data authenticity. This study has found that the relationship between reading motivation, students' acceptance of digital reading and digital reading behaviors, there is a need for future research about how mobile technologies can be used to enhance learners' motivation and learning outcome.

Acknowledgements

We would like to thank the geography teacher Mr.lai who supported us in completing the experiment and questionnaire.

References

Bennett, S., & Maton, K. (2010). Beyond the 'digital natives' debate: Towards a more nuanced understanding of students' technology experiences. *Journal of computer assisted learning*, 26(5), 321-331.

Chen, J. V., Yen, D. C., & Chen, K. (2009). The acceptance and diffusion of the innovative smart phone use: A case study of a delivery service company in logistics. *Information & Management*, 46(4), 241-248.

Chu, H. C., Hwang, G. J. & Tsai, C. C. (2010). A knowledge engineering approach to developing mindtools for context-aware ubiquitous learning. *Computers & Education*, 54, 1, 289–297.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology).

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International journal of man-machine studies*, 38(3), 475-487.

Friedewald, M., & Raabe, O. (2011). Ubiquitous computing: An overview of technology impacts. *Telematics and Informatics*, 28(2), 55-65.

Guthrie, J. T., & Wigfield, A. (1999). How motivation fits into a science of reading. *Scientific studies of reading*, 3(3), 199-205.

Hwang, G. J., Tsai, C. C., & Yang, S. J. (2008). Criteria, Strategies and Research Issues of Context-Aware Ubiquitous Learning. *Educational Technology & Society*, 11(2), 81-91.

Hwang, G. J., & Tsai, C. C. (2011). Research trends in mobile and ubiquitous learning: A review of publications in selected journals from 2001 to 2010.*British Journal of Educational Technology*, 42(4), E65-E70.

Jones, C., & Healing, G. (2010). Net generation students: agency and choice and the new technologies. *Journal of Computer Assisted Learning*, 26(5), 344-356.

Kennedy, G., Judd, T., Dalgarno, B., & Waycott, J. (2010). Beyond natives and immigrants: exploring types of net generation students. *Journal of Computer Assisted Learning*, 26(5), 332-343.

Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of documentation*, 61(6), 700-712.

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61-68.

Morgan, P. L., & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional children*, 73(2), 242-255.

Morrow, L. M., Rand, M. K., & Smith, J. K. (1995). Reading aloud to children: Characteristics and relationships between teachers and student behaviors. *Literacy Research and Instruction*, 35(1), 85-101.

Nicholas, D., Rowlands, I., Clark, D., Huntington, P., Jamali, H. R., & Ollé, C. (2008, July). UK scholarly e-book usage: a landmark survey. In *Aslib Proceedings* (Vol. 60, No. 4, pp. 311-334). Emerald Group Publishing Limited.

Sakamura, K., & Koshizuka, N. (2005, November). Ubiquitous computing technologies for ubiquitous learning. In *Wireless and Mobile Technologies in Education, 2005. WMTE 2005. IEEE International Workshop on* (pp. 11-20). IEEE.

Saadé, R., & Bahli, B. (2005). The impact of cognitive absorption on perceived usefulness and perceived ease of use in on-line learning: an extension of the technology acceptance model. *Information & Management*, 42(2), 317-327.

Shin, D. H., Shin, Y. J., Choo, H., & Beom, K. (2011). Smartphones as smart pedagogical tools: Implications for smartphones as u-learning devices. *Computers in Human Behavior*, 27(6), 2207-2214.

Uemukai, T., Hara, T., & Nishio, S. (2004, March). A method for selecting output data from ubiquitous terminals in a ubiquitous computing environment. In *Distributed Computing Systems Workshops, 2004. Proceedings. 24th International Conference on* (pp. 562-567). IEEE.

Wigfield, A. (1996). A Questionnaire Measure of Children's Motivations for Reading. Instructional Resource No. 22.

Wigfield, A. (1997). Reading motivation: A domain-specific approach to motivation. *Educational Psychologist*, 32(2), 59-68.

Wigfield, A., & Guthrie, J. T. (1995). Dimensions of Children's Motivations for Reading: An Initial Study. Reading Research Report No. 34.

Wigfield, A., & Guthrie, J. T. (2000). Engagement and motivation in reading. *Handbook of reading research*, 3, 403-422.

Wigfield, A., Guthrie, J. T., Tonks, S., & Perencevich, K. C. (2004). Children's motivation for reading: Domain specificity and instructional influences. *The Journal of Educational Research*, 97(6), 299-310.

Wellings, J., & Levine, M. H. (2009). The Digital Promise: Transforming Learning with Innovative Uses of Technology. In *New York: Joan Ganz Cooney Center at Sesame Workshop*.

Yahya, S., Ahmad, E. A., & Jalil, K. A. (2010). The definition and characteristics of ubiquitous learning: A discussion. In *International Journal of Education and Development using Information and Communication Technology*, 6 (1), 117-127.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Visualizing Ubiquitous Learning Logs Using Collocational Networks

Kousuke Mouri [1], Hiroaki Ogata[2], Noriko Uosaki[3], SongRan Liu[4]
[1,2]*Faculty of Arts and Science and the Graduate School of Information Science and Electrical Engineering, University of Kyushu, Japan*
[3]*Center for International Education and Exchange, Osaka University, Japan*
[4]*Advanced Technology and Science, University of Tokushima, Japan*
*{mourikousuke, hiroaki.ogata, n.uosaki, lb90518}@gmail.com

Abstract: This paper describes a system that can be used to visualize some ubiquitous learning logs using collocational networks to discover several learning patterns. Visualization of the system is based on vast amount of learning data in ubiquitous learning environment. Ubiquitous Learning Log (ULL) is defined as a digital record of what learners have learned in the daily life using ubiquitous technologies. It allows learners to log their learning experiences with photos, audios, videos, location, RFID tag and sensor data, and to share and to reuse ULL with others. This paper will reveal about the relationship between the ubiquitous learning logs and learners by using network graph and collocational networks. Also, this paper will explicate the system through which learners can grasp their learning time, histories, knowledge and location.

Keywords: ubiquitous learning, network graph, time-map, information visualization, collocational networks

1. Introduction

Recently, researchers in the educational engineering area have been studying focusing on ubiquitous themes. For example, CSUL (Computer Supported Ubiquitous Learning) or context aware ubiquitous learning (u-Learning) have been constructed using computing technologies such as mobile devices, QR-code, RFID tag and wireless sensor networks (Hwang et al., 2008; Ogata & Yano, 2004). These learnings take place in a variety of learning space such as classroom, home and museum. Also, the cutting-edge technologies can provide the right information using the contextual data like location, surrounding objects and temperature.

Therefore, many researchers have been focusing on effective learning with ubiquitous technologies. We have developed ubiquitous learning system called SCROLL (System for Capturing and Reminding of Learning Log) (Ogata et al., 2011). The system will support international students to learn their target languages. Traditionally, international students take memos when they have learned something in their daily lives. However, if the notes have not been taken in detail, they can neither actively recall what they have learned, nor the location where they learned them. Therefore, we have proposed SCROLL which enables learners to recall their past learning experiences by saving them to the system with location, photo, or video as digital records.

Also, these learning dataset include spatiotemporal data. Spatiotemporal data usually contain the states of an object, an event or a position in space over a period of time. These datasets might be collected at different locations, various time points in different formats. It poses many challenges in representing, processing, analysis and mining of dataset due to complex structure of spatiotemporal objects and the relationships among them in both spatial and temporal dimensions (K.Venkateswara Rao et al., 2011, 2012).

Similarly, it poses many issues about relationship between the learners and the ubiquitous learning logs due to complex structure of the ubiquitous learning logs in SCROLL. In addition, it is important for learners to recognize what and how they have learned by analyzing and visualizing the past ULLs, so that they can improve what and how to learn in future (Ogata et al., 2011). To tackle these issues, it is necessary to reveal relationships between the learners and the ubiquitous learning logs.

Therefore, this paper proposes a method to visualize relationships between the learners and the ubiquitous learning logs using Time-map, network graph and collocational networks. The objective of this study will reveal what and how the learners learned language in their daily lives. Therefore, this paper is aimed to recommend appropriate learning patterns and trends for the learners, using collocational networks.

2. Related Works

2.1 Learning Analytics and Knowledge

In recent years, Learning Analytics and Knowledge (LAK) has been drawing an attention from researchers of such fields as educational engineering, information science and network science. To date, Course Management System (CMS) and Learning Management System (LMS) enabled us to record learners' access logs onto server. The Learning Analytics (LA) aims for practical use based on learning mechanisms revealed by visualizing, mining and analyzing vast amount of learning data (Ferguson 2012). This paper focuses on the Social Learning Analytics (SLA), a subset of the LAK (Buckingham 2012). The SLA puts forward presenting appropriate information to learners at the appropriate timing through the Dashboard in real time. As a new challenge, this paper aims to reveal about relationships between learners and learning logs on spatiotemporal fields.

Therefore, this paper is expected to contribute to educational improvement and strategies below.

- This study facilitates the analysis of learners by visualizing all data on spatiotemporal.
- This study enables future prediction about learners and learning environment from visualized learning logs.

2.2 Time-map

Time-map is a library of javascript, which collaborated with Google map and SIMILE (Semantic Interoperability of Metadata and Information in unLike Environments) TimeLine (SIMILE project). SIMILE focuses on developing robust, open source tools that empower users to access, manage, visualize and reuse digital assets. The time-map function means that the user can scroll the timeline and then the Google map will display the learning logs recorded during learners' selected period. It is designed to help learners to reflect what they have learned. For example, if a learner clicks his learning logs on timeline, Google map will display their positions as shown in Figure 1. After visualizing log information, Time-map will facilitate learners to reflect on their logs with spatio and temporal information. They are able to grasp their learning context and time zone. Also, it is a possibility that the geographic information is a clue of recalling what they have learned.

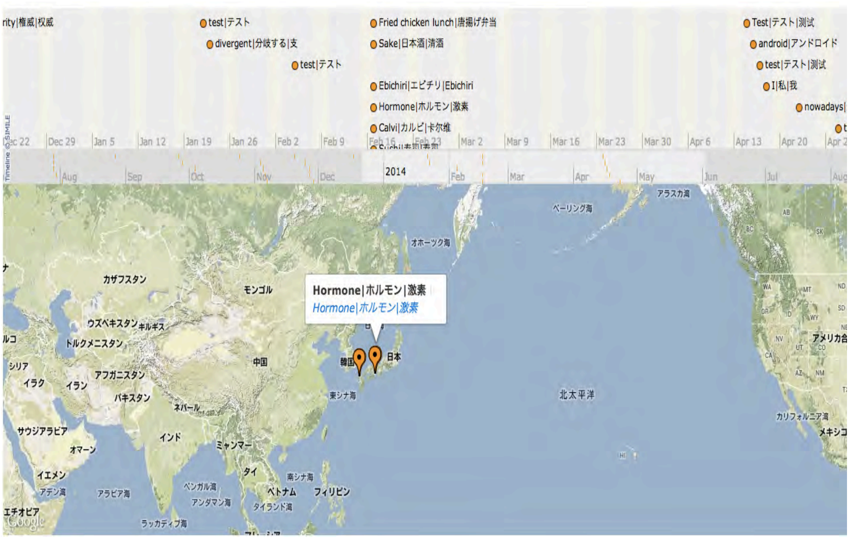


Figure 1: Time-map

3. Design of the system

3.1 SCROLL

With the evolution of the mobile device, people prefer to record learning contents using mobile devices instead of taking memos on paper. Most of the language learners have their own learning notes. In this paper, learning log is defined as a recorded form of knowledge or learning experiences acquired in our daily lives.

One of the objectives of SCROLL is to support international students in Japan to learn Japanese language from what they have learned in formal and informal setting. It adopts an approach of sharing user created contents among users and is constructed based on a LORE (Log-Organize-Recall-Evaluate) model which is shown in Figure 2 (Ogata et al., 2011).

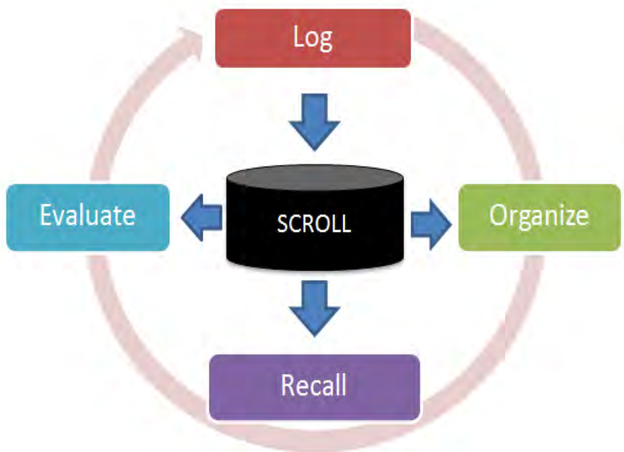


Figure 2: LORE model in SCROLL

SCROLL is a client-server application, which runs on different platforms including Android mobile phones, PC and general mobile phones shown in Figure 3. The server side runs on Ubuntu 12.04.2 and it is programing using Java, Spring MVC and Mybatis. The developed software for Google phone is a native java application based on Android SDK (Li et al., 2012).

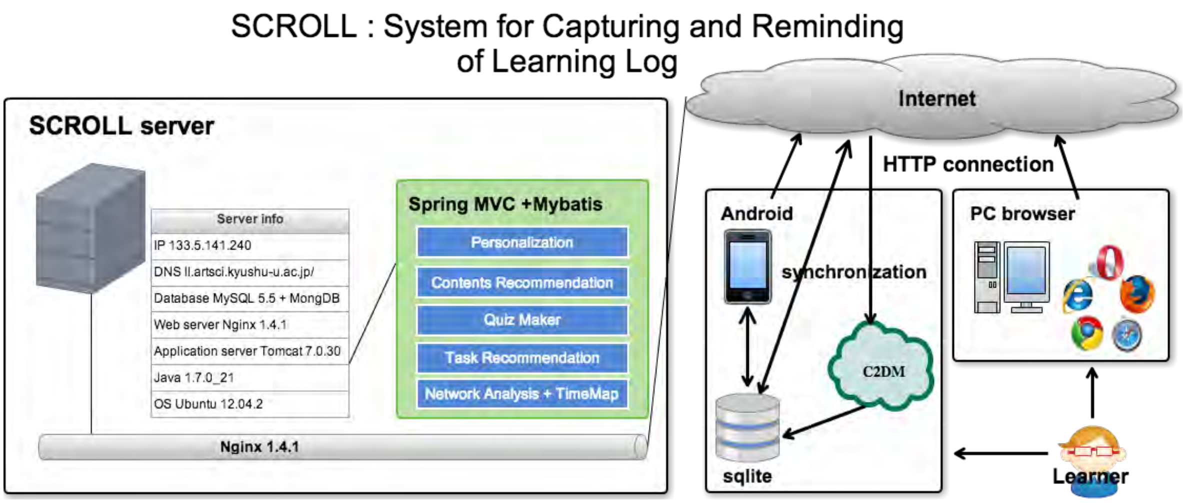


Figure 3: The architecture of SCROLL

3.2 Collecting a ubiquitous learning log on SCROLL

The learners can record some learning language such as English, Japanese and Chinese with a photo using android device and SCROLL as shown in Figure 4. Figure 5 shows a learning log on android device.

The learning log includes meta-data such as author, language, created time, location (latitude and longitude) and tag. The learners will record or review a learning log using these functions on android device. Such iterative learning is supported by our quiz function on SCROLL. There are three types of quizzes generated automatically by the system, which are yes/no quiz, text multiple-choice quiz and image multiple-choice quiz. Figure 6 shows an image multiple-choice quiz interface generated automatically based on the meta-data of ULLs.

Figure 4: Learning note

Figure 5: A learning log

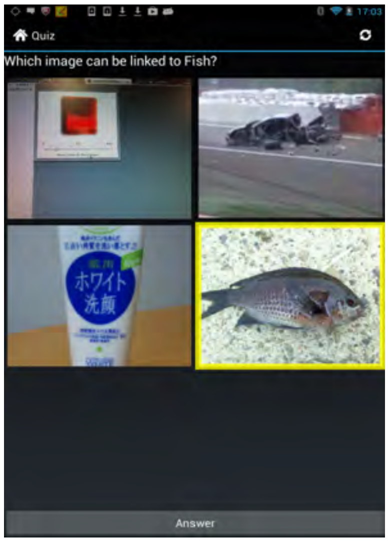


Figure 6: Quiz

3.3 *Structure based on network graph in SCROLL*

To reveal several relationships between the learners and knowledge or knowledge and location, we have uniquely defined them as three-layers structures as shown in Figure 7.

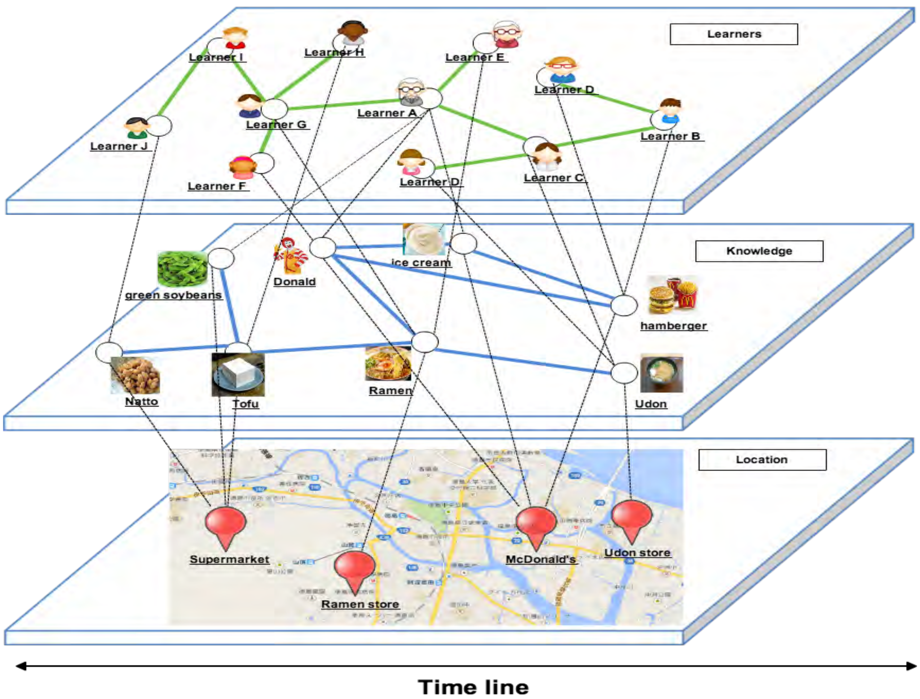


Figure 7: Three-layer structure in SCROLL

The upper layer contains each author in order to confirm position of own or other learners.

The intermediate layer contains the knowledge that learners learned. Also, some fields of learning tasks can be included in this layer. For example, some task-based learning in ubiquitous learning environment can be carried out using knowledge and event (Mouri et al., 2013; Sharon 2013). The scalability of the layers can be enhanced and the field of visualization can be widened by linking one's own learning logs to the knowledge learned by doing tasks.

The lowest layer contains data such as location and time. In order to realize spatiotemporal visualization of our learning logs, nodes on the intermediate layer are linked to the nodes on the lowest layer.

Analysis by categorizing three-layers has following advantages:

- Places with a large number of links to the related knowledge are the places where they can learn a lot of knowledge. For example, if a certain supermarket or convenience are related with a lot of knowledge such as natto, green soy beans, tofu, miso soup, and cup noodle, by analyzing relationships between the knowledge and the location the System can provide learners with a valuable learning information.
- Knowledge which is related to many places is the knowledge which we can learn in various places. For example, if a learner experience tea ceremony of a traditional Japanese culture at the university in Japan, a set of tea ceremony related knowledge (eg. tea, seize: to sit in the correct manner on a Japanese tatami mat) can be learned in other various places. The tea can be learned by purchasing at the supermarket and the seiza can be learned at the martial art gym.

3.4 Collocational networks in SCROLL

Collocational networks are two-dimensional networks which contain interlinked collocation, i.e. word which occur together in a text. The concept of collocational networks originates in an article by Williams (Williams 1998). In his study, Williams uses the network as a corpus linguistic tool in order to create specialized dictionaries. Also, Magnusson describes an important to visualize most central concept in the text. There are some collocational networks on the SCROLL. For example, if a learner A learned natto (a traditional Japanese food made from fermented soybeans) at the supermarket, he/she might learn other food at the same time. Therefore, the collocational network can show relationships between the knowledge and time. Similarly, it can show collocational relationships between knowledge and place and place and time.

In this paper, we propose a method using collocational networks in order to predict their learning patterns and trends in the future and to expand their field of view. Using recorded collocational data on SCROLL, this study have constructed the collocational networks.

Firstly, the collocational relationships between knowledge on the intermediate layer will link in time-series order what they have learned. For example, if the learner A learned a tofu in the next learning after studying a natto, the natto and tofu on the intermediate layer will be connected. Also, if the learner B learned a green soy beans after studying a natto, the natto and the green soy beans on the intermediate layer will be connected in the same way. By linking their knowledge and knowledge in the next learning, SCROLL can be predicted knowledge that they might be able to learn in the next learning.

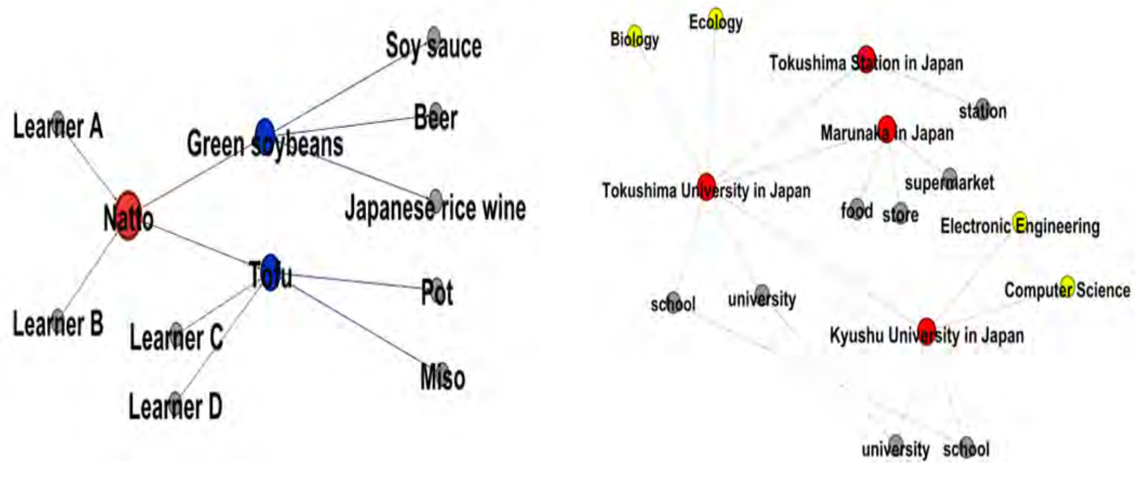


Figure 8: Relationships between knowledge, place and time on collocational networks

However, their notable knowledge is a possibility which is connected many edges. For example, after the learner A and B learned the natto, there are two learning path as described above. That is, there are two paths of "they learn the tofu after learning the natto" and "they learn the green soy beans after learning the natto" as shown in figure 8 (left network image). Figure 8 is one of an example of some learning patterns, and there are many learning patterns in actually in the ubiquitous learning environment.

Secondly, in this paper, we attempt to categorize the names, such as hospital, university and restaurant of nearest place where they have learned, using Google place api from the latitude and longitude of past ubiquitous learning logs. Also, some buildings acquired from its place include a lot of attributes. Hence, it is necessary to construct collocational network including the place and the attributes.

For example, if a learner A learned the word such as "electronic engineering" and "computer science" at the kyushu university in japan after studying some word such as "ecology" and "biology" at the tokushima university in japan, it is created edge to connect "tokushima university" to "kyushu university" on the lowest layer. On the other hand, a learner B learned another word at the tokushima station near the university after studying some words such as "ecology" and "biology" at the tokushima university in japan, it is created edge to connect "tokushima university" to "tokushima station" on the lowest layer. In addition, prediction of the next learning place is determined by the number of the attributes of the place. For example, the collocational network shown in Figure 8 (right network image) is connected from the tokushima university to tokushima station, marunaka (Supermarkets' name in japan) and kyushu university. The attributes of the tokushima university include "university" and "school". Similarly, the tokushima station includes "station", and the marunaka includes "supermarket", "shop" and "food", and the kyushu university includes "university" and "school". By comparing to attributes of three types, it is evident that the attributes of the tokushima university and kyu-shu university are same.

Therefore, there are a possibility that two learning place are high relationship. However, there is a possibility that the distances between two places are a far. That is, it is necessary to calculate the distances between current position of the learner and target learning place. Also, as described above, to expand a field of their view regarding the place is effectiveness in order to understand learning situation and context.

4. Implementation

This section describes ways of the implementation of the system for visualizing the three-layer structure using network graph using collocational networks and Time-map.

4.1 System for visualizing network graph in SCROLL

4.1.1 How to create node or connect edge on three-layers

Firstly, system for visualizing network graph will create authors' node on the upper layer. To date, the number of learners in SCROLL is approximately three thousand people.

Secondly, the system will create knowledge node on the intermediate layer. Then, the system will connect authors' node related to knowledge node that learners have learned. For example, if learner A learned a learning log like “natto”, “tofu (bean curd)” and “sushi”, the system will connect "learner A" of node on the upper layer to "natto", "tofu" and "sushi" on the intermediate layer.

Thirdly, the system will create location node on the lowest layer. Then, the system connect knowledge node on the intermediate layer to node of the location on the lowest layer. For example, if the learner A have learned knowledge of "natto" at the supermarket in Japan, the system will connect "natto "on the intermediate layer to the latitude and longitude of "supermarket" on the lowest layer.

4.1.2 Color of visualized nodes

The learners might get confused when they recognize past learning logs because there might be too many of visualized nodes. Therefore, it is definitely necessary to establish some criteria for distinction of each node. To effectively distinguish kind of each node, we defined as Table 1 below using node color.

Table 1: Color to distinguish the kinds of nodes

Node	Layer	Node color
Learner’s own name	Upper layer	Pink
Names of other learners	Upper layer	Blue
Representative learners	Upper layer	Green
Knowledge of learners	Intermediate layer	Yellow
Location of learners	Lowest layer	Red

- Pink color node shows the learner’s own name on the upper layer. If connecting the pink node to yellow node on the intermediate layer, edge color will be decided as pink so that they can be easily recognized as the learner’s own logs.
- Blue color nodes show the names of other learners on the upper layer. If connecting the blue node to yellow color node on the intermediate layer, edge color is decided blue color.
- Green color node shows the names of famous or representative learner on the upper layer. If connecting the yellow node to the green node on the intermediate layer, edge color will be decided as green color.
- Yellow nodes represent both the learner own knowledge and the knowledge of other learners. For example, the learner can recognize his own knowledge because edge between the learner own name on the upper layer and the knowledge on the intermediate layer is pink color. In addition, the learner might discover knowledge of other learners related to own knowledge.
- Red color node shows the location of learners on the lowest layer. The node includes latitude, longitude and created time.

4.2 Combining network graph and Time-map

The interface of network graph for visualizing relationships between the learners and ubiquitous learning logs is shown in Figure 9. The learners can recognize relationships between own/others author and knowledge by using the network graph interface. Figure 9 shows an example of interface on collocational network based on knowledge. The network layout consists of using two basic layouts and an original layout.

The first layout consists of using Yifan Hu multilevel layout (Y.F Hu, 2001, 2005). It is a very fast algorithm with a good quality on large graphs. It combines a force-directed model with a multilevel algorithm to reduce the complexity. The repulsive forces on one node from a cluster of distant nodes are approximated by a Barnes-Hut calculation (Barnes and P. Hut., 1986), which treats them as one super-node.

The second layout consists of using the random network. It is simple algorism generating them randomly on the graph after filtering some nodes, and then the system will connect relationships related between node and node.

The third layout consists of using original layout we have developed. As shown in Figure 9, the layout regards x axis as time axis. In this figure 9 case, the knowledge that they might be studying in the next learning after studying the natto will be generated to right side in constant interval (Next knowledge are tofu, coffee, router and kimchi).

Recommendation objects in Figure 9 are shown rankings in the learning trends in order to expand a field of their view from visualized ubiquitous learning logs on the network graph. By arranging the in-degree centrality in the high order from the ubiquitous learning logs that they might study in the next learning, the learners are able to recognize famous or representative learners and important knowledge.

Time-map function in Figure 9 consists of the timeline and Google map. It represents the shift of learning history in accordance with lapse of time. The learners might forget the learning logs when and where they have learned before. Therefore, the system can remind the learners of them by combining timeline with map. The system will remind them of their learning logs recorded during the specified period of time by showing them on the timeline (default: two month before and after the setting time). Besides, the system will lead them to be aware of knowledge recorded right before or after the knowledge of their interest which was recorded by other learners. Therefore, it will give them a clue on what to learn in the next learning.

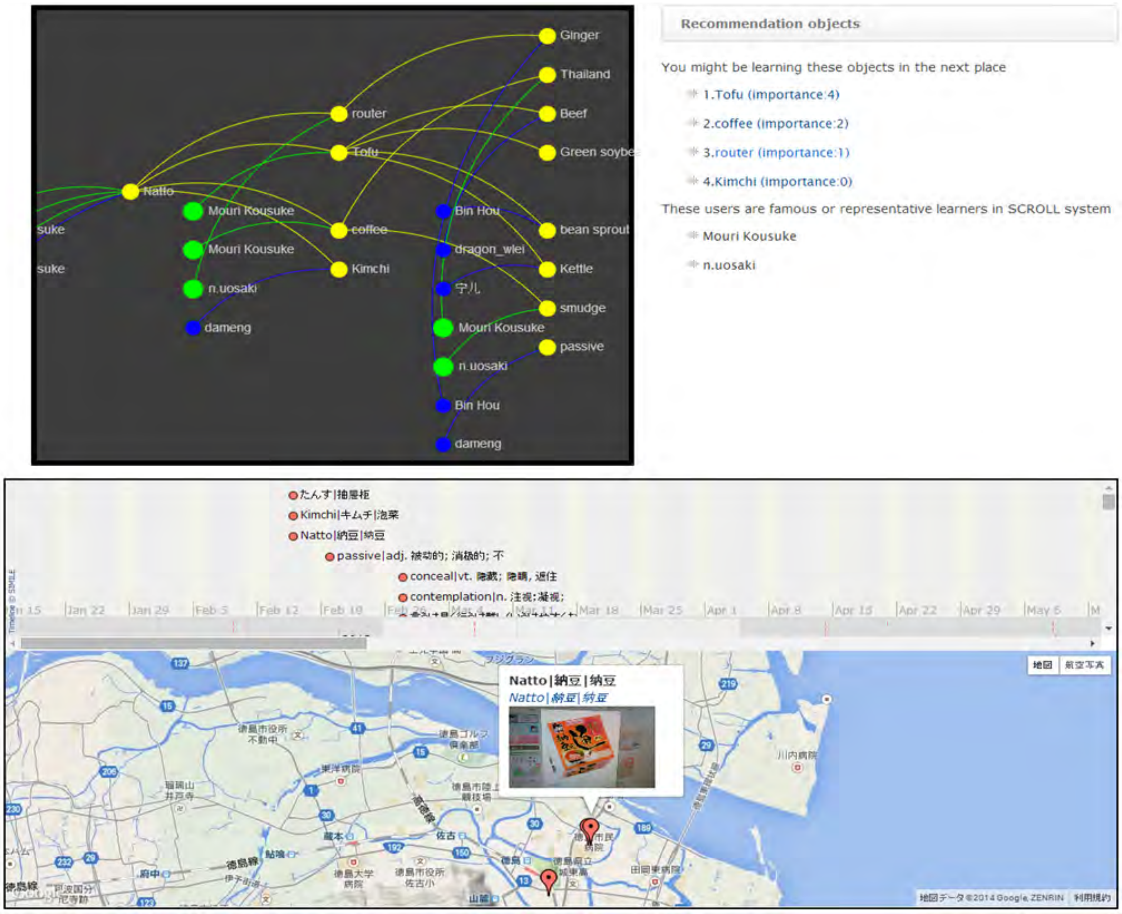


Figure 9: Interface of this system

5. Conclusion and Future work

This paper described the system for visualizing relationships between the learners and the learning logs, using collocational networks. International students can add their knowledge as the learning log in SCROLL, and then SCROLL can provide learning contents to recall what they learned based on their learning contexts. By using the system that we proposed, the international students can discover the knowledge related to others learners and the interesting knowledge.

In the future, we will develop a new function so that the system can analyze various situations focusing learning analytics such as network analysis (Freeman 1978; Shane 2014), decision tree (Bitner 2000) and association rule (Florian 2005).

Acknowledgements

Part of this research work was supported by PRESTO from Japan Science and Technology Agency, and the Grant-in-Aid for Scientific Research No.26350319 from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan.

References

Bitner T. (2000). Rough sets in Spatiotemporal data mining, Proceedings of International workshop on Temporal, Spatial and Spatiotemporal Data Mining, Lyon, France.

Buckingham, S.S & Freguson, R. (2012). Social Learning Analytics, Journal of Educational Technology & Society, 15 (3), 3-26.

Florian, V., & Sanjay, C. (2005). Mining Spatio-Temporal Association Rules, Source, Sinks, Stationary Regions and Thoroughfares in Object Mobility Databases, Technical Report Number 574, the university of Sydney, 187-201.

Freeman, L.C. (1978). Centrality in social networks conceptual clarification, Social Networks 1, 3, 215-239.

Ferguon, R. (2012). The State of Learning Analytics in 2012: A Review and Future Challenges, Technical Report KMI-12-01, Knowledge Media Institute, The Open University, UK.

Hwang, G. J., Tsai, C. C., & Yang, S. J. H. (2008). Criteria, strategies and research issues of context-aware ubiquitous learning. Educational Technology and Society, 11(2), 81-91.

J. Barnes and P. Hut. (1986). A Hierarchical O (n log n) Force-Calculation Algorithm, Nature, 324(4), 446-449.

K.Venkateswara R, A.Govardhan & Dr.K.V.Chalapati R. (2011). Discovering Spatiotemporal Topological Relationships, The second international workshop on Database Management Systems, DMS-2011,Springer Proceedings LNCS-CCIS 198.

K.Venkateswara R & A.Govardhan. (2012). Spatiotemporal data mining: Issues, Task and Applications, International Journal of Computer Science and Engineering Survey, Vol.3, No.1, Feb2012, 39-52.

Li, M., Ogata, H., Hou, B., Uosaki, N., & Yano, Y. (2012). Personalization in Context-aware Ubiquitous Learning-Log System, Mobile and Ubiquitous Technology in Education, 41-48.

Mouri,K., Ogata, H., Li, M., Hou, B., Uosaki, N., & Liu, S. (2013). Learning Log Navigator: Supporting Task-based Learning Using Ubiquitous Learning Logs, Journal of Research and Practice on Technology Enhanced Learning (RPTEL), 8(1), 117-128.

Ogata, H., & Yano, Y. (2004). Context-aware support for computer-supported ubiquitous learning. Proc. of IEEE International Workshop on Wireless and Mobile Technologies in Education, 27-34.

Ogata, H., & Yano, Y. (2004). Knowledge awareness map for computer-supported ubiquitous language-learning.Proc. of IEEE International Workshop on Wireless and Mobile Technologies in Education, 19-26.

Ogata, H., Li, M., Bin, H., Uosaki, N., El-Bishoutly, M., & Yano, Y. (2011). SCROLL: Supporting to share and reuse ubiquitous learning logs in the context of language learning. Research and Practice on Technology Enhanced Learning, 6 (3), 69-82.

Sharon, A. (2013). Storyline: A task-based approach for the young learner classroom, ETL journal, 67(1), 41-51.

Shane, D., Dragan, G., George, S., & Srecko, J. (2014). Current State and Future Trends: A Citation Network Analysis of the Learning Analytics Field, International conference on Learning Analytics and Knowledge (LAK) ACM, 231-240.

SIMILE project: http://www.simile-widgets.org/timeline/

Williams, G. (1998). Collocational Networks: Interlocking Pattern of Lexis in a Corpus of Plant Biology Research Articles, International Journal of Corpus Linguistics 3, 151-171.

Y. F. Hu & J. A. Scolt. (2001). A Multilevel Algorithm for Wavefront Reduction, SIAM Journal on Scientific Computing, 23(4), 1352-1375.

Y. F. Hu. (2005). Efficient and high quality force-directed graph drawing, The Mathematica Journal, vol. 10, 37–71.

The Research of China's Policies and Practices of Life-long Learning in U-learning Environment

Bingqian JIANG[a*], Jun XIAO[b], Jing LENG[a], Xiaoqing GU[a]
[a]*Department of Education Information Technology, East China Normal University, China*
[b]*Shanghai Engineering Research Centre of Open Distance Education, Shanghai Open University, China*
*51130104045@student.ecnu.edu.cn

Abstract: Recently, life-long learning and u-learning are significant concepts in the area of technology-enhanced learning those have changed learning. This paper begins with an overview of u-learning and life-long learning. And then, it focuses on the review of China's national policies of life-long learning. By using Keller's Attention, Relevance, Confidence, Satisfaction (ARCS) model, it finally analyzes Shanghai Life-long Learning Network as the practice case.

Keywords: life-long learning; u-learning; notional policies; learning platform

1. Introduction

The 21st century is an era of knowledge economy and one reform after another has been happening in the education and learning area. Nowadays, the idea of life-long learning has won more and more popularity among people from all walks of life. Building a life-long learning system is no longer just an idea or concept, but evolves into an actual practice of education reform and social activities from which the system is enriched and becomes mature. With the rapid development of computer technology and network communication technology, a new leaning mode based on ubiquitous computing technology-ubiquitous learning, emerges as the times require. Ubiquitous learning is very flexible and practical as anyone who wants to learn could receive education in any form and at any time. This learning concept offers a brand-new life-long learning mode and provides new opportunity for the development of life-long learning, which greatly promotes the building of life-long learning system.

This paper first explores life-long learning theory and ubiquitous learning theory and then makes an interpretation of the future development of current life-long learning from the perspective of China's policy planning. Using Shanghai-an economic center city as the research setting and combining ubiquitous leaning mode, the author investigates the current status of Shanghai's life-long learning system which has adopted ubiquitous learning mode. Then we analyze the design of public service platform open to life-long learners as the case of the infrastructure construction of national policies.

2. Literature Review

2.1 Ubiquitous Learning Overview

As technology becomes more powerful, ubiquitous learning (u-learning) has been a mainstream way of learning in recent years. There are different views in defining and characterizing u-learning, but most researchers believed that u-learning derived from the concept of ubiquitous computing. The term "ubiquitous computing" was coined by Mark Weiser, described as "the calm technology, which recedes into the background of our lives". His vision allows people and the environment with the combination of various computational technologies to exchange information and services at anytime and anywhere (Weiser, 1991). Then, "ubiquitous" was used into the field of education, promoting the generation of "u-learning". Birgit Bomsdorf pointed out, that "ubiquitous computing leads to ubiquitous learning allowing to embed individual learning activities in everyday life." (Bomsdorf, 2005)

To some extent, u-learning combines the strength of e-learning and m-learning. Dey Casey (2005) formulated the view of "u-learning = e-learning + m-learning". Although there are various views of u-learning, a broader definition of it is "anywhere and anytime learning". U-learning is intended to build a learner-centered intelligent environment, supporting learners use all kinds of terminals to access learning content regardless of time and place. In u-learning environment, learning will be more likely to happen, learning support services also will be more humanization, and learners' interaction with the environment will be further strengthened. U-learning truly reflects the "people-oriented" learning conception and brings innovation of learning styles.

2.2 U-learning and Life-long Learning

2.2.1 The related concepts of life-long learning

"Life-long learning" and "life-long education" are inseparably two concepts. Both of them are proposed and evolved in the basis of human social change, under the background of human society transformation from industrial economy to knowledge economy. Life-long education thought that education should be throughout the beginning and end of life. In other words, people can also accept education via kinds of ways after the completion of compulsory education or basic education. It breaks through the framework of formal school education, regards education as a continuous process in personal life, and implements a uniform from preschool to old age throughout the educational process.

Compared with "life-long education", "life-long learning" emphasizes learner-centered, focuses on the diverse needs of learners. From the perspective of individual learners, learning is not only the need of earning a living, but also is the personal development need. It has become an essential part of one's life. 1st Global Conference of Life-long Learning proposed that life-long learning is the 21st survival concept (Longworth & Davies, 1994). People obtain knowledge, change behavior, and develop ability by continuous learning. Promoting life-long learning is an important means to deal with the world of social development needs.

U-learning refers to education is no longer limited by time and space, realizing "anywhere and anytime learning" from a technical point of view. Life-long learning emphasizes a type of learning concept, and the u-learning focuses on the learning support from technology. But they have consistency in concept hierarchy – the meaning of continuous learning. U-learning is more flexible and practical that learners can accord their characteristics and needs to choose the most suitable learning for their own, and the learning time, place, content and ways are all personal. In some sense, u-learning makes it possible to access learning network and get learning service anytime and anywhere. It provided a new learning model for life-long learning, adapted to the needs of life-long learning, is also expected to be a strong support for the built of the learning society.

The concept of the “learning society” emerged with two authors in the 1960s and 1970s. It is closely related to life-long learning and life-long education. After much debate, UNESCO (2005:60) offered the following as a view of the learning society:

Thus, learning, as a phenomenon may generalize at all levels of our societies and offer a model for organizing the time, work and lives of our institutions. Such an evolution illustrates a paradigm shift. On the one hand, education and learning can no longer be confined to a set and settled space-time, but may develop over a lifetime. On the other hand, the human actor must be put at the heart of the continuing process of knowledge acquisition and communication.

Therefore, building a learning society and cultivating life-long learning ability are necessary conditions for implementing life-long education.

2.2.2 National policies of life-long learning

Life-long learning theory has brought profound changes in the field of education; u-learning has given new idea for it. In order to put forward life-long learning and the construction of a learning society, government has been drawn up many polices.

Currently, China is using IT to propel modernization and industrialization, promoting life-long education actively, thereby constructing the life-long learning system. In 2002, *the 16th National Congress of the Communist Party of China* clearly put forward that constructing a learning society would be an important goal of building well-off society in an all-around way in the first 20 years of this century. Subsequently, the national Ministry of Education had further made explicit strategic plan of "the education information construction projects": promoting community education actively, constructing life-long learning system rapidly, and facilitating learning society in the paper of "*2003-2007 Year Education Promotion Motion Plan*". In 2010, *the Outline of China’s National Medium-and Long-Term Program for Education Reform and Development (2012-2020) (the Education Reform and Development Outline)* proposed “focusing on the construction of Chinese education information public service system and learning society support service system”, in order to realize rapid development of China’s education. After that, the Ministry of Education authorized *China’s Ten-Year Plan for Educational Informatization(2011-2020)* to confirm the tasks of *the Education Reform and Development Outline*. Meanwhile, the *Ten-year Plan* asked to build flexible, personalized IT learning environment which can meet learning needs of different groups and provide high-quality educational resources for everyone.

Under the guidance of central government policies, local governments at various levels have developed a series of policies in response to this trend. Shanghai as China's largest city, with a cosmopolitan, open and strategic vision, on behalf of China's most advanced development. As early as in 1991, Shanghai had already proposed to create a learning society and to practice life-long learning. In 2003, Shanghai Propaganda Department, Information Commission and other two departments jointly issued *Opinions on Promoting the City's Community and Cultural Information Integrated Services Construction*. Combing with the degree of Shanghai Educational Informatization, this issue asked to use modern information technology to construct educational public service platforms throughout the city, and to motivate community informatization. In 2006, Shanghai multiple government published *Guiding Opinions on Promoting the Construction of a Learning Society* on the basis of Shanghai’s realities of situation. Two goals are proposed in this document: initially set up "Anybody can Learn

Anytime and Anywhere" framework of learning society; built diverse, multi-leveled, accessible and open life-long learning system which provides formal education and leisure culture education. *The Outline of Shanghai Medium-and Long-Term Program for Education Reform and Development (2010-2020)* is another typical policy that is proposed in 2010. This policy is aimed at to promoting and integrating high-quality educational resources, improving the learning infrastructure and service system, building IT public service platforms”. And this aimed at establishing 21st century city u-learning environment, providing personalized learning service for learners; thereby bring about leapfrog development in the education informatization.

3. Methodology

As guided by national policies, China is building the learning society, and setting up public learning platforms is an important means and guarantee to realize it. Relying on abundant learning resources and learner-centered design, good learning platforms can provide rich learning experience for learners, and stimulate learning motivation. The life-long learning, u-learning and relative polices all emphasize the personal learning support system of learning platforms.

This paper synthesizes learners’ motivation and advantages of u-learning, and analyses the construction of learning platforms with ARCS model. The ARCS model is based upon the macro theory of motivation and instructional design developed by Keller (1979, 1983). Because of the boring content and short learning support, distance learning had high drop-out and low completion rates. The ARCS model was intended to change this phenomenon and stimulate learner’s motivation originally. Although many changes have taken place, the ARCS model is significant in learning design.

The ARCS model defines four major conditions (Attention, Relevance, Confidence, and Satisfaction) that have to be met for people to become and remain motivated.

3.1 Attention

The first condition, attention, is an element of motivation and is also a prerequisite for learning. The motivational concern is for getting and sustaining attention. Although stimulate and sustain motivation refers to many other aspects and through the whole learning process, drawing and sustaining learners’ attention are also the first and the most important step.

3.2 Relevance

Relevance is closely related to the learners. It may not come from the learning content; it can come from something is taught. According to expectations and value theory, the relevance of learners’ learning objectives and learning content can determine their learning depth. To the extent that a course of instruction offers opportunities for an individual to satisfy these and other needs, the person will have a feeling of perceived relevance.

3.3 Confidence

Some people never quite achieve success even when the odds are in their favor; others always seem to excel through no matter what the odds. Differences in confidence, the third major component of the

ARCS model, can influence a student's persistence and accomplishment.

3.4 Satisfaction

This category incorporates research and practices that help make people feel good about their accomplishments. According to reinforcement theory, people should be more motivated if the task and the reward are defined, and an appropriate reinforcement schedules used. There are several factors that contribute to one's level of confidence, such as feedback, rewards and evaluation.

In summary, these four categories form the basis of the ARCS model. It is always further improved with the development of education informatization. In this paper, we use this model to analysis the construction of a learning platform – Shanghai Life-long Learning Network.

4. Analysis and Results

In 2000, Shanghai began to construct "ten learning websites and one educational resource center" according to the trend of education informatization. Following the goal of initially establishing "Anybody can Learn Anytime and Anywhere" framework of learning society "Anyone learns at Anywhere and Anytime" by 2010, Shanghai started to build digital life-long learning system on the basis of "ten learning websites and one educational resource center". Shanghai Life-long Learning Network as an important part of this system has the most extensive study objects and the most comprehensive learning resources.

Shanghai Life-long Learning Network is a highly interactive educational website which opens to people of all educational levels and backgrounds within Shanghai province. The network provides people the opportunity of accessing a vast range of high- quality educational resources, and embodies an integration of multiple functions such as courseware search, learning, exchange, testing and evaluation. It is a large-scaled online learning platform that can help users effectively learn, investigate, manage and collect a great variety of information and materials. Figure 1 is the using flow chart of Shanghai Life-long Learning Network.

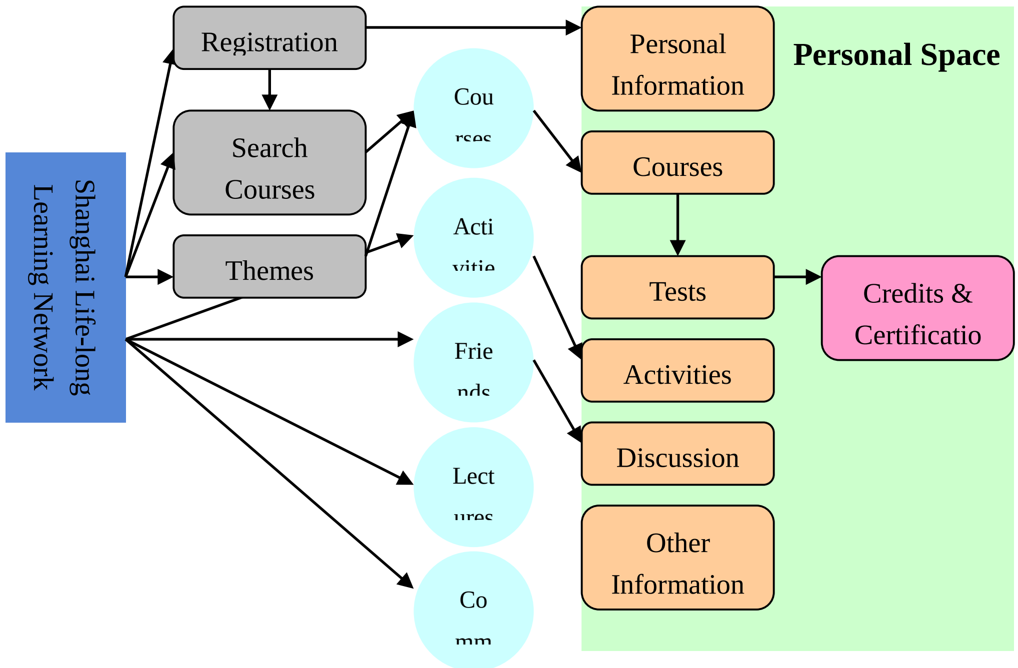


Figure1. The Using Flow Chart of Shanghai Life-long Learning Network

4.1 Courses Learning

Shanghai Life-long Learning Network provides a vast range of learning resources. People can use key words, contents or the various guidance of the platform to find the courses they want. The rich learning resources cover basic education, higher education, vocational education and life-long education, including linguistic literacy, information technology, career guides, and other various themes. Learners can choose the modes by themselves to learn the courses they appreciate, such as text reading, video browsing, audio lectures, three screen courseware, and micro learning.

Besides, there are “activities”, “reading”, “learning community” learning modules to provide learners with a wide range of popular e-books and interactive activities in the Shanghai Life-long Learning Network. For example, 2014 Shanghai community family talent show will be hold recently.

Additionally, the personalized resource recommendation system in Shanghai Life-long Learning Network can recommend high-quality courses with high evaluation by similar users on the basis of the learners’ learning records, vocation, and interests as the figure 2 shows. The personalized resource recommendation system is more convenient for users to check the courses, which saves a lot of browsing and course finding time to a large degree and allows more users to be involved in the Life-long Learning Network.

Rich learning resources and activities will attract learners’ attention, and make it easier for learners to acquire relevant knowledge and satisfy their basic motives. And the recommendation system could help learners find relevant courses. These designs meet the “Attention” and “Relevant” requirements of ARCS model.

Figure2. The Personalized Resource Recommendation

4.2 Learning Process Management

As figure 3 shows, learners are always being remained or encouraged by learning process management during their learning process. By comparing the personal learning process to the common, Shanghai Life-long Learning Network help advanced learners to build self-confidence and encourage backward learners to do much better. Learning process management is an important way to sustain learners' attention and help them gain confidence.

Figure 3. Learning Process Management

4.3 Personal Learning Space

Personal learning space is to allow learners easily master and manage their personal information. With chased curriculums presented in this space, it records all the information of learners in Shanghai Life-long Learning Network, including learning partners, online activities, discussion, and tests and so on. This information can set up a personal file for learners and record their learning activities. It can also help attract their attention to the study and remind the learning situation.

Figure 4. Personal Learning Space

4.4 Learning Interaction

In the learning process, Shanghai Life-long Learning Network provides an "Activities-Forum-Lecture" three-dimensional interactive learning system for learners to confirm the learning success. Learners can join into the course forum to discuss with other students as soon as they choose it. And this can help them find their partners with same interests. Activities are the good way for learners to present their learning outcomes and share something with others. Lectures give opportunities for learners to communication with experts, and let them answer confusion posed by learners. Learning interaction can not only urge learners to complete courses from each other, also can bring confidence and satisfaction for learners.

4.5 Learning Motivation

In order to sustain learners' attention and desire their learning satisfaction, Shanghai Life-long Learning Networks established the "Credits-Certificates-Titles" the integrated motivation system as figure 5 shows. Users will win credits and certificates through learning. The accumulated credits can also be used to change in-kind incentives.

Figure 5. Motivation system of "Credits-Certificates-Titles"

In addition, Shanghai Life-long Learning Network also provides for each learner a personal learning passport shown in Figure 6. Personal Learning Passport records history and learning outcomes of learners, demonstrates learners' awards in it. Learners can also access other people's learning passport, motivate each other and learn together.

Figure 6. Personal Learning Passport

According to ARCS model, Shanghai Lifelong Education Network can successfully attract and sustain learners' attention. Its rich learning resources and various presentation ways makes it easier to find relevant and interesting knowledge. It also has a complete record and motivate system to help learners build up confidence and feel satisfaction.

Since the official opening up of the Shanghai Lifelong Education Network on the 14th of April, 2009, the number of hits to the website within the first two weeks has reached 136,000 and the free online registration has reached 13, 2113. It is a successful social practice that supported by China's national policies.

5. Conclusion

The development of life-long learning in u-learning environment has attracted a great deal of attention in recent years that policies are proposed and infrastructure is constructed. Under the guidance of national policies, China has made efforts for the learning society to give life-long learners good learning environment. As the economy center of China, Shanghai is the largest city with the most advanced technologies and ideas. The construction of Shanghai Life-long Learning Network is a good example. Such platform reflects China's policies, meanwhile, has the potential to inform new policy-making and provides makers a vision towards drawing up better policies.

Acknowledgements

This paper is part of "Shu Guang" project which is supported by Shanghai Municipal Education Commission and Shanghai Education Development Foundation(13SG56). Besides thanks for the support of Shanghai Science and Technology Commission scientific research project "Shanghai Engineering Research Centre of Open Distance Education" (13DZ2252200).

Reference

Saadiah Yahya, Erny Arniza Ahmad & Kamarularifin Abd Jalil (2010). The definition and characteristics of ubiquitous learning: A discussion. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, 6(1), 117-127.

Bomsdor,f B. (2005). Adaptation of learning spaces: supporting ubiquitous learning in higher distance education. Retrieved from http://drops.dagstuhl.de/opus/volltexte/2005/371/pdf/05181.BomsdorfBirgit.Paper.371.pdf

Mark Weiser. (1991). The computer for the 21st century. Retrieved from http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html

Casey, D. (2005). *u-Learning = e-Learning + m-Learning*. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005, 2864-2871.

The Ministry of Education of the People's Republic of China (2004). 2003-2007 Year Education Promotion Motion Plan.(In Chinese)

The Ministry of Education of the People's Republic of China (2010). The Outline of China's National Medium-and Long-Term Program for Education Reform and Development (2012-2020). (In Chinese)

The Ministry of Education of the People's Republic of China (2011). China's Ten-Year Plan for Educational Informatization(2011-2020). (In Chinese)

Shanghai Propaganda Department, Information Commission et.al. (2003)Opinions on Promoting the City's Community and Cultural Information Integrated Services Construction. (In Chinese)

Shanghai Multiple Government (2006). Guiding Opinions on Promoting the Construction of a Learning Society. (In Chinese)

Shanghai Multiple Government (2010). The Outline of Shanghai Medium-and Long-Term Program for Education Reform and Development (2010-2020). (In Chinese)

Jiefang Daily(2007). To build a learning society with the framework of "Anybody can Learn Anytime and Anywhere" in 2010 in Shanghai. Retrieved from http: //www. jasq. cn/show. aspx?id=403&cid=39. (In Chinese)

Fu Daoming, Xu Fuyin (2007).Ubiquitous learning in pervasive computing environment. *China Educational Technology*, 24(6): 94-98. (In Chinese)

Phonic Social Network Software Scaffolds Language Learning in Ubiquitous Learning Environment

Huawen WANG, Jing LENG & Xiaoqing GU*
Department of Educational Information Technology, East China Normal University, China
*xqgu@ses.ecnu.edu.cn

Abstract: Many kinds of applications were developed to support learners in ubiquitous learning environment. Among that, social networking software (SNS) plays an increasingly important role in various educational settings. In this paper, we integrated a phonic SNS *papa* into language leaning in ubiquitous learning environment. The research investigated the general usages of *papa* and examined that to what extent the use of *papa* could effectively improve language learning. The results indicated that the use of *papa* as a SNS does promote the language learning among lower age students. The ability to speak or talk is also improved.

Keywords: Ubiquitous learning, phonic social network, language learning

1. Introduction

Ubiquitous Learning (u-Learning) refers to a learning model in which the learners can learn anytime and anywhere with the aid of mobile devices and wireless communication (Phumeechanya & Wannapiroon, 2014). The u-learning environment is supported by embedded and invisible mobile devices in daily life (Ogata & Yano, 2003), allowing students to immerse themselves fully into different learning situations (Chin & Chen, 2013) and to interact with others at any time (Weiser, Gold, & Brown, 1999).

With the popularity of mobile devices, many kinds of applications were developed. Most of u-learning applications were extended from ubiquitous computing projects and later were focused on language learning systems to support teaching and learning activities (Chin & Chen, 2013) such as Japanese Polite Expressions Learning Assisting System (JAPELAS), Japanese Mimicry and Onomatopoeia Learning Assisting System (JAMIOLAS) and Language-learning Outside the Classroom with Handhelds (LOCH) (Yin et al.,2004; Yahya, Ahmad, & Jalil, 2010). Among u-learning applications, social networking software (SNS) plays an increasingly important role in various educational settings. Educators are exploring how SNS can be used as teaching and learning tools.

Nevertheless, many related researches focused on English as a foreign language (EFL) or second language learning (SLL) and almost all the information was represented on the social network by text and pictures. This study tries to investigate the potential use of phonic social networking application and how it improves language learning on ubiquitous learning environment. More specifically, the research questions are:

1. What are students' general practices or uses of *papa*?
2. To what extent the use of phonic SNS could effectively improve language learning?

2. Literature review

SNS, including communication tools and interactive tools, provides the basis for community driven content and social networking. Introducing SNS into classroom and other educational settings results in further empowerment of learners in a new way of communicating, collaborating, and interacting, which has been recognized as potentially powerful enabling tools for educational use. Learning through SNS is advocated by pedagogical theories such as authentic learning. There are four types of learning to ensure authentic learning through SNS: action, situated, incidental and experimental learning (Liu, 2009). Action learning is a social process of solving the difficulties, by involving that learners are doing things and thinking about what they are doing as well as a practical process where students learn by doing, by observing and imitating the experts, and by getting feedback from teachers and their friends. Situated learning is similar to action learning because learners are sent to school-like settings to learn and understand new concepts and theories to be applied later on in practice. Knowledge is developed through the authentic activities and important social interactions. Cognitive apprenticeship methods try to enculturate students into authentic practices through activities and social interaction in a way similar to that evident in craft apprenticeship. Incidental learning in education contributes to unintentional or unplanned learning that results from other activities (Kerka, 2000) through observation, repetition, social interaction, and problem solving (Rogers, 1997). Knowledge from incidental learning develops self-confidence and increases self-knowledge in learning. Incidental learning usually happens in the process of completing tasks using computers (Cahoon, 1996) and/or in the online environment (McFerrin, 1999). One form of the experiential learning is outdoor education which means an outdoor program to apply their new learning during an outdoor experience upon returning to the job in order to gain more insights through challenging activities. Learners integrate thoughts and actions with reflection from the outdoor experiences.

Compared to learning language specifically vocabulary by listening, talking and reading, learning words from abstract definitions and sentences taken out of the context of normal use, which is how vocabulary has often been taught, is slow and generally unsuccessful (Miller & Gildea, 1987). As people generally learn words and practice speaking in the context of ordinary communication by expressing his/her feeling through the words s/he has just learned, it is necessary to provide the context to learners for better language learning. A number of studies examine the use of multimedia tools or online SNS for language teaching and learning, particularly studies on reading comprehension and vocabulary learning. Research of Kabilan et al. indicates that university students consider Facebook as a useful and meaningful learning tool to support, enhance and strengthen English language learning (Kabilan, Ahmad, & Abidin, 2010). Aydin (2014) explored the level of learners' interactions with their teachers on Facebook when learning English as a foreign language and found that students preferred passive behaviors regarding their interactions with their teachers. Arnold & Paulus (2010) integrated Ning, a publicly available SNS, into a blended course. The outcomes showed that the site effectively served as an information repository and the blogs and discussion forums promoted reflection and review of each other's work. HJEnglish, the most successful second language learning social network in China, engaged users into real English context in synchronous and asynchronous ways.

In a nutshell, SNS has positive effects and can be used on language learning. However, a limited number of studies focus on the first language learning through SNS. For lower age learners, their literacy is little. Being unable to type on the SNS is a great barrier to communication and language learning. The birth of phonic SNS provides a new tool of language learning for lower age learners.

3. Methodology

3.1 Participants

The participants included one teacher, 43 first grade students of 6 or 7 years old from a primary school in Shanghai. All the parents of the 43 students are well-educated and they can tutor their children to use digital products properly. The teacher, a female who was about 26 years old, has a 2-year working experience as a literacy teacher. She was keen on computer-assisted instruction and often integrates learning tools into the class activities. In this research, she interacted with students through *papa*.

3.2 Research Design

In this research, we chose *papa* as the tool to enhance language leaning. In addition to the common functions of posting, following, sharing, comment, like, and/or adding to favorites of SNS, *papa* also has recording function. After capturing photos of the real-life contexts, the learners can upload them to *papa* and record descriptions for the photos. The 43 students were assigned to use *papa* on iPad during the whole fall semester in 2013.

According to the new National Curriculum Standard of China, first language learning in primary school emphasizes (1) words, vocabularies, sentences and articles; (2) emotional attitude and values; and (3) communication and social skills. In order to achieve those goals, Huang X. H., a primary school teacher agreed that students should seize every opportunity to learn language though (a) textbooks, (b) extracurricular books, (c) peopled around, and (4) the internet. Based on the requirements for first language learning, the participating teacher led the lesson design with the researchers' guidance and support. The research consisted of the following main activities:

- Students practiced describing the reading materials the teacher provided once a couple of days.
- Students took photos they were interested in in real-life contexts. They then uploaded the pictures to *papa* and constructed verbal sentences with the newly acquired expressions and idioms.

Both the teacher and students could comment, share, like, and/or add others' posts to favorites.

3.3 Data collection and analysis

The students were numbered from 1 to 43 randomly. At the end of the semester, contents from participants waere transcribed from recordings on *papa* into words in Excel for analysis. We calculated the quantitative data to know students' general use of *papa*. Qualitative analysis came from scaling process of the posts to know to what extent the use of phonic SNS could effectively improve language learning. Likert scale of 5 points was used to scale the content of post. The scaling scheme was adapted from YaoYing's research about lower grade students' language competence (YAO & SHI, 2013). The scheme consisted of two parts, (1) Practice reading the provided materials and (2) Describe your own photos. As for the first one, the teacher usually posts photos of a textbook page and assigns students to practice reading the sentences, paragraphs or articles. Each part had three dimensions to evaluate students' language competence. The scaling process included eliminating posts that were not related to the designed activities and scaling by two assistant researchers according to the scale scheme. Before they started, they got approximately two hours of training to understand the scale and elaborate on the scale process.

4. Findings and Discussion

4.1 Survey results

At the beginning of the semester, a survey was conducted among all the students' parents to learn whether the family owned any iPads and whether they would like to encourage their children to learn on iPad. 43 students whose parents had iPads and were very glad to share them with their children were chosen to compose an experimental class which was assigned to do this research. The parents were willing to engage in supervising the students' study out of class and ensuring their kids to acquire the basic skills to operate iPads. Once there were students who had problems using iPads, the teacher or even the parents would offer 1:1 help.

4.2 Students' general practices or uses of papa

Research question 1: What are students' general practices or uses of papa?

Profile page, "following" and posts are common features of SNS. In this research, students' profile pages were personalized, with all of the 43 students uploading photos and 27 students changing the background of their profiles. All participants in this research followed each other's account on papa. As for "practicing reading the provided materials", a student posted the records of reading together with the task requirements written on a picture. After the other students saw the picture and listened to the records, they would either click "like" or give some comments. The teacher would also listen to all the posts, and if necessary she would give positive reviews or point out some mistakes so that they could be corrected next time. As for "describe your own photos", students could post any photos of beautiful sceneries or an interesting stories they took together with their records of photo description and wait for other people's comments. In this way, the teacher and students could communicate with each other about their works asynchronously. The examples are as follow:

Teacher comments

to Student 13: You used a perfect words to describe the fruitfulness in autumn. But next time please speak slowly.

to Student 24: Good reading! I can see your improvement! But "琴" is pronounced as "qin" not "qing".

Student 7 comments to Student 34: Don't laugh when telling a story!

Student 16 comments to Student 28: How can you think of the metaphor?

An obvious interaction among participants on papa is clicking "like". The data indicated that everybody got "like" and some students were much more active than others, as Student 20 and Student 7, got most "like" (3.5 and 3.2 "like" respectively), and Student 5 and Student 2 got least "like" (0.7and 0.3 "like" respectively). Everybody is engaged in interaction on *papa*.

Except clicking "like", *papa* achieved a higher level of visible student-to-teacher interaction compared to student-to-student interaction because the teacher has to trace the students' post and know the students' progress. It's her responsibility to correct the errors or mistakes of students in time. The teacher takes on roles as organizers, prompters, participants, counselors or investigators.

4.3 Papa as a SNS that facilitates language learning

Research question 2: To what extent the use of phonic SNS could effectively improve language learning?

To figure out to what extent the use of phonic SNS could positively improve language learning, we calculated the average scores in term of month from the aspect of context, language skill and emotional attitude. Overall, students had a slow and steady improvement during the 5 months according to the values in figure 1. As a phonic social networking software, *papa* can transmit not only information through words and pictures but also emotion through voice and sounds. Thus at the beginning of the semester, the young students achieved lower scores in emotional attitude but improved greatly after the activity of “practice reading the provided materials” though papa. For the content and language skill, practice is the key influential factor. Through practicing, the student minimum skipped and additive words so as to read more fluently with fewer stutters and pauses.

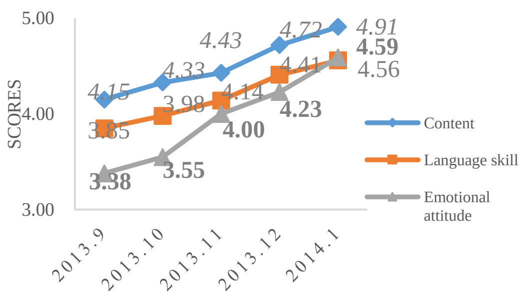


Figure 1. Scores of Reading Practice

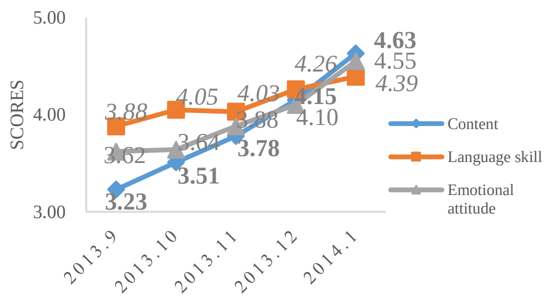


Figure 2. Scores of Photo Describing

We also calculated the average scores in term of month from the aspect of context, language skill and emotional attitude. On the whole, the scores indicated that students had an obvious and steady improvement during the 5 months according to the values in figure 2, especially the “content” and the “emotional attitude”. Whenever the students in a real-life context pertained to the newly learned phrases or idioms, sentences or stories are constructed with them and the stimulated passion and emotion were real. As the teacher said, the social practice also contributes to the content and language skill for the teacher and companions can correct it.

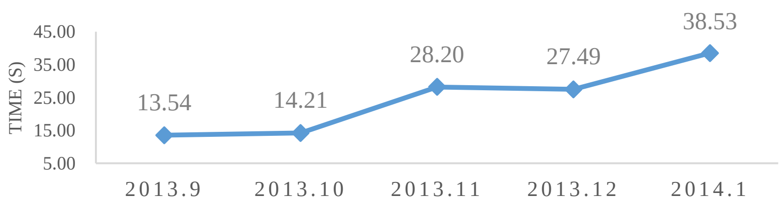


Figure 3. Average time of “describe your own photos”

In the research there is also an unintended outcome which is showed in figure 3. The average time of “describe your own photos” is gradually increasing from 13.54 seconds in Sep. 2013 to 28.2 seconds in Nov. 2013 and peaks at 38.53 seconds in Jan. 2014, except a slight drop in Dec. 2013. The trend resembles to earlier research findings of Fillmore (1983) who defined “good language learner” as a student who is talkative, eager to communicate with anyone, highly verbal and had mouth that seemed to operate non-stop around the clock. Therefore, if students are willing to say or they do have something to say, at least they have the opportunity to become “good language learner” .We can judge that the students’ speaking and talking ability are improved in this research.

5. Conclusion

Based on the findings it can be concluded that *papa* has the potential to be used as a language learning tool in ubiquitous learning environment in line with the current trend and the use of *papa* as a SNS does promote the language learning among lower age students. In interpreting the results of this study, we must pay attention to a number of limitations. One is that tradition language learning can also improve students' competence. Therefore the future research should focus on the controlled trial which the controlled group do not use *papa* to examine if *papa* do make improvement in language learning. Beyond that the interactions between students is also worthy to investigate.

Acknowledgements

We would like to thank the participants of this study for allowing us to use their data and sharing their experiences with us. We are also grateful for the reviewers' valuable comments.

References:

Arnold, N., & Paulus, T. (2010). Using a social networking site for experiential learning: Appropriating, lurking, modeling and community building. *The Internet and Higher Education, 13*(4), 188-196.

Aydin, S. (2014). Foreign language learners' interactions with their teachers on Facebook. *System, 42*, 155-163.

Cahoon, B. B. (1996). Computer skill learning in the workplace: A comparative case study.

Chengjiu Yin, Hiroaki Ogata, and Yoneo Yano, JAPELAS: Supporting Japanese Polite Expressions Learning Using PDA towards Ubiquitous Learning, The Journal of Information and Systems in Education, Vol.3, No.1, pp.33-39, 2004.

Chin, K., & Chen, Y. (2013). A mobile learning support system for ubiquitous learning environments. *Procedia-Social and Behavioral Sciences, 73*, 14-21.

Fillmore, L. W. (1983). The language learner as an individual: Implications of research on individual differences for the ESL teacher. *DOCUMENT RESUME ED 228 890 FL 013 679*, 157.

Folstein, M. (1999). MANIA, AGITATION AND ALZHEIMER DISEASE. *The American Journal of Geriatric Psychiatry, 7*, 5.

Kabilan, M. K., Ahmad, N., & Abidin, M. J. Z. (2010). Facebook: An online environment for learning of English in institutions of higher education? *The Internet and Higher Education, 13*(4), 179-187.

Kerka, S. (2000). Incidental Learning. Trends and Issues Alert No. 18.

Liu, T. Y. (2009). A context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning, 25*(6), 515-527.

McFerrin, K. M. (1999). Incidental Learning in a Higher Education Asynchronous Online Distance Education Course..

Miller, G. A., & Gildea, P. M. (1987). *How children learn words*: na.

Ogata, H., & Yano, Y. (2003). How ubiquitous computing can support language learning. *Proceedings of KEST*, 1-6.

Phumeechanya, N., & Wannapiroon, P. (2014). Design of Problem-based with Scaffolding Learning Activities in Ubiquitous Learning Environment to Develop Problem-solving Skills. *Procedia-Social and Behavioral Sciences, 116*, 4803-4808.

Rogers, A. (1997). Learning: Can We Change the Discourse?. *Adults Learning (England), 8*(5), 116-117.

Weiser, M., Gold, R., & Brown, J. S. (1999). The origins of ubiquitous computing research at PARC in the late 1980s. *IBM Systems Journal [H.W. Wilson - AST], 38*(4), 693.

Yahya, S., Ahmad, E. A., & Jalil, K. A. (2010, 2010-01-01). *The definition and characteristics of ubiquitous learning: A discussion*.

YAO, Y., & SHI, Y. (2013). An Empirical Analysis of Test of Language Expression Ability of Students in the Lower Grades of Primary Schools. *Applied Linguistics, 3*, 14.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

We are going to the ZOO! Virtual Badges in Formal out-of-school 1:1 Learning Journey with Smartphones

Ivica BOTICKI[a*], Jelena BAKSA[a], Peter SEOW[b] & Chee-Kit LOOI[b]
[a]*University of Zagreb, Faculty of electrical engineering and computing, Croatia*
[b]*National Institute of Education, Singapore*
*ivica.boticki@fer.hr

Abstract: This paper presents experiences from a seamless mobile learning project in Singapore. Although the project included a variety of seamless mobile learning designs, this paper focuses on only one, and that is a mobile learning application SamEx in support of a specific learning scenario – an outdoor ZOO field trip. The paper describes SamEx design by focusing on virtual badges gathered by the students during their ZOO trip. The trip was structured by the teachers and scaffolded by SamEx system contextually triggered questions and prompts. The paper describes experiences from the ZOO trip done with SamEx, gives an elaborate example of a student's learning experience during the trip, and concludes by examining different types of student profiles according to their badge usage and SamEx social engagement.

Keywords: virtual badges, 1:1 mobile learning, out-of-school learning, field trips

1. Introduction

This study presents and examines SamEx, a mobile learning system used by 350 students in a variety of formal and informal learning scenarios in a primary school in Singapore. Students use SamEx to capture media such as pictures, video clips and audio recordings and share them with their peers through discussions.

Although SamEx has been used in a variety of formal and informal learning designs in the project, the focus of this paper is its usage during a mobile learning field trip to Singapore ZOO (Figure 1). 350 students equipped with smartphones and SamEx system embarked on this journey to learn as designed by their teachers. Prior to the trip, scaffolds were set by the teachers and researchers in the form of contextually triggered questions and prompts. By moving through the ZOO, students were prompted to engage in a variety of tasks, collect photos, videos and audios and share them with their peers through SamEx.

As the trip progressed, students were collecting points in return for their digital content submission, question answers, mobility and social interaction. These points were materialized in the form of virtual badges, depicted in one of the SamEx screens.

2. Virtual Badges in Technology Enhanced Learning

2.1 Theoretical Background

Coming from the computer gaming world, badges are earned to indicate the achievement of certain level of skills, acquisition of knowledge, or participation in an activity (Young, 2012). As one implementation option, badges indicate the achieved competence level as defined by the issuer. For example, the integration of badges into existing software is supported by the Mozilla Open Badge Infrastructure (Mozilla, 2013). In the social media context they have five social psychological functions: goal setting, instruction, reputation, status/affirmation, and group identification (Antin &

Churchill, 2011). Thus, they have proven useful in applications which traditionally lack credit systems, such as web sites like Huffington Post and TripAdvisor reward community effort content moderation via digital badges.

Badges are nowadays integrated into numerous educational learning tools (Moore, 2013; Sharples et al., 2013), including Khan Academy, BuzzMath and CodeAcademy. However, there are still doubts on whether and how badge scores contribute to the overall student grade in online learning environments (Hakulinen, Auvinen, & Korhonen, 2013). One study shows that ability and motivation of learners have to be considered when choosing the right kind of badges to be used and the kinds of effect they could have on critical learner motivations (Abramovich, Schunn, & Higashi, 2013). TRAKLA2 confirms that and states more research is needed in balancing the badge achievement criteria so that they maximize beneficial learning practices while minimizing harmful side effects; and to understand why the same set of badges had different effects on different populations (Hakulinen et al., 2013).

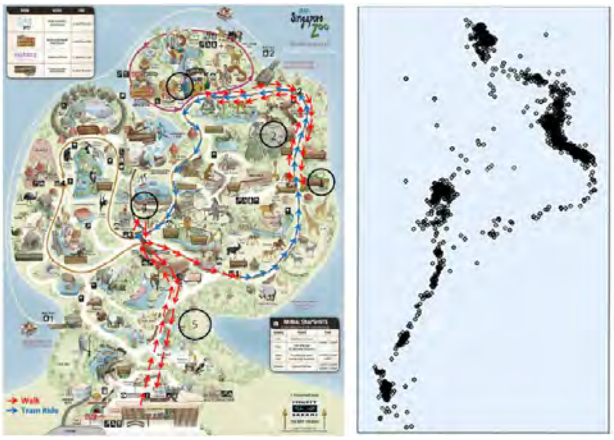


Figure 1. Route of the ZOO trip (left) and students' contributions (right)

3. Virtual Badges in Technology Enhanced Learning

3.1 Theoretical Background

Coming from the computer gaming world, badges are earned to indicate the achievement of certain level of skills, acquisition of knowledge, or participation in an activity (Young, 2012). As one implementation option, badges indicate the achieved competence level as defined by the issuer. For example, the integration of badges into existing software is supported by the Mozilla Open Badge Infrastructure (Mozilla, 2013). In the social media context they have five social psychological functions: goal setting, instruction, reputation, status/affirmation, and group identification (Antin & Churchill, 2011). Thus, they have proven useful in applications which traditionally lack credit systems, such as web sites like Huffington Post and TripAdvisor reward community effort content moderation via digital badges.

Badges are nowadays integrated into numerous educational learning tools (Moore, 2013; Sharples et al., 2013), including Khan Academy, BuzzMath and CodeAcademy. However, there are still doubts on whether and how badge scores contribute to the overall student grade in online learning environments (Hakulinen, Auvinen, & Korhonen, 2013). One study shows that ability and motivation of learners have to be considered when choosing the right kind of badges to be used and the kinds of effect they could have on critical learner motivations (Abramovich, Schunn, & Higashi, 2013). TRAKLA2 confirms that and states more research is needed in balancing the badge achievement criteria so that they maximize beneficial learning practices while minimizing harmful side effects; and to understand why the same set of badges had different effects on different populations (Hakulinen et al., 2013).

3.2 SamEx Mobile Learning Application

SamEx was designed to support self-directed and collaborative learning activities and provides a participatory platform for students to contribute, share, and give feedback. Students can use it to take a picture to collect data or post information they found to be useful for their learning. These postings are shared with other students who can review, give comments and evaluate by giving “Likes” to the contribution.

For the purpose of this study, activities were designed for primary school students who used SamEx throughout a 1-year period. In addition to collecting, storing and accessing multimedia artifacts (Figure 1), SamEx can store contextual users’ information for potential educational use. Depending on the current time and users’ location, the system allows question prompts (Figure 3) to be displayed on students’ smartphones potentially facilitating or scaffolding learning tasks. Students can therefore be guided in outdoor mobile learning trails or just prompted periodically in connection with their homework observations or other work they are recommended or required to pursue outside school.

Figure 1. Media capture in SamEx

Figure 2. Digital badges in SamEx

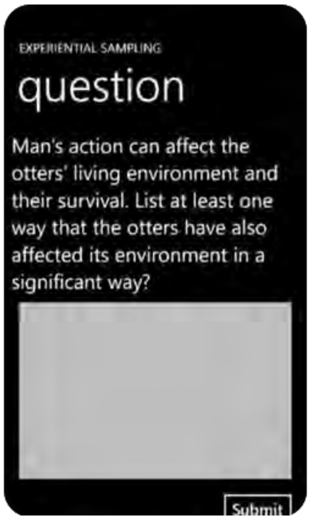


Figure 3. A location-aware question prompt

To reward students’ activity, SamEx leverages on its own badge system, an extrinsic motivational tool (Figure 2**Error! Reference source not found.**). By collecting media, answering location-aware questions, providing comments to other students’ questions and liking other students’ work, students take part in a game to accumulate points leading to the earning of badges.

4. The Analysis Framework for the Study

Throughout our studies, we discovered our students belong to the four fundamental groups according to badge system usage: (1) Badge Hunters, (2) Sharers, (3) Dodgers and (4) Explorers. The students were classified in one of the four groups by performing qualitative and quantitative analysis of their media artifacts, answers, comments and likes. We based our decision by closely observing the behaviour patterns for each student in our data sample (introduced in the following chapter). It is important to mention that there were some borderline cases where a particular student could be placed in two different categories.

5. A Primary School Mobile Learning ZOO Trip

We focus on a whole grade level of primary grade students who are equipped with smartphones with unlimited internet data plans. There are more than 350 students who were given a mobile device with SamEx mobile application preinstalled and preconfigured for use in and out of school. The study employs Design-Based Research (DBR) to develop a deeper understanding of the processes involved in implementing seamless mobile learning. With iterative cycles of studying the processes and outcomes of interventions in building teacher capacity, lesson and technology design, we can refine the processes to develop a program for designing technology enhanced learning environments and develop strategies in and out of the classroom (Phillips, 2006). The phases of the DBR approach along with the initial observation and findings are listed in Table 1.

Table 1. Use of SamEx though five main phases of Design-Based Research

Phase	Research/activity design
Phase1A (Pilot phase in the end 2012)	Study a naturalistic process of using SamEx.
This Phase 1B (Pilot phase in the end 2012)	Study a naturalistic process of using SamEx.
Phase 2 (Jan 2013)	Study a naturalistic process of using SamEx. Incorporated badges in SamEx but did not inform the students.
Phase 3 (Feb 2013)	**320 Primary 3 students used SamEx in a combination of indoor and outdoor environment in the Zoo to learn about animals and plants.**
Phase 4 (Feb-May 2013)	Study a naturalistic process of using SamEx to document students' self-directed use of SamEx
Phase 5 (June 2013)	Over a 4-week holiday, the students were assigned a task to grow a seed-

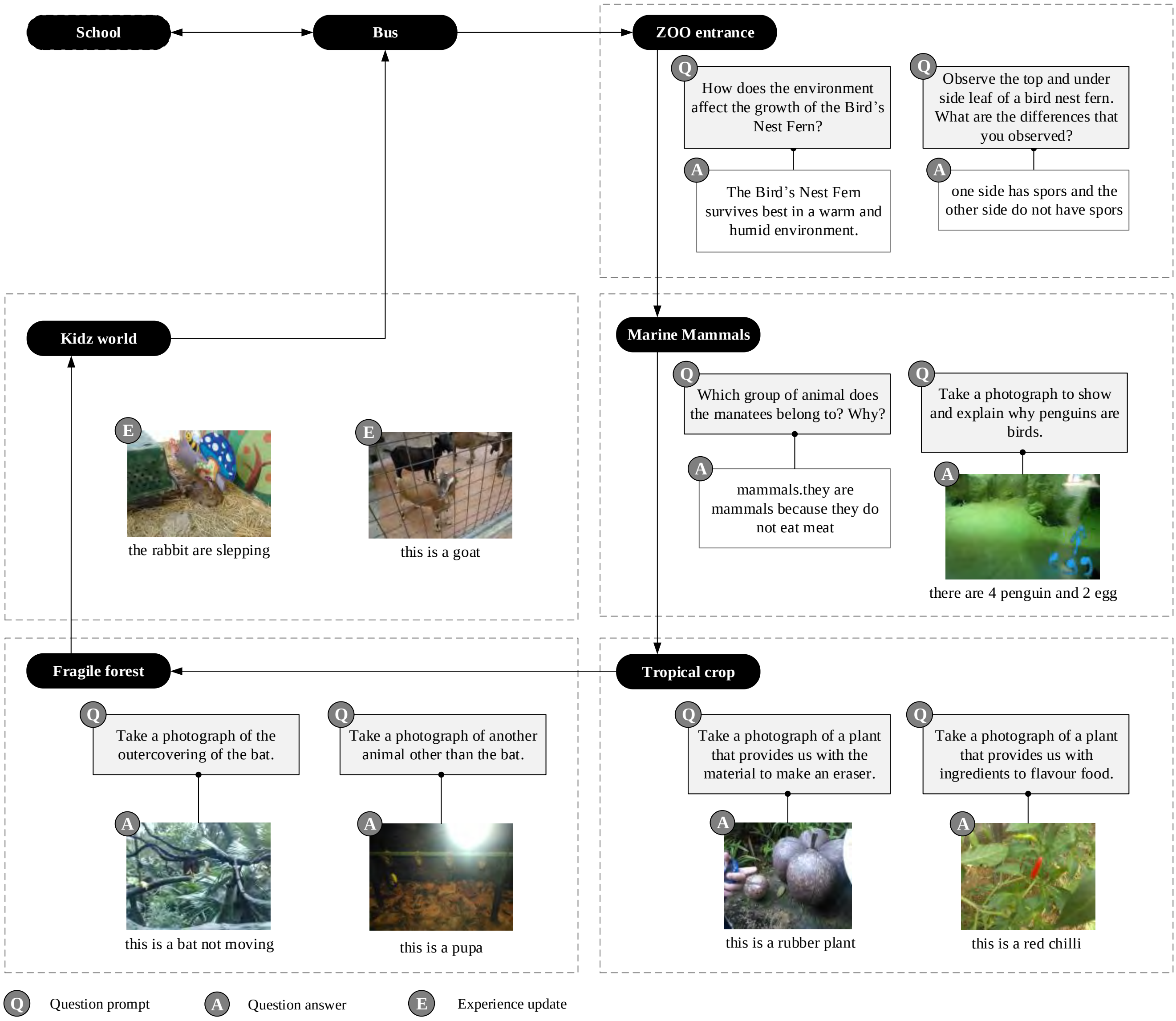


Figure 4. SamEx used in a ZOO trip: a sample of contributions and answers by one student

6. Reflections and Conclusions

Badge Hunters were identified by very a large amount of low quality data over a short period of time. They are only interested in attaining high levels of badges and they only respond to extrinsic motivation and do not care about quality of contributions. Sharers are on the other hand interested in sharing with their peers while earning their badges and their participation consists of higher quality contributions. They make meaningful contributions and ask good questions. Dodgers are not interested in earning badges at all.

Table 2. Percentage of students in each category for class P3A

Category	%
Sharers	42.86
Dodgers	35.71
Badge hunters	16.67
Explorers	4.76

Unfortunately, a very low percentage of students (Table 2) was placed in the ideal category of Explorers. They actively participate in SamEx by generating high quality contributions, sharing their observations, initiating conversations with other students and are trying to gain knowledge collaboratively from their peers.

This means badges can currently only encourage the first two groups of students to participate. However, Badge Hunters will stop participating once they achieve their desired level of badges. Both Badge Hunters and Sharers are not interested in learning collaboratively since there is no observable learning with their peers.

Nevertheless, it is important to mention that some of the students who were identified as Sharers have a good potential of becoming Explorers. They usually try to learn collaboratively, only to be discouraged by the lack of feedback from their classmates. To help them bridge that gap between Shares and Explorers, all students need to be encouraged to participate more actively, especially in the tasks which involve interaction with their peers

References

Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education?: it depends upon the type of badge and expertise of learner. *Etr&D-Educational Technology Research and Development*, *61*(2), 217–232. doi:10.1007/s11423-013-9289-2

Antin, J., & Churchill, E. (2011). Badges in social media: A social psychological perspective. In *CHI 2011* (pp. 1–4). Retrieved from http://uxscientist.com/public/docs/uxsci_2.pdf\nhttp://uxscientist.com/?sort=post_date&page=7

Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment. In *Proceedings of the 2013 Learning and Teaching in Computing and Engineering Conference* (pp. 47–54). IEEE. doi:10.1109/LaTiCE.2013.34

Moore, M. G. (2013). Independent Learning, MOOCs, and the Open Badges Infrastructure. *American Journal of Distance Education*, pp. 75–76. doi:10.1080/08923647.2013.786935

Mozilla. (2013). *Open Badges*. Retrieved from http://openbadges.org/about/

Phillips, D. (2006). Assessing the quality of design research proposals. *Educational Design Research*, 93–99.

Sharples, M., McAndrew, P., Weller, M., Ferguson, R., FitzGerald, E., Hirst, T., … Whitelock, D. (2013). *Innovating Pedagogy*. The Open University UK.

Young, J. R. (2012). “ Badges” Earned Online Pose Challenge to Traditional College Diplomas. *The Chronicle of Higher Education*, pp. 1–7. Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ1002829

Building an Online Collaborative Learning Community in Ubiquitous Learning Environment

Ru ZHANG [a*], Jing LENG [b], Xiaoqing GU [c], Guanfeng FU [d], Huawen WANG[f]
[a,b,c,d,f] *Department of Education Information Technology, East China Normal University, China*
*ruzhang88@foxmail.com

Abstract: The paper aims at understanding to what extent the learners could engage in online collaborative discussion and generate sustained, productive knowledge building discourse in ubiquitous learning environment in a social context. Online discourse data in an inquiry case and a survey responses were analyzed to examine learners' effort for idea improvement and attitudes towards online building community respectively. The findings indicate that in ubiquitous learning environment discussions are strengthened by social networks and the collaborative learning community can be built on the social interactions and collective knowledge advancement by generating productive knowledge building discourse.

Keywords: Ubiquitous learning, online asynchronous discussion, collaborative learning community, knowledge building

1. Introduction

With ever increasing advanced capabilities of personal devices and popularity of novel web services, more learners have engaged in online collaboration learning to build shared understanding through social networking and computer-mediated communication, in addition to searching for required information. Rather than just acting as a repository for online reading, online asynchronous discussion forum has become an effective avenue for collaborative inquiry through identifying, examining and reflecting upon ideas (Collison, Elbaum, Haavind & Tinker, 2000). Moreover, it enables a collaborative space for the creation of positive collaboration learning environment and community culture (Wei, 2013; Xia, Fielder, & Siragusa, 2013), where a wide range of learners engage in an inquiry about a problem and collective idea is improved through the progressive collaborative inquiry from the perspective of knowledge building. However, not all the discussions is helpful in creating and sustaining an online community of inquiry. Advances in Internet and ubiquitous computing enable a new means for building collaborative learning community through synchronous discussion where learners could use multiple devices to access learning materials and to support their learning activities, anytime and anywhere.

The paper aims at understanding to what extent the learners could engage in online collaborative discussion and generate sustained, productive knowledge building discourse in ubiquitous learning environment in a social context. Following an overall description of the asynchronous discussion platform, Content analysis on an inquiry case in *Zhihu* was to examine learners' effort for idea improvement during the process of collaborative discussion. A survey responses were also analyzed to present learners' attitudes towards the formation and sustainment of learning community. This paper also tries to enlighten researchers and practitioners on effective ways for promoting learners' active participation in online discussion and facilitating the collaborative knowledge building processes. Specifically, the research questions are as follow:

1. How do learners engage in the discussion in ubiquitous learning environment?
2. During online asynchronous discussion activities, what are the characteristics of the learners' behavior contributing to idea improvement in the social context?
3. What are the learners' attitudes toward building online communities through online asynchronous discussion?

2. Literature Review

2.1 Online collaborative learning

Collaboration is commonly used to encourage learners' engagement and enhance their understanding through collaborative inquiry in online learning environments. In the past several decades, there has been an upsurge in research on computer-supported collaborative learning (CSCL). To attempt to initiate learners into a knowledge creating culture, researchers have developed the conceptual framework of knowledge building in which learners work together through CSCL to advance collective knowledge and understanding (Scardamalia & Bereiter, 2006). Knowledge building is defined as "the production and continual improvement of ideas of value to a community" (Scardamalia & Bereiter, 2003), which focuses on the social process of collaborative inquiry in asynchronous discussion environment where learners post ideas and comments to generate questions, co-construct explanations, and revise their ideas. In the process, learners' ideas are regarded as conceptual artefacts of inquiry to be discussed, inter-connected, revised, and superseded for achieving idea improvement (Stahl, 2006). Such collective idea improvement is of great value to online collaborative learning community of inquiry where learners collaboratively build upon each other's knowledge, resulting in the creation or modification of community knowledge through discourse and discussion.

2.2 Online learning environments

Asynchronous discussion forums have been used in a wide range of formal and informal learning setting. Despite the generalized BBS in daily lives, typical instructional management software environments, like Blackboard, Moodle and WebCT are also available to asynchronous discussion forum. It is considered an extension of instructional practices that promotes dialogue, reflection, knowledge construction and self-assessment (Gao et al., 2013). Although many researchers regard online Q&A discussion forums as a valuable learning resource (e.g., Cheng, Liu, & Shieh, 2012), the quality of online discussion and communication were questioned. Thomas (2002) argued that they “might not be the best technology to support the interactive and collaborative processes essential to a conversational model of learning”. The main problem is that they may fail to foster naturally productive discussions that are focused, interactive and in-depth (Gao et al., 2013), thus failing to promote collaborative learning (Guzdial & Turns, 2000) and militating against deepening inquiry (Scardamalia & Bereiter, 2006). Because the isolation among the learners in asynchronous discussion, can hinder them from interacting and building knowledge together, discussions can be invaluable toward creating and sustaining an online community of inquiry (DeNoyelles et al., 2014). In view of this, it’s need to integrate the emerging technologies to design online learning environments that have the capability to facilitate complex learning and sustain a strong sense of community that supports students’ learning socially and cognitively (DeNoyelles et al., 2014; Gao et al., 2013; Wei & Chen, 2006). Some systems are designed or introduced to enable collaborative learning through discussion forums. Guzdial and Turns (2000) designed a computer-mediated anchored forum to increase the effectiveness of sustained, broad participation and on-topic discussions by using anchors. An anchor is a topic that students find worthy of discussion in the forum, but the anchors are static and not initiated by students. Lin, Hou et al.(2013) used a popular SNS, Facebook, to support students’ asynchronous online discussions of project-based learning activities and found that the essential social nature of Facebook play an important role. In the same way, LeNoue et al. (2011) suggest that a well-designed social networking site can provide probability for diversified opportunities for learners by facilitating both broad and deep interactions.

2.3 Using ubiquitous computing tools to support discussion activity

The rapid developments of wireless networking technology and powerful mobile devices have led to the landscape of ubiquitous computing as Weiser proposed, and have given rise to the emergence of ubiquitous learning environments where learners can become completely immersed in the learning process or learning activity in any situation (Jones & Jo, 2004). Supported by proper technological affordances and activity structures in ubiquitous learning environments, learners can easily access rich resources and actively participate in learning activities anytime and anywhere through any device (Jones & Jo, 2004; Wei & Chen, 2006). Thus we unconsciously and effortlessly harness their digital abilities as effort-saving strategies for achieving the benefits of distributed intelligence (Pea & Maldonado, 2006). Ubiquitous learning environment was featured by Yahya et al. (2010) in the following characteristics: 1) Permanency, 2) Accessibility, 3) Immediacy, 4) Interactivity, 5) Context-awareness, and 6) Adaptability. Previous studies have proved the great potentiality of ubiquitous computing tools in improving discussion activities accordingly. Researchers have indicated that using mobile devices to support discussion activities may promote acquisition and exchange of authentic learning experience, improve collaborative interaction encourage active learning (Lan, Tsai, Yang, & Hung, 2012; Wei & Chen, 2006). For instance, Yang and Lin (2010) integrated tablet computers with personal digital assistants (PDAs) together to facilitate collaborative discussions and information sharing, and improve learners’ abilities on classifying plants. Wong et al.(2010) presented a seamless learning environment with mobile device where learners can use smart phones to capture photos of real-life contexts pertaining to Chinese idioms, and engage in in-class or out-of-class sharing and discussions in personal and social learning spaces. Therefore, various types of discussions and learning activities are enabled in ubiquitous learning environment.

3. Methodology

This research employed the method of content analysis to analyze an inquiry case and survey response, to explore the formation and sustainment of a community of inquiry by examining the process and motivation about collaborative knowledge building in asynchronous discussion forum, *Zhihu*.

3.1 Research context

Zhihu, is a well-known open online question and answer discussion forum based on social networking in China, where many authentic, complex and novel problems are being discussed by the public. It emphasizes on posting problems and seeking help from peers, facilitating collaborative thinking, and building a community. Meanwhile, a mobile App called *Zhihu Daily* has been designed to present dozens of valuable contents generated in the process of discussion. According to *Alexa* (a web traffic tracking subsidiary of amazon, http://www.alexa.cn/) , the number of Daily Pageviews per Visitor in *Zhihu* (http://www.*Zhihu* .com/) and *Zhihu Daily* (http://www.*Zhihu* daily.net/) is higher than similar Q&A sites as *Baidu Knows* (http://zhidao.baidu.com/) and *IASK* (http://iask.sina.com.cn/), see in table 1. It means that users in *Zhihu* tends to be more engaged and deep thinking in the discussions.

Table 1: the number of Daily Pageviews per Visitor in August 11, 2014.

	Baidu Knows	*IASK*	*Zhihu*	*Zhihu Daily*
Daily Pageviews per Visitor	4.21	1.45	5.75	6.1

Zhihu and *Zhihu Daily* can be accessed using a web browser or by using the mobile app, which helps learners to easily engage in a community of inquiry about a problem via different digital devices (e.g. smart phones) anytime and anywhere. *Zhihu* has adopted typical features of search, endorse, wiki and follow in social networking. Thus as an asynchronous discussion forum, its features enable learners engage in collaborative inquiries to solve authentic and in-depth problems in a more interactive way, such as raising questions for inquiry about a problem; clarifying questions by revising anyone’s question; endorsing and commenting quality answers to actively discuss and construct deep explanation; following up an answer for a discussion in a new thread through commenting; searching and following related problems or inquiry topic, proper learners to engage in the inquiry and a community of inquiry. It turns out that diverse ideas are generated and exchanged, learners engage in sustained progressive discourse to revise, combine, synthesize and produce rise-above ideas to advance the collective knowledge in the online platform. Gradually, a knowledge-building community has been built and their social networks have been reinforced through continuously interaction, thus contributing to effective knowledge work in the community space. In this study, *Zhihu* is described as a ubiquitous learning environment for its nature of interaction and community of inquiry, where learners can easily get immersed in the inquiry about a contextual problem and achieve the sustainment of community by means of collaborative effort.

Figure 1. Features of the platform *Zhihu*.

3.2 Data collection and analysis

Our data collection was comprised two parts: notes in an inquiry and a survey in *Zhihu*. Content analysis were employed to analyze online interaction in the inquiry and survey response. In the study, a typical scientific inquiry case in *Zhihu* were analyzed to explore how learners engage and advance knowledge in discussion. And the problem in the inquiry is:

Birds'「Knees」bend backward, human knees bend forward, why did they evolve so differently?

The problem was posted on December 30 2013, but the inquiry on it were lasted until now, along with 500 learners following the inquiry. For this study, all the responses (including questions, answers, and comments) involved in the inquiry were collected. Thus we collected a total of 260 notes on the day of August 13 2014, among which 221 notes were comments to quality or interesting answers, excluding the notes without any obvious meaning. But not all the notes were address the problem, some were about social interaction within the community. So we divides each note to idea unit—the smallest unit of text that conveyed a distinct idea or meaning, and each unit were coded into the coding categories (shown in Table 2.).

To understand deeply the motivation of sustainment in the discussion, we analyze a survey "Why did you participate in discussions in *Zhihu*?" posted in *Zhihu*. There are over 300 users participate by answering, commenting and endorsing on the responses to it. We collected 73 responses and used qualitative coding techniques, to code and analyze the responses. We then categorized the codes to explore learners' attitudes and explain the sustainment of building online community in asynchronous discussion.

Table 2. Coding categories

Categories	Description
progressive inquiry	Raising questions or making explanations to the problem
Social interaction	Complimenting, appreciating the answer, greeting, connecting or sharing with other learners and expressing emotions

4. Results and discussion

4.1 How did learners engage in addressing problem in the discussion?

Although the inquiry was lasted about one year, the discussions were continued with active participation and timely responses. In the inquiry, the ideas generated in the inquiry received 2448 endorsements. And once a theory or personal idea was posted, learners can be quickly informed to response to the idea. The results of coding showed most discourse activities in the inquiry(65%) were related to community building through social interaction, such as complimenting, expressing emotions, appreciating the answer, greeting, connecting or sharing with other learners using the features of social networking(e.g. @ sb.) and so on. 38% of notes were devoted to extending or deepening the inquiry, such as continuing a thread by arguing or quoting from other's messages, referring explicitly to other messages and raising a new discussion thread, stating learners' personal ideas. Progressive inquiry was evident in the notes addressing the problem that were characterized in a network of learners' interaction represented in Figure 1. The network revealed how the learners set forth their ideas and question, and addressed increasingly complex problems. And the ideas were improved in extended discourse about reviewing progress, identifying the synthesis, rise-above new ideas and lines for the collaboration inquiry. For example, the learner (XKC) answer made an explanation on bird's bones of leg, leading other learners to the understanding that birds are digitigrade. Based on this understanding, learners generated further problems and statements of what they needed to know in comments, such as: "If dogs' knees bend backwards"? (by CG). They could also comment on an existing thread to build up the argument related to the discussion thread, such as the learner (LL).

4.2 How did learners change their understanding within their discourse space?

In the inquiry, learners used the public space to invite peer input, and to generate potential solutions on the problem. After the first question about the problem of birds' knees was posted, questions on it were clarified and refined constantly in seven times by modifying or raising new questions. And in the process, the problem was clarified from a misunderstanding to generate a series of good questions, such as "Look carefully, you will find that (the joint bending backwards) is not the knees, but the ankle joint."[by ZH], "Why would humans followed with their heel?" [by YY], "Comparing to the way of plantigrade, what's the advantage in way of digitigrade? ”[by YY] .These results indicate that learners changed their understanding of birds’ knees from a misunderstanding towards a more scientific view.

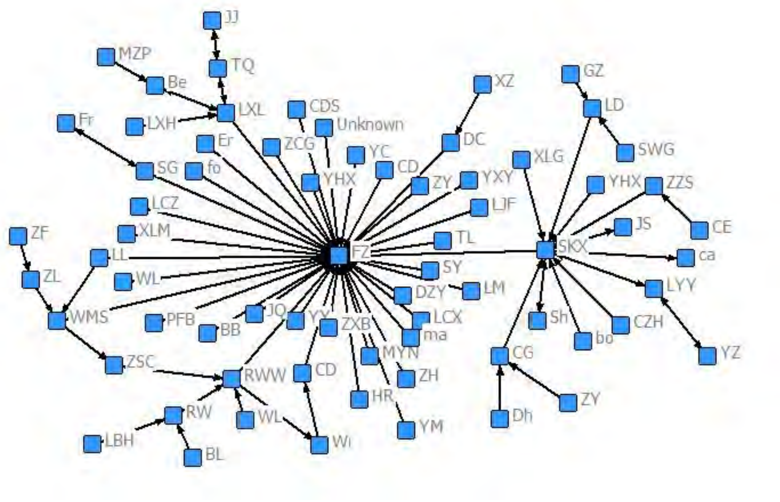


Figure 1. Network of learners’ interaction in the inquiry.

Table 3 reports the number of notes (a) raising factual problems that’s seeking for factual information for peers to address; (b) raising explanatory problems that’s explanations about “why”, “how” and so on. Empirical data such as experimental results, personal observations, and life experiences were found in 16 notes (64%) among all the explanation in the main discussion thread. According to the two patterns using empirical data proposed by Zhang et al. (2007), empirical data can be used in descriptive or explanatory approach.10 notes were a description of experiments, observations, or experiences without elaboration of ideas, other 6 notes were integrated empirical data to justify and improve learners’ ideas and received most endorsements and comments during the inquiry, that is productive to the knowledge building community (Zhang et al., 2007). So there’re 5 discussions thread continued and 11 new discussion threads raised in these comments altogether in addition to these notes to examine and interpret these evidence critically. Meanwhile, other expert resources were introduced as evidences to support their ideas, which are helpful to deeper understanding in the community. These results indicate how learners changed their understanding of birds’ knees from an observed phenomena in real lives towards a more scientific view.

Table 3: Number of problems in the inquiry.

	Number of notes raising factual questions	Number of notes raising explanatory questions
Total in the inquiry	11	5
Total in comments	9	2

4.3 What are the learners’ attitudes toward building online communities?

What is most essential to the sustainment of online learning community to advance collective knowledge like *Zhihu*, is the participation of its learners. What motivates learners to essentially provide knowledge services for others without any explicit rewards and teachers’ instruction?

Most of our understanding of learners' attitudes comes from the survey "Why did you participate in discussions in *Zhihu?*" in it. In the survey, most learners gave multiple reasons. By far, the most often mentioned reasons were enjoy helping, learning, hobby, which is in accordance with prior research (Lou, Fang, Lim, & Peng, 2013) expectedly. Thus a platform supporting ubiquitous learning can be characterized from their survey responses to explain the community building. We will discuss each in below.

4.3.1 Supporting collaborative inquiry on authentic problems

Although various novel internet technologies have emerged, the argument that "it is difficult to find a place where you would feel funny in the Internet" was common in the survey responses. For learners in *Zhihu*, it's a place that can "acquire knowledge in reality along with novel ideas" to meet curiosity and interests, thus getting immersed. Authentic problem held some of learners' learning motivations and provided question sources for inquiry from a cognitive perspective. It's important implications for developing a community where learners can pursue their own questions and contribute ideas for inquiry to build on ideas, concepts and explanations about problem through discussion in society. So the community supporting the collaborative work of improving ideas on authentic problem is critical. Regarding *Zhihu*, Said one respondent:

> *When there is a related problem in front of you, you can express and examine your ideas (in Zhihu).*

The problems contribute greatly to collaborative inquiry, as a respondent stated:

> *At the beginning (I) contributed a lot of answers to the questions relating my own profession. It's not only to let others understand the area more comprehensively, but also to tease my knowledge better. Then the problems in this area were almost answered. But I found I know little about other area…… Participating in Zhihu can be considered as an inspiration for me to read more books and to explore more.*

4.3.2 Distributed Expertise

A progressive inquiry intends to engage the community in a shared process of knowledge advancement and to convey simultaneously the cognitive goals for collaboration (Muukkonen, Hakkarainen, & Lakkala, 2004). Not only delivering tasks or productions on time, employing socially distributed cognitive resource in the inquiry process is high valued by many respondents in the survey. With the help of social networking, learners can easily assemble right experts to engage in or deepen the inquiry and even generate new knowledge by introducing empirical data and expert resources to build social knowledge network and advance collective knowledge. Diversity in expertise among participants, and interaction with expert cultures, promotes knowledge advancement (Muukkonen, Hakkarainen, & Lakkala, 2004). So benefiting from the distributed expertise, the platform impresses learners that valid solution to problem is right there and it can meet their need of just-in-time learning anywhere. A respondent stated his online learning needs:

> *We need systematic fragmentation information…… There's someone answer your question systematically once question posted in Zhihu without much time……An anonymous users in it can provide awesome goods, and it's second to none.*

4.3.3 Peer help

Like-minded peers were frequently mentioned in the responses. Peer help with numerous pedagogical advantages (Wei & Chen, 2006) is of great helpful when Web-based learning environments have become quite common without enough teachers' instructions. Many respondents indicated the ability to dialogue with friends with common topics increased the activity of the discussion, and encouraged a sense of community with helping each other. Hew and Cheung (2011) believed that it's helpful for knowledge advancement when peers provided comments, showed appreciation and encouragement, and summarized the discussions. The Internet spirit of openness, equality, collaboration and sharing was used to express the cognitive value of social collaboration the cognitive value of social collaboration in a response:

> *Zhihu is open and equal to all people without any differences in professional and identity. Each questioner would need and respect the exchange of ideas and you. Collaboration: A question will be answered by all kinds of people, eventually integrate of the most appropriate answer through collaboration. Sharing: a useful answer, would pass to people in most need through other people.*

A certified public accountants stated the benefits considerably from peer help:

> *My expertise shared through Zhihu and Zhihu Daily were recognized and praised by many people including knowledgeable expert. This motivation inspired by the recognition is far more than that in other Internet platforms.*

5. Conclusion and implications

In this study, we analyzed an online collaborative inquiry in social context to examine learners' efforts for idea improvement by generation of deepening questions, social interactions focusing on idea advancement, constructive use of empirical data and authoritative sources in notes. The findings show that the social networking provide possibility for learners to continue discussions, facilitate ideas to spread and sustain community of inquiry in communal space through encouragement, drawing in participants, endorsement, appreciation. In ubiquitous learning environment, the community has been sustained by enabling learners to access inquiry process anytime and anywhere. And supporting with collaborative inquiry on authentic problems, distributed expertise, peer help, increased question generation and consequent knowledge construction can lead to a virtuous cycle within the community through productive knowledge building discourse, where learners can ask better questions, construct their own knowledge when participating in collective idea improvement.

As for the limitations in the study, there's only an inquiry case analyzed to examine the process of idea improvement. And adding more inquiry cases is need. And it is important to use more diverse methods to analyze online synchronous discussions in ubiquitous learning environment. As social learning interactions play an important role, in addition to the use of quantitative content analysis, future researchers should include methods such as social network analysis and sequential analysis.

References:

Cheng, S. S., Liu, E. Z. F., & Shieh, R. S. (2012). Identifying the Indicators Attracting Users to Online

Question and Answer Discussion Forums. Social Behavior and Personality: an international journal, 40(2), 283-292. doi: 10.2224/sbp.2012.40.2.283

Collison, G., Elbaum, B., Haavind, S. & Tinker, R. (2000).Facilitating online learning: effective strategies for moderators. Madison, WI: Atwood Publishing.

DeNoyelles, A. A. D. U., Mannheimer Zydney, J. J. Z. U., & Baiyun Chen, B. C. U. E. (2014). Strategies for Creating a Community of Inquiry through Online Asynchronous Discussions. Journal of Online Learning & Teaching, 10(1), 153-165

Gao, F., Zhang, T., & Franklin, T. (2013). Designing asynchronous online discussion environments: Recent progress and possible future directions. British Journal of Educational Technology, 44(3), 469-483. doi: 10.1111/j.1467-8535.2012.01330.x

Guzdial, M., & Turns, J. (2000). Effective Discussion through a Computer-Mediated Anchored Forum (437): Lawrence Erlbaum Associates, Inc. (Reprinted.

Jones, V., & Jo, J. H. (2004). Ubiquitous learning environment: An adaptive teaching system using ubiquitous technology. In *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (No. s 468, p. 474).

Lan, Y., Tsai, P., Yang, S., & Hung, C. (2012). Comparing the social knowledge construction behavioral patterns of problem-based online asynchronous discussion in e/m-learning environments. Computers & Education, 59(4), 1122-1135. doi: 10.1016/j.compedu.2012.05.004

Lin, P., Hou, H., Wang, S., & Chang, K. (2013). Analyzing knowledge dimensions and cognitive process of a project-based online discussion instructional activity using Facebook in an adult and continuing education course. Computers & Education, 60(1), 110-121. doi: 10.1016/j.compedu.2012.07.017

Lou, J., Fang, Y., Lim, K. H., & Peng, J. Z. (2013). Contributing high quantity and quality knowledge to online Q&A communities. Journal of the American Society for Information Science and Technology, 64(2), 356-371. doi: 10.1002/asi.22750

Muukkonen, H., Hakkarainen, K., & Lakkala, M. (2004). Computer-Mediated Progressive Inquiry in Higher Education. In T. S. Roberts (Ed.), Online Collaborative Learning : Theory and Practice (28-53). Hershey PA: Information Science Pub. (Reprinted.

Pea, R. D., & Maldonado, H. (2006). WILD for Learning: Interacting Through New Computing Devices Anytime, Anywhere. In R. K. Sawyer (Ed.), (427-441). New York, NY, US: Cambridge University Press. (Reprinted.

Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In Encyclopedia of Education (2nd Ed., p. 1370-1373). New York: Macmillan Reference, USA

Scardamalia, M., & Bereiter, C. (2006). Knowledge Building: Theory, Pedagogy, and Technology. In K. Sawyer (Ed.), Cambridge Handbook of the Learning Sciences (97-118). New York: Cambridge University Press. (Reprinted.

Stahl, G. (2006).Group cognition: Computer support for building collaborative knowledge. Cambridge, MA: The MIT Press.

Wei, F., & Chen, G. (2006). Collaborative mentor support in a learning context using a ubiquitous discussion forum to facilitate knowledge sharing for lifelong learning. British Journal of Educational Technology, 37(6), 917-935. doi: 10.1111/j.1467-8535.2006.00674.x

Wong, L., Chin, C., Tan, C., & Liu, M. (2010). Students' personal and social meaning making in a Chinese idiom mobile learning environment. Educational Technology & Society, 13(4), 15

Xia, J., Fielder, J., & Siragusa, L. (2013). Achieving better peer interaction in online discussion forums: A reflective practitioner case study. Educational Research, 23(1)

Yahya, S., Ahmad, E. A., & Jalil, K. A. (2010). The definition and characteristics of ubiquitous learning: A discussion. In *International Journal of Education and Development using Information and Communication Technology (IJEDICT*.

Yang, J. C. Œ. L. (2010). Development and evaluation of an interactive mobile learning environment with Shared Display Groupware. Educational Technology & Society, 13(1), 195

Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. Educational Technology Research and Development, 55(2), 117-145. doi: 10.1007/s11423-006-9019-0

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Identifying User's Perceptions Toward Integrating Mobile Applications in Science Education

Hyo-Jeong SO[a*], Hye-Gyoung YOON[b], Hyungshin CHOI[c], Heung-Chang LEE[a] & Kyudong PARK[a]

[a]*Creative IT Engineering, Pohang University of Science & Technology(POSTECH), Korea*
[b]*Science Education, Chuncheon National University of Education, Korea*
[c]*Computer Education, Chuncheon National University of Education, Korea*
*hyojeongso@postech.ac.kr

Abstract: Situated under the macro-context of the Korean government's initiative to provide mobile- and cloud-based learning resources in schools, the goal of this study is to examine potential users' perceptions toward integrating mobile applications in science education. The sample includes 632 students, 68 teachers and 141 parents from 11 Primary schools in Korea. Results reveal some important implications regarding users' acceptance of mobile applications for teaching and learning. First, the majority of students do not frequently access and use mobile devices, which might be associated with the school policy to ban or limit the use of mobile devices in the classroom. This finding may imply a conflict between the enacted macro-level policy and the reality of schools. Second, the teacher group is the most conservative in their perceptions toward integrating mobile applications in science education. Teachers' conservative attitude may be associated with the lack of perceived advantages of using mobile applications in teaching and learning. Lastly, there was a clear pattern observed in users' perceptions about science topics that are effective or ineffective in integrating mobile applications. Science phenomena that require a long duration of observation were perceived to be highly relevant to integrate mobile applications, supporting the continuity of observation experiences across physical contexts and time scale. Further, we found that mobile applications that provoke students' imaginations and curiosity, where students do not feel forced to learn, but are intrinsically motivated to learning science in daily life, are likely to be readily accepted and used by student users.

Keywords: mobile earning, science education, digital textbook, needs analysis

1. Introduction

For the past decade, leveraging the unique affordances of mobile devices such as connectivity, immediacy and multi-functionality, several research studies have explored the impact of mobile learning applications (Motiwalla, 2007). Despite the promises of mobile learning, however, the use and adoption of mobile devices and applications in schools have been slow, and we have not witnessed many successful examples of mobile applications or systems widely adopted in classrooms. Still, most-widely used mobile applications in educational settings focus on simple communication, presentation, and productivity functions. There is a lack of successful examples of mobile applications that have fully utilized the affordances of mobility to create a new form of learning experiences. As to the use of mobile technologies in science education, Avraamidou (2008) makes a critical remark that while several research studies presented the positive impact of mobile applications in science education, there is a critical need to "design technology-enhanced curriculum materials that can be implemented and used prevalently in the classroom" (p. 361). We concur with this remark that without widespread acceptance and use of mobile-enhanced curriculum materials among individual users, the promise of mobile learning cannot be realized.

The research project presented in this paper was part of the large-scale project to develop mobile applications that could be prevalently integrated in Primary Science curricula. At a macro-level, this study is situated under the context of the large-scale ICT in education policy by the Korean Government that has planned the development and implementation of digital textbooks in schools. The key impetus underlying the “Digital Textbook for All by 2015” initiative is to promote self-directed learning practices with the use of digital resources and content. To date, the digital textbooks have been piloted in over 100 research schools to examine design and implementation issues (MEST & KERIS, 2012).

While the Korean government’s plan is to introduce and implement digital textbooks to all schools by 2015, the majority of schools may not be ready to adopt digital textbooks. When moving from traditional printed textbooks to digital textbooks, schools may face transitory issues such as changes in teaching and learning practices, building necessary socio-technological infrastructures, and stakeholders’ adoption of digital textbooks. That is, the transition toward digital textbooks involves not only the adoption of technology-based tools, but also changes in socio-technical structures and cultural practices surrounding the use of such a new form of technologies (Bielaczyc, 2006).

Understanding such transitory issues, we explored a hybrid form of learning where mobile applications are used to support learning with printed textbooks, prior to the full adoption of digital textbooks in schools. The ultimate goal of this project was to develop a series of mobile applications that could be used in conjunction with printed textbooks in Primary Science curricula. As an initial step toward this goal, we conducted a needs analysis study to examine potential users’ perceptions, expectations, and concerns toward using mobile applications in science education. While the main user group includes primary school students, this study also included teachers and parents as important actors affecting the use and adoption of mobile applications in science education.

2. Methods

2.1 Participants

Three stakeholder groups, namely students, teachers and parents, were the participants of this study. As shown in Table 1, the sample includes 632 students, 68 teachers and 141 parents from 11 Primary schools in Korea. We employed a purposive sampling method to recruit participants in various geographical areas in Korea. First, we recruited participants from two schools that were piloting the use of digital textbooks in the science curriculum at Grades 3 and 4 levels. We purposely included the pilot schools since one of the main goals of this project was to examine the possibility of integrating mobile applications as complementary or supporting resources to the science digital textbook. Second, we purposely recruited teachers who participated as authors of the national science textbook since content knowledge and teaching experiences in a science subject area were critical to evaluate the feasibility and needs for developing mobile applications relevant to science curricula. Students and parents in the recruited teachers’ classes were invited to participate in this research via an invitation letter explaining the purpose and process of the research study.

Regarding demographic profiles of the participants, the student group includes 342 Grade 3 students (54.1%) and 290 Grade 4 students (45.9%). The gender distribution was nearly even, 330 male (52.2%) and 302 female (47.8%) students. The teacher group includes diverse age ranges: 19.1% of teachers under 30 years old, 36.8% of those age 31-40, 36.8% of those age 41-50, and 7.4% of those 51 and older. Nearly three quarters of teachers are female. The parents group includes about 60% of those age 31-40 and 40% of those age 41-50. Almost 88% of parents who responded to our survey were female.

Table 1: Survey Participants

	Digital Textbook Pilot School	Non-Pilot Schools	Total
Students	187	445	632
Teachers	51	17	68
Parents	141	-	141

2.2 Data Collection & Analysis

This study employed both survey and interviews to identify users' perceptions toward mobile application integration in the science curriculum. The research team developed three different sets of survey instrument for a respective group. We used a Likert scale of 1 (strongly disagree) to 5 (strongly agree) in the teacher survey, whereas a four-point Likert scale (no mid-point) was used in the student survey and parent survey. Common factors measured across the groups are a) usage patterns of using mobile devices, b) prior experiences of using mobile devices and mobile applications in science learning, and c) perceived expectations and concerns toward mobile learning. Specifically, factors examined in a respective group are as follows:

- *Study survey*: a) student's interests in science learning, b) mobile device usage, and c) expectation toward using mobile applications in science lessons
- *Teacher survey*: a) mobile device usage, b) prior experience of using mobile applications for teaching, c) challenges in teaching science, d) topics in the science curriculum suitable for integrating mobile applications, and e) potential concerns and issues with mobile learning
- *Parent survey*: a) mobile devices usage, b) prior experiences of using mobile applications for child's learning purposes, c) expectation toward using mobile applications in science lessons, and d) areas in need for using mobile applications at home

In addition to the survey method, we conducted interviews with students and teachers to collect qualitative data concerning the factors abovementioned. The interview sessions were conducted in different formats to accommodate the participants' needs and schedules. For the student group, we conducted four focus group interview sessions with 28 students at the digital textbook pilot schools. We also interviewed 20 teachers, including 8 teachers from the digital textbook pilot school, and 12 teachers who were the authors of the national science textbook. We used a focus group interview format with the teachers from the digital textbook pilot school, whereas the remaining teachers were tele-interviewed via phone or web-conferencing platform, due to their geographical diversity. We were not able to interview parents due to a low number of parents who volunteered to participate in an interview session.

In the student interview, we asked questions concerning student's interest, motivation and self-efficacy in science learning, prior experiences of using the science digital textbook, using mobile devices in daily life, expectation toward using mobile applications in science lessons, and ideas/topics perceived to be suitable for using mobile applications in the science curriculum. The teacher interview includes questions regarding their prior experiences teaching the science subject, prior experiences using mobile devices (including both general personal use and teaching/learning purposes), challenges with using the printed mode of textbook, and topics/activities perceived to be highly relevant to integrate mobile applications.

For data analysis, SPSS was used to statistically analyze survey data. Interview sessions were audio-recorded and transcribed. We used an open-coding method to identify key themes emerged from the interview data.

3. Results

3.1 Student Group

3.1.1 Mobile Device Usage Patterns

To understand how Primary school students use mobile devices in daily life, we examined the types of mobile devices, the frequency of usage and the types of activities with mobile devices. It was found that while students typically use smartphones more frequently than other mobile devices such as tablet PCs and portable game players, the overall usage of mobile devices is not high. One-quarter of students reported that they do not use smartphones at all. Nearly two-thirds of students use smartphones less than two hours per day. Top three activities that students do with mobile devices are: 1) playing games (26.3%), 2) communication (24.5%), and 3) watching video clips (18.8%).

3.1.2 Perceived Expectations

Next, we examined the students' perceived expectations about using mobile applications in science lessons. Overall, the students have positive perceptions of using mobile applications together with the printed textbook. Nearly 80% of students reported that it is not difficult to learn how to operate mobile devices, which imply that young users at this age do not perceive technical difficulties as a hindrance to learning with mobile devices. They perceived that science lessons could be more fun (89.4%), easier to learn (86.5%), and include more diverse types of activities (92.2%), with the use of mobile applications. During the focus group interview, we also observed students' positive attitude toward using mobile applications in science lessons. Students mentioned that with the use of mobile applications, "lessons could be more lively", "It would be easy to conduct dangerous experiments", and "learning with mobile applications would be more fun than learning with the traditional textbook". However, about 40% of students expressed some concern that lessons could be distracting with the use of mobile applications.

3.1.3 Relative Advantages: Digital Textbook vs. Printed Textbook

We examined how students perceived relative advantages of digital textbooks and printed textbooks. The students mainly mentioned multimedia features such as "we can watch video clips" and "I don't need to take notes" as advantages of using a digital textbook. Limitations of digital textbooks mostly involve technical issues (e.g., frequent error message, difficult to use) and health issues (e.g., affecting eyes and vision). In addition, as a way to compare and contrast digital vs. printed textbooks, we asked the students to generate ideas about science learning activities that are often limited in the traditional printed textbook, but could be better learned with a mobile application. Three distinctive categories of mobile application emerged from student-generated ideas. We categorized them according to the functional framework of mobile applications by Patten et al. (2006):

- *Referential*: scientific term search, learning scientific phenomenon or terms in daily life
- *Data collection*: observation of living and non-living things, raising a virtual pet like Tamagochi, complex/dangerous experiment (e.g., air current, dissection, natural disaster)
- *Microworld*: exploration, discovery-oriented games (e.g. survival in a deserted island, journey to the planet)

3.2 Teacher Group

3.2.1 Mobile Device Usage Patterns

Teachers reported that they use smartphones more frequently than other types of mobile devices. However, when we asked about what device or platform they use to search for teaching-related information and resources, the majority of teachers (97%) responded that they use a computer (i.e., desktop, notebooks). In the survey, about 47% of teachers reported that they had prior experiences using mobile devices for teaching purposes. Interview data reveal that while some teachers had prior experiences using mobile devices in science lessons, most usage patterns involve simple activities like term search, taking a picture, and video recording.

3.2.2 Science Topics Suitable for Mobile Learning

Next, we examined what areas/topics in the science curriculum teachers perceived to be suitable and effective for integrating mobile applications. The question item listed all topics in the Primary 3 and 4 Science curricula, and the teachers were asked to indicate the degree of expected effectiveness for each topic, on a five point Likert scale. From this data, we were able to identify a clear pattern in the nature of science topics that were perceived to be effective or ineffective in integrating mobile applications. Topics perceived to be highly effective are *Life cycles of animals* (86.8%), *Life cycles of plants* (86.8%), *Change of earth surface* (83.8%), and *Volcano and earthquake* (83.8%), which are dealing with scientific themes or phenomena that are difficult to observe in a real-life context due to its dynamic nature and gradual changes for a long period. On the contrary, topics perceived to be ineffective in integrating mobile learning applications include *Separation of mixture* (55.8%), *Weight of object* (52.9%), *Use of magnet* (44%), and *Materials and matter* (39.7%), which tend to cover scientific phenomena easily observable and replicable in school contexts.

3.2.3 Relative Advantages: Digital Textbook vs. Printed Textbook

Among 12 teachers interviewed, three of them indicated that they did not have any difficulties using the traditional printed textbooks. It appeared that teachers often restructure and repurpose the content of textbooks to suit their needs, rather than simply follow the given format and structure. Teachers perceived limitations of printed textbooks when they a) teach topics requiring scientific experiments difficult to be conducted in school contexts, b) teach scientific phenomena involving a long duration of observation and field trips, and c) face difficulties accommodating various questions from students. Some teachers indicated that students do not completely understand instructions in textbooks when conducting an experiment. Digital content like video clips and audio materials explaining procedural steps were suggested to solve this problem.

3.3 Parent Group

3.3.1 Mobile Device Usage Patterns

There was no dominant type of devices used by parents for searching education-related information. About 46% of parents reported using a computer (i.e., desktop, notebooks) whereas about 42% of them used smartphones to search for information related to their child's education. Regarding prior experiences, nearly 76% of parents reported that they had not used any mobile application to teaching their children, indicating an overall low usage of mobile learning applications among parent users.

3.3.2 Expectations

On the whole, the parents had positive expectations about using mobile devices and applications in science lessons. Parents responded that science lessons could be more fun (89.4%), easier to learn (86.5%), and include more diverse activities (92.2%) with the use of mobile devices and applications. Regarding the question "Lessons could be distracting", the percentage of parents who agree with this item (53.95%) was slightly more than that of parents who disagree (46.1%). Similar to the student group, parents tend to perceive the use of mobile devices not technically challenging for children. About 75% of parents indicate that it is not difficult for children to operate mobile devices.

4. Discussion & Conclusion

The main goal of this research was to identify potential users' perceptions toward integrating mobile applications in Primary Science curricula. We examined perceptions of three stakeholder groups, with a consideration that not only students but also both teachers and parents are also important actors affecting the use and adoption of mobile learning applications in science education. Technology use and acceptance is a long-lasting issue that has drawn much research interest. While several researchers have examined determinants of technology adoption in educational settings (e.g., Teo, & van Schaik, 2012), this study is different in that it has a focus to the use and adoption of mobile technologies particularly in the context of science education.

Overall, this research reveals some important implications regarding users' acceptance of mobile technologies for teaching and learning. First, while access to mobile devices is the foundational condition to meet prior to the use of any mobile learning resources and application, we found that the majority of students do not frequently access and use mobile devices. This finding may be associated with the school policy to ban or limit the use of mobile devices in the classroom. Contrary to the Korean Government's plan to introduce more mobile- or cloud-based learning resources, students are not allowed to use a mobile device as a learning tool in schools, which may indicate a conflict between the enacted macro-level policy and the reality of schools.

Second, we found that the teacher group is the most conservative in their perceptions toward integrating mobile applications in science education. Unlike the student and the parent groups who generally showed positive attitudes and toward mobile learning, the teachers expressed several concerns regarding the use of both mobile applications and digital textbooks. We also found that many teachers prefer to use a computer rather than a mobile device to search for teaching-related information and resources. Teachers' conservative attitude may be associated with the lack of perceived advantages of using mobile applications in teaching and learning. This is consistent with the general Technology Acceptance Model where perceived usefulness and perceived ease of use are important factors affecting user acceptance of technology (Davis, 1989).

Lastly, this study also reveals that there are certain areas or topics in science curricula that are perceived to be more suitable and effective for integrating mobile applications than other topics. Science phenomena that require a long duration of observation were perceived to be highly relevant to integrate mobile applications, supporting *the continuity of observation experiences* across physical contexts and time scale. Curricula topics involving static phenomena that could be sufficiently covered with printed textbooks and physical objects were perceived to be less relevant to integrate mobile applications. We also found that mobile applications merely reflecting curricula content are not likely to attract student users' motivation to use them. The students generated several interesting ideas of mobile applications that could be potentially developed as promising learning tools. Using such ideas, mobile applications that provoke students' imaginations and curiosity, where students do not feel forced to learn, but are intrinsically motivated to learning science in daily life, are likely to be readily accepted and used by student users.

While the present study examined users' general perceptions, we see needs to conduct future research into mobile technology acceptance and adoption that particularly consider unique characteristics and issue associated with mobile technologies. Such research will contribute to better understanding of the complexity of users' perception and adoption of mobile devices and applications for teaching and learning (Sarker & Wells, 2003). In conclusion, we believe that this initial user study provides some important insight to how the three groups of users perceive the role of mobile technology and applications in the context of science education, and to what factors should be considered to develop mobile applications reflecting users' needs.

Acknowledgements

This research is supported by the Korea Foundation for the Advancement of Science and Creativity.

References

Avraamidou, L. (2008). Prospects for the use of mobile technologies in science education. *AACE Journal, 16*(3), 347-365.

Bielaczyc, K. (2006). Designing social infrastructure: Critical issues in creating learning environments with technology. *The Journal of the Learning Sciences, 15*(3), 301-329.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319-340.

Ministry of Education, Science and Technology & KERIS (2012). Adapting education to the information age (white paper), Retrieved on 10 August 2014 from http://english.keris.or.kr/es_ac/es_ac_210.jsp

Motiwalla, L. F. (2007). Mobile learning: A framework and evaluation. *Computers & Education*, *49*(3), 581-596.

Patten, B., Arnedillo Sánchez, I., & Tangney, B. (2006). Designing collaborative, constructionist and contextual applications for handheld devices. *Computers & Education, 46*(3), 294-308.

Sarker, S., & Wells, J. D. (2003). Understanding mobile handheld device use and adoption. *Communications of the ACM, 46*(12), 35-40.

Teo, T., & van Schaik, P. (2012). Understanding the intention to use technology by preservice teachers: An empirical test of competing theoretical models. *International Journal of Human-Computer Interaction*, *28*(3), 178-188.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Integration of multiple external representations in chemistry: a requirements-gathering study

Prajakt PANDE, Sanjay CHANDRASEKHARAN
Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, India
*prajakp@hbcse.tifr.res.in

Abstract: Multiple external representations (MERs) are crucial in the learning and practice of chemistry. Representational competence (RC), the ability to simultaneously process, integrate and transform between MERs, marks expertise in chemistry. A major strand of chemistry education research attributes students' difficulties in learning to difficulty in understanding MERs, particularly the ability to imagine the various inter-connections between them. A dominant model of RC is Johnstone's model of three thinking levels, which describes three different levels of representations in chemistry (symbolic equations, molecular models and reaction phenomena), and treats cognitive load as the core problem underlying student difficulties with MERs. This model is used to design a number of computer interventions in chemistry, mostly focusing on lowering cognitive/working memory load, by simultaneously displaying on a screen, molecular animations, graphs and equations. In contrast to this classical information processing framework, our theoretical approach seeks to understand the internal cognitive mechanisms that support the processing of MERs, using recent cognitive theories such as distributed and embodied cognition. At the intervention level, we focus on achieving integration of MERs, through enactive/embodied interaction design approaches (such as fully interconnected and manipulable interfaces). Before developing the actual form factor of the interventions, we wanted to characterize student difficulties, and how students navigated through existing MERs. For this, we presented a categorization task to students, where 3D molecular animations (depicting only molecular level reaction dynamics, without symbols, text and other representations), graphs, chemical equations and videos of some chemical reactions were given to 6 chemistry undergrad students. Eye-tracking was used to obtain fine-grained data about participants' gaze and eye movement patterns while they viewed these representations. Addition of molecules, molecular aggregation, heat source and increase in velocity of the molecules were frequently attended-to features. Only one student made chemically meaningful groups with animations. Her eye-movement analysis reveals systematic mapping of animation features to chemical equations and other representations.

Keywords: Representational competence (RC), multiple external representations (MERs), chemistry education, molecular animations, categorization, eye-tracking

1. Introduction

Chemical phenomena are understood at multiple levels of detail (electronic configuration, stereo-chemistry or spatial conformation of molecules, stoichiometric ratios etc.), using multiple external representations (MERs) such as reaction mechanisms, molecular diagrams, graphs and equations at each level. A critical aspect of learning chemistry is developing expertise over these MERs. The ability to generate and use MERs in an integrated fashion (for conceptualization, discovery and communication) is indicative of expertise in chemistry. This skill-set is collectively known as representational competence (RC) in chemistry (Kozma & Russell, 1997).

Many problems and difficulties in teaching/learning chemistry are attributed to difficulties in understanding the different MERs in chemistry (Johnstone, 1991, 1993; Kozma & Russell, 1997; Gilbert & Treagust, 2009a & b). The achievement of RC, through many representational transformations, as well as the integration of MERs, is central to learning chemistry.

Currently dominant theoretical approaches, for instance Johnstone's model of three thinking levels (Johnstone, 1982) and versions thereof, describe the different levels of MERs (such as symbolic level, molecular level, etc.) based on the type and level of detail of the chemical phenomenon represented in those representations. This model is combined with working memory models, to develop a cognitive-load-based characterization of RC in chemistry, which is used to also address student difficulties in dealing with chemistry MERs. This approach considers the simultaneous consideration of MERs by a learner as increasing her cognitive load, while an expert is better able to minimize/handle cognitive load by employing cognitive strategies such as information chunking.

Johnstone's model has inspired and guided the development and use of a number of computer interfaces in chemistry teaching/learning, typically used for making sense of MERs and also developing concept, phenomenon and procedure understanding using MERs (Kozma & Russell, 1997; Kelly & Jones, 2008). Many of these interfaces focus on the simultaneous, dynamic, display of MERs on a screen. However, the effectiveness of such computer interfaces in helping develop RC has been mixed. One possible reason for this could be that these designs are guided by classical information processing theories of cognition, where the role of the interface is to decrease the learner’s cognitive load, particularly working memory load. Emerging theories of cognition, such as distributed and embodied cognition, postulate that the roles played by external representations are wider than decreasing cognitive load. For instance, external representations can support operations that are difficult, and sometimes impossible, to do in imagination (Kirsh, 2010). Further, actions could be a way of promoting integration (Chandrasekharan, 2009).

Current characterizations of RC in chemistry, and the interventions inspired by them, do not seek to provide a detailed understanding of the cognitive mechanisms underlying the processing of MERs, and thus offer only a rather superficial account of MER integration. Our research attempts to characterize RC by developing models of the cognitive mechanisms underlying the processing of MERs, particularly integration of MERs (which is how we define RC), and suggest design principles for interventions. Our theoretical approach, as well as interaction designs, are inspired by distributed and embodied cognition perspectives. In this paper, we report findings from an ongoing requirements-gathering phase for an intervention design. For this, we presented a categorization task to students, where 3D molecular animations (depicting only molecular level reaction dynamics, without symbols, text and other representations), graphs, chemical equations and videos of some chemical reactions were given to 6 chemistry undergrad students.

We used Tobii X2-60 static eye-tracker to capture fine-grained data on student eye-movement and gaze patterns across MERs presented to (and handled by) them. This data provides a deeper understanding of how students move through the representations (see e.g. Figure 2). Our preliminary analysis confirms earlier reports on novices' surface-feature-based exploration of MERs, but adds details of eye-gaze and movement patterns. Students struggled/failed in mapping dynamic features from animations to corresponding features in equations.

2. RC characterization and investigation approaches in chemistry

A significant strand of research in chemistry education reports descriptions of students’ use of multiple representations, transformations of these representations, and the difficulties students face while doing both of the above. Studies show that students lack a clear understanding of basic concepts such as oxidation numbers, ionic charge, atoms and atomic structure, formal rules for writing molecular formulae, as well as meaning of subscript numbers and brackets and coefficients (Garforth, Johnstone & Lazonby, 1976; Savoy, 1988), and eventually fail to associate the symbols and numbers with substances and phenomena (Yarroch, 1985; Herron and Greenbowe, 1986; Nurrenbern & Pickering, 1987; Hinton & Nakhleh, 1999; Sanger & Phelps, 2007). Ben-Zvi, Eylon, & Silberstein, (1988) propose that students' thinking about phenomena relies primarily on perceptual/sensory information but since current pedagogical practices hardly provide perceptual/sensory assistance, students do not understand chemical symbols in terms of their macro and micro-level instantiations.

A more direct approach to characterize RC describes expert-novice differences in the use of MERs of chemical phenomena and their transformations (Kozma & Russell, 1997; Kozma, 2003), by using (in combination or in isolation) the influential working memory model (novices have less skills to manage cognitive load; Johnstone, 1982), context and practice (novices are less exposed to these; Ben-Zvi, Eylon, & Silberstein, 1988) and conceptual understanding (novices lack rich conceptual ground to counter cognitive load; Cook, 2006; Nitz, Nerdel & Prechtl, 2012). Kozma and Russell (2005) identified specific skills among chemistry experts, viz., (a) using representations to describe chemical phenomena, (b) generating and/or selecting appropriate MERs according to specific needs,

(c) identifying and analyzing different features of MERs, (d) comparing and contrasting different MERs, (e) making connections across different representations, relating/mapping features between MERs, (f) understanding that the MERs correspond to phenomena but are distinct from them, and (g) using MERs to support claims, draw inferences, and make predictions.

Experts also seem to better transform between static (such as equation & graphs) and dynamic representations (such as reaction mechanisms; Wu & Shah, 2004; Kelly & Jones, 2008; Nakhleh & Postek, 2008) while students face difficulties in producing static representations (e.g. sketches; Madden, Jones & Rahm, 2011) of the (imagined) dynamic particulate interactions. Understanding chemical phenomena involves building of internal (mental) models that simulate the behaviors of many individual molecules/atoms, their collective behaviors and properties (Levy & Wilensky, 2009) and effects of various parameters on such behaviors.

To improve students' conceptual understanding by linking and transforming between representations, interventions (guided by the above research and theoretical approaches) have focused on the use of computer interfaces, mostly based on the classical information processing approach to cognition, particularly Baddeley's working memory model (e.g. SMV Chem, visChem, 4M:Chem, EduChem HS, eChem, etc.). These interfaces, seek to lower the load on students' memory, by allowing learners to view multiple representations simultaneously on screen.

In contrast to this traditional memory-based approach focusing on simultaneous display, recent work, such as the Connected Chemistry Curriculum, focuses on interlinking representations through *manipulable* simulations and animations. Based on the Netlogo 2D interface, the manipulability feature may help students better transform between static and dynamic representations. The developers of this curriculum report, through control-experimental group studies, that the curriculum improves generation, handling and understanding of chemistry MERs, particularly the submicroscopic ones, among students, when compared to conventional text/lecture based curricula (Stieff & Wilensky, 2003; Stieff & McCombs, 2006). Such manipulable simulation interfaces have often been coupled with other scaffolds (such as exercises, quizzes, activities and teacher guides; Zhang & Linn, 2011) and have been effective in improving students' representations and understanding. Computer interfaces for RC assessment have also been explored, and could prove useful in characterizing RC. For instance, by examining students' use of a multi-representational molecular mechanics animation using eye-tracking, researchers show that students mainly use graphical and model representations in animations, often ignoring the equations (Stieff, Hegarty & Deslongchamps, 2011).

Despite this rich set of studies, there is no theoretical account of the cognitive mechanisms underlying RC, or efforts to develop interventions based on recent cognitive models.

3. Study and Research Questions

Here we report an ongoing requirements-gathering phase for an intervention design. The studies in this phase seek to characterize specific problems faced by students, and how they navigate through existing MERs. 3D molecular animations (depicting only molecular level reaction dynamics, without symbols, text and other representations), graphs, chemical equations and videos of some chemical reactions were given to 6 chemistry undergrad students. Our primary research question was:

How do students understand the reaction dynamics through animations and other representations? At the operational level, our questions include:

- What are the commonly/frequently attended-to features in bare/raw animations, and how are they used to establish links with other representations?
- Do students make chemically meaningful groups/correspondences between bare/raw animations and other representations?
- What are the correctly and incorrectly referred-to features in bare/raw animations, and other representations, in cases of chemically meaningful correspondences?

4. Materials, sample and methodology

Materials for the categorization experiment included, for five pre-determined general chemical reactions, bare 3D molecular animations, laboratory videos, chemical equations and graphs. We developed bare/raw 3D molecular animation for five general chemical reactions. Each animation depicts only the molecular dynamics/mechanism of that reaction, and does not have any other embedded representations, such as text, narrative, or symbols. Free and open videos of the five chemical reactions (being performed in laboratories) were procured from on-line sources. Equations and approximate graphs were generated using an image editing software. Each representation was validated for content, conceptual and representational correctness, by two chemistry experts and one cognitive science expert. Figure 1 shows the preparation and execution of the experiment in detail.

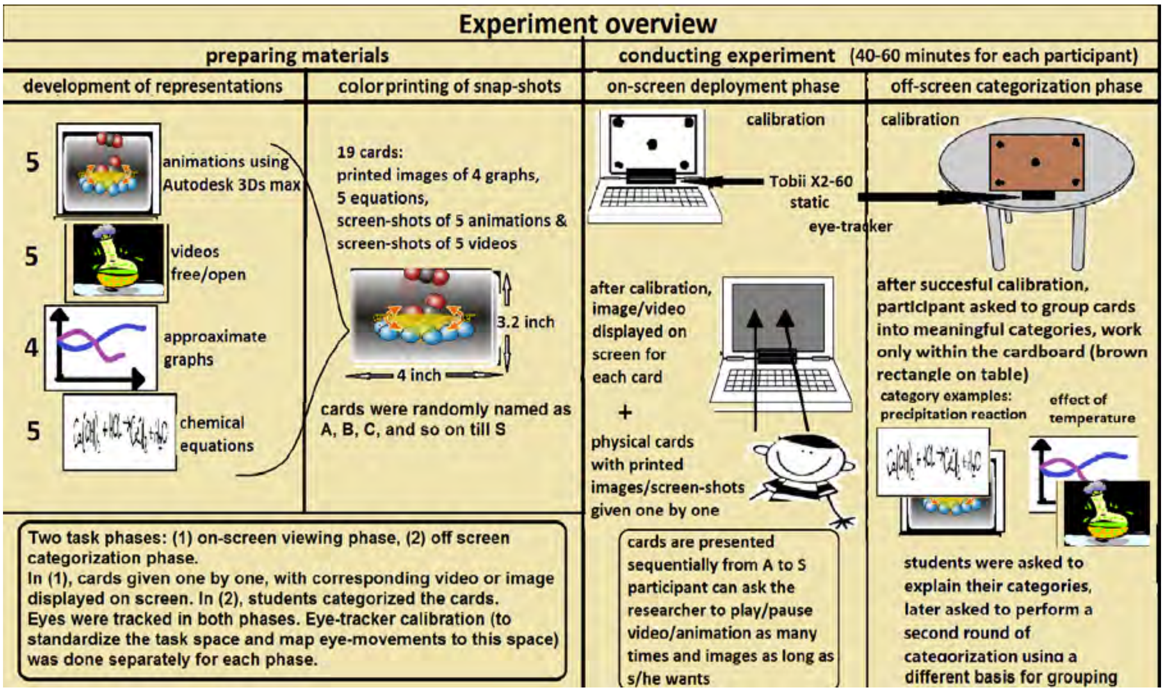


Figure 1. Material development and experimental design details

Six chemistry undergrad students (3 girls) from a nearby college participated in the categorization experiment; each of them performed the task individually.

Sources of data collection: (a) for on-screen phase – dynamic eye-movement and fixation data superimposed on the screen-capture video, (b) for off-screen phase – categories made by the participants, their verbal justifications, dynamic eye-movement and fixation data superimposed on the top-view video of the task-process, and side-view video recording of the categorization and justification sessions. The study session ranged from 40-60 minutes for each participant.

5. Findings and discussion

The number of animations used in chemically meaningful ways, and the explanations of categories (see table 1), suggest students found it difficult to understand animations.

All students, except G2, tended to map surface-features of animations to other MERs. For instance, the heat source in an animation was mapped with 'Δ' (symbol signifying 'heating') in an equation (e.g. equation for effect of temperature on NO_2 equilibrium) and/or with burner shown in videos. Addition of molecules, molecular aggregation, breaking of molecules, illumination of heat source and increase in velocity of the molecules on heating were the most frequently referred-to features in animations. Students often said, '...they (animations) all look alike...' while explaining their categories. None focused on the structure of individual molecules in the animations.

Graphs play a mediating role for linking static (equation) and dynamic (animation) content. An ideal way of examining the graph could be moving across its slope, and imagining the dynamics of the corresponding molecular behavior. None of the students' eye-movements match such a pattern, nor did any of the students group graphs in chemically meaningful ways, except for B2. However, his eye-movement patterns are not indicative of simulating molecular/reaction behavior from graphs.

Table 1: students' usage of animations in the first round of categorization (each given 5 animations)

Participant	B1	B2	B3	G1	G2	G3
Total animations used (correctly + incorrectly)	5	5	3	5	5	5
Animations used in feature-based + chemically meaningful way	5	3	2	5	4	4
Animations used in only surface-feature-based manned	5	2	2	5	1	4
Animations used in only chemically meaningful manner	0	1	0	0	3	0

G2's results were close to expert-level, so we present brief findings from a preliminary analysis of the task only for G2. She made 4 categories with the 19 representations, placing 4 out of 5 animations with 4 chemical equations (NO_2 equilibrium, precipitation, AgCl-NH_3 equilibrium and aqueous cobalt chloride equilibrium), each animation corresponding to one equation. One pair was incorrect (cobalt chloride equilibrium paired with animation representing neutralization reaction). The remaining animation was grouped with two graphs and two videos (all four representing effect of temperature on equilibrium).

While viewing the animations, G2 tended to follow only one molecule at a time, and was attentive to changes in the number molecules, and their bonding and breaking. G2 did systematic scanning of the equation, looking for molecular formulas and subscripts. Figure 2 shows one segment of G2's gaze sequence, while she viewed an equation. For instance, G2's second fixation (point No. 2) was on the subscript (aq) that denotes aqueous state of KNO_3, at fixation point 3, G2's attention is on subscript (s) that denotes the solid state of PbI_2. Immediate next fixation (point No. 4) occurs on subscript (aq), aqueous state of KI. A similar pattern is observable while G2 viewed other equations. G2 seems to trace the states of elements before and after their displacement. There are both forward (in the direction of the reaction) and backward (opposite to the direction of reactions) eye-movements. Such movements may generally be associated with sequential imagination of reaction dynamics. Her chemically meaningful categories, combining animations and equations, and her gaze pattern, suggests that she imagined the step-wise dynamics of molecular mechanisms, using the symbols in the equations. A detailed analysis of her gaze pattern, and comparison with other participants' gaze patterns, is ongoing. We expect to present details of this analysis at the workshop.

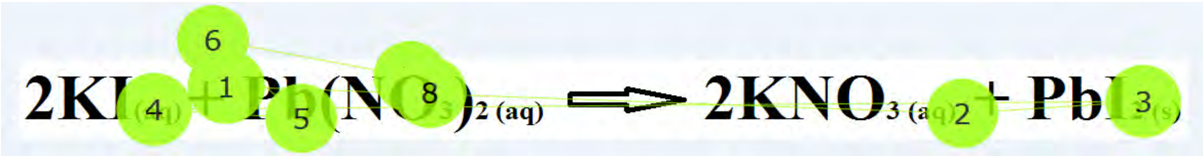


Figure 2. An instance of student G2's eye-gaze sequence

6. Conclusion

Students have difficulties in chemically relating animations to other representations such as graphs and equations. They tend to focus on surface features and ignore important dynamic features in animations. Molecular dynamics are difficult to understand, and integrating them with equations and graphs requires generating dynamic features using static equations and graphs. This imagination of dynamics, and the cognitive mechanisms underlying such imagination, appear critical to the development of RC in chemistry. A preliminary study was done to characterize student difficulties, as a requirements-gathering phase for interaction designs for gaining RC. The study replicated some results from literature, and adds further details about how students move their eyes as they navigate (through) the MERs. Further analysis would help isolate eye-movement and navigation patterns related to RC.

References

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1988). Theories, principles and laws. *Education in chemistry*, *25*(3), 89-92.

Chandrasekharan, S. (2009). Building to discover: a common coding model. *Cognitive Science*, *33*(6), 1059-1086.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science education*, *90*(6), 1073-1091.

Garforth, F. M., Johnstone, A. H. & Lazonby, J. N. (1976). Ionic Equations: Difficulties in Understanding and Use. *Education in Chemistry.*, 13, 72–75.

Gilbert, J. K., & Treagust, D. (Eds.). (2009a). *Multiple representations in chemical education* (Vol. 4). Springer.

Gilbert, J. K., & Treagust, D. (2009b). Towards a coherent model for macro, submicro and symbolic representations in chemical education. In J. K. Gilbert & D. Treagust (Eds.), *Multiple representations in chemical education,* 333–350. The Netherlands: Springer.

Herron, J. D. & Greenbowe, T. J. (1986). What Can We Do About Sue: A Case Study of Competence. *Journal of Chemical Education*, *63(6),* 528.

Hinton, M. E., & Nakhleh, M. B. (1999). Students' microscopic, macroscopic, and symbolic representations of chemical reactions. *The Chemical Educator*, *4*(4), 158–167.

Johnstone, A. H. (1982). Macro and microchemistry. *School Science Review*, *64*(227), 377-379.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of computer assisted learning*, *7*(2), 75-83.

Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education, 70*(9), 701–705.

Kelly, R. M. and Jones, L. L. 2008. Investigating students' ability to transfer ideas learned from molecular animations of the dissolution process. *Journal of Chemical Education*, 85: 303–309.

Kirsh, D. Thinking with External Representations. *AI and Society*. Springer: London, (2010) 25:441–454.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of research in science teaching*, *34*(9), 949-968.

Kozma, R., (2003). Material and social affordances of multiple representations for science understanding, *Learning and Instruction, 13*(2): 205-226.

Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualization in science education*, 121-145. Springer: 2005.

Levy, S. T., & Wilensky, U. (2009). Crossing levels and representations: The Connected Chemistry (CC1) curriculum. *Journal of Science Education and Technology*, *18*(3), 224-242.

Madden, S. P., Jones, L. L., & Rahm, J. (2011). The role of multiple representations in the understanding of ideal gas problems. *Chemistry Education Research & Practice.*, *12*(3), 283-293.

Nakhleh M. B. & Postek B. (2008). Learning Chemistry Using Multiple External Representations .In J.K. Gilbert et al., (eds.), *Visualization: Theory and Practice in Science Education*, 209–231. Spinger: 2008.

Nitz, S., Nerdel, C., & Prechtl, H. (2012). Modeling the Relationship between Representational Competence and Domain Knowledge. In E. de Vries & K. Scheiter (Eds.), *Staging Knowledge and Experience: How to Take an Advantage of Representational Technologies in Education and Training?*, 160-162.

Nurrenbern, S. C. & Pickering, M. (1987) Concept Learning versus Problem Solving: Is There a Difference? *Journal of Chemical Education, 64(6)*, 508.

Sanger, M. J., & Phelps, A. J. (2007). What are students thinking when they pick their answer? A content analysis of students' explanations of gas properties. *Journal of Chemical Education, 84*(5), 870–874.

Savoy, L. G. (1988). Balancing Chemical Equations. *School Science Review*, *69(249)*, 713-720.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction*, *29*(1), 123-145.

Stieff, M., & McCombs, M. (2006). Increasing representational fluency with visualization tools. In *Proceedings of the 7th international conference on Learning sciences*, 730-736. International Society of the Learning Sciences.

Stieff, M., & Wilensky, U. (2003). Connected chemistry—incorporating interactive simulations into the chemistry classroom. *Journal of Science Education and Technology*, 12(3), 285-302.

Zhang, Z. H., & Linn, M. C. (2011). Can generating representations enhance learning with dynamic visualizations?. *Journal of research in science teaching*, *48*(10), 1177-1198.

Wu H. K., Shah P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, 88(3):465–492.

Yarroch, W. L. (1985). Student understanding of chemical equation balancing. *Journal of Research in Science Teaching*, *22*(5), 449-459.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

How does representational competence develop? Explorations using a fully controllable interface and eye-tracking

Aditi KOTHIYAL[a*], Rwitajit MAJUMDAR[a], Prajakt PANDE[b], Harshit AGARWAL[c], Ajit RANKA[d], Sanjay CHANDRASEKHARAN[b]

[a]*Inter-disciplinary program in Educational Technology, Indian Institute of Technology Bombay, India*
[b]*Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, India*
[c]*Indian Institute of Technology Roorkee, India*
[d]*Birla Institute of Technology and Science Pilani, India*
*aditi.kothiyal@iitb.ac.in

Abstract: Representational competence (RC), defined as "the ability to simultaneously process and integrate multiple external representations (MERs) in a domain", is a marker of expertise in science and engineering. However, the cognitive mechanisms underlying this ability, and how this ability develops in learners, are poorly understood. In this paper, we present a fully manipulable interface, designed to help school students develop RC, and a pilot eye and mouse tracking study, which sought to develop a detailed understanding of how students interacted with our interface. We developed an analysis methodology for eye and mouse tracking data that characterizes the interaction process in analytical terms, and operationalizes the process of MER integration. We present preliminary results of applying our analysis methodology to student data obtained in our pilot study.

Keywords: multiple external representations, representational competence, distributed cognition, embodied cognition, equations, graphs

1. Introduction and Related Work

Representational competence (RC) is defined as "the ability to simultaneously process and integrate multiple external representations (MERs) in that domain" (Pande and Chandrasekharan, 2014). MERs are used extensively in science and engineering, and students have difficulties in learning owing to problems in working with MERs (Pande and Chandrasekharan, 2014 has a review). Students understand and are able to use and generate graphs and equations independently (Sherin, 2001; Hammer, Sherin and Kolpakowski, 1991). However students often have difficulty understanding how the two representations are related and can be used together (Kozma and Russell, 1997; Knuth, 2000). This indicates that there is a clear need for development of RC among students.

Computer interfaces with MERs have been widely used for the improving conceptual, phenomenon and procedural understanding in science and engineering (Rutten, van Joolingen and van der Veen, 2012). Despite this, the effectiveness of available computer interfaces for learning has been mixed (Ainsworth, 2006; Rutten, van Joolingen and van der Veen, 2012; Bodemer et al, 2004). One possible reason for this is that interface design is currently guided by information processing theories of cognition, wherein the role of the interface is to decrease the learner's cognitive load, particularly working memory load (Ainsworth, 2006; van der Meij and de Jong, 2006). However, emerging theories, such as distributed and embodied cognition (Glenberg, Witt and Metcalfe, 2013), postulate that external representations play more roles than decreasing cognitive load (Kirsh, 2010; Kirsh and Maglio, 1994). Further, actions could be a way of promoting integration of MERs (Chandrasekharan, 2009). Tangible interfaces, based on embodied cognition theories, have been used for learning (Marshall, 2007). But there is no consensus on how such representations should be combined for effective integration, the benefits of various approaches, or the cognitive effects of combining representations (Marshall, 2007).

Finally, there is a dearth of research which focuses directly on the development and assessment of RC using computer interfaces. Examples are Johri and Lohani (2011), Stieff, Hegarty and Deslongchamps (2011) and Wilder and Brinkerhoff (2007), and these are also based on working memory load design principles. Approaching the RC development problem from new theories of cognition could help in developing better interaction designs that facilitate MER integration.

In this paper, we report on the design of such a computer interface. We applied insights from embodied and enactive theories of cognition, particularly common coding and tool use (Maravita and Iriki, 2004) and theories of how building and manipulation of external models could lead to conceptual change and discovery (Chandrasekharan, 2009) to identify interaction features that will result in the integration of MERs and the development of RC.

The interface is designed for self-learning by a grade 7 student, and includes specific tasks that encourage exploration. We developed a stable initial prototype of the interface and performed a pilot study to understand the interaction process in detail. We recorded student eye movements and mouse clicks using an eye-tracker with the goal of developing a way to capture the RC development process. Our specific research question (RQ) was: "How can eye tracking data analysis give us more insight into the process and mechanism of MER integration?"In this paper, we report preliminary results of our ongoing work towards answering this RQ.

2. Design of the Interface

We chose the concept of oscillation of a simple pendulum as the medium to examine the development of RC. This is because the concept is easy to understand for a 7^{th} grade student, and we didn't want conceptual complexity to interfere with the learners' integration of representations.

Our learning objectives (LOs) for this interface were that the student should understand (i) the idea of equation and graph as dynamic entities (ii) the idea of equation as a controller of systems, and (iii) different numerical-spatial and dynamic-static transformations and develop an integrated internal representation, consisting of the physical system, equation and graph.

Our design, unlike simulation models with similar elements, such as Netlogo (Wilensky, 1999) and PhET (Perkins et al, 2006), is derived from basic research, particularly education research examining RC, and recent cognitive science theories and models, including distributed and embodied cognition, that investigate the cognitive roles played by different kinds of representations and their underlying cognitive/neural mechanisms (Marshall, 2007; Kirsh, 2010; Kirsh and Maglio, 1994; Chandrasekharan, 2009). One feature derived from basic cognition research is the full manipulation of the interface, which seeks to promote integration of MERs. This link is derived from an embodied cognition idea - that actions and manipulation, i.e. motor control, requires integrating multiple cognitive and perceptual inputs, and feedback loops. This suggests that actions and manipulations performed on MERs in an interface would trigger/prime the neural processes involved in integration of inputs; thus it would help in integrating the multiple representations as well. This line of thinking led to making the equation components manipulable. This also introduces the controller role of the equation, a feature not seen in other interactive visualizations.

In this design, students control and 'enact' the equation, and integration is hypothesized to result from this control feature. Thus the (eventual) testing of the development of RC based on our design would also involve testing this hypothesis, and by extension, the cognitive theory that underlies it. Applying these cognitive theories to our interface leads to features such as full learner manipulation of the pendulum via clicking and dragging, controlling the equation parameters using vertical sliders, and complete interconnection between the three modes. By contrast, a PhET pendulum simulation **Error! Reference source not found.** does not have the equation and graph, and there is only one interaction on the pendulum, while the other variable is manipulated via horizontal sliders. The design of the interface evolved through three iterations and was based on a set of design principles from distributed and embodied/enactive cognition theory (Kirsh, 2010, Kirsh and Maglio, 1994, Chandrasekharan, 2009) which are shown in Table 1, along with our operationalization of these principles. Other mappings are possible.

In order for the LOs to be met, students need to be able to do the following: (i) Map a physical system to a graph, (ii) Map a physical system to an equation and (iii) Map an equation to a graph. We designed a series of three tasks, requiring the student to manipulate the equation and pendulum to match a given graph. We hypothesized that these tasks were complex enough to result in extensive exploration and manipulation of the interface by the student, leading to the three representations being integrated.

Screenshots of the first two versions of the interface are shown in Figure 1, while a screenshot of the final version used in the pilot study is shown in Figure 2.

Table 1: Design principles and operationalization

Principle	Operationalization
External representations allow processing not possible/ difficult to do in the mind.	The interface plots the graph of the equation/motion of the pendulum for various lengths and initial angles of the pendulum.
Cognition emerges from ongoing interaction with the world.	The interface is fully manipulable, i.e., the learner can control the pendulum, equation and graph, to see how change in each affects the other elements.
Features of the world are used directly for cognitive operations. Hence the interface features should support integration directly.	The interface has the physical system, equation and graph, along with different numerical values. The dynamicity of elements, and their interconnections are made transparent, so that learners can integrate across spatial-numerical and dynamic-static modes.
The active self is critical for integration of features.	The exploration on the interface is guided by tasks which the learner must do.
Action patterns can activate concepts, hence actions and manipulations of the representations should be related to existing concepts.	The learner can interact with the pendulum by changing its length and initial angle by clicking and dragging the mouse. The parameters in the equation can be changed using vertical sliders - moving up/down increases/decreases parameter values. This is related to the finding that numbers are grounded by associating small magnitudes with lower space and larger magnitudes with upper space (Fischer, 2012). By contrast, a PhET pendulum simulation (Perkins et al, 2006) does not have the equation and graph, and there is only one interaction on the pendulum, while the other variable is manipulated via horizontal sliders. These interactions distinguish our interface from other variable manipulation simulations, wherein the mode by which values are changed (slider, input box or multiple options) is not relevant. Our interface seeks to make the learners do actions that mimic the behaviour of the system, so that the system can be 'enacted' - the learning is thus through a form of participation with the system.
The interface should allow coupling of internal and external representations.	The task requires student to match a given graph. Learners change the parameters of the pendulum/equation to generate the graph, and visually match the task graph to their graph. This develops learner's imagination and coupling between their internal model and the external representation.

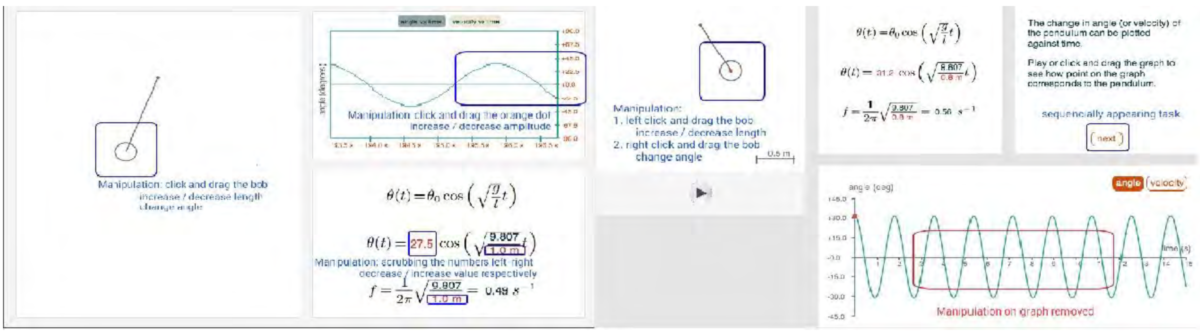


Figure 1. First version of interface (L). All 3 components (pendulum, graph and equation) are manipulable. The second version of the interface (R) only pendulum and equation manipulable.

3. Methodology

A pilot study was done with the broad research goal of developing an analysis methodology -- i.e. how to characterize interactions with our interface, and how to connect this to RC. Our specific RQ was, "How can eye tracking data give us more insight into the process and mechanism of MER integration?"

Our (convenient) sample consisted of twelve (6 female) 7th grade school students from two urban schools in Mumbai. Each student was allowed to work independently with the interface for as long as he/she wished, proceeding through the screens and tasks by clicking the "Next" button. When students had a question the experimenter provided appropriate hints. When the students indicated that the tasks were completed or that they wished to quit, they were interviewed regarding their background and their impressions of the interface. They were then administered an offline assessment task.

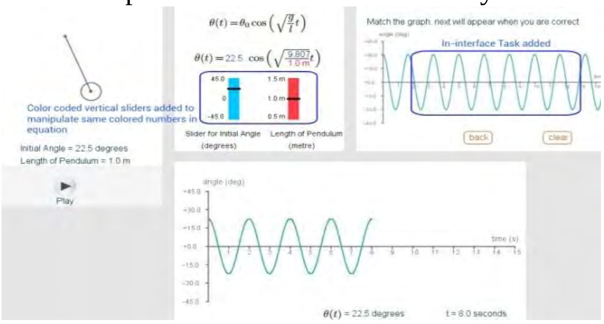


Figure 2. Final interface with sliders and tasks

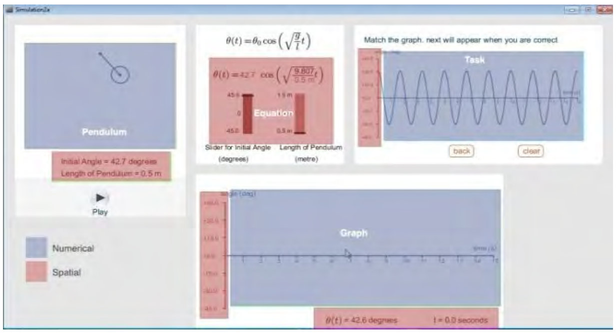


Figure 3. AOIs used for the analysis.

Our data sources were:

1. Eye Tracker: Eye movements recorded using a Tobii X2-60 (static) eye-tracker, capturing how students' loci of attention shifted as they explored the interface.
2. Assessment task: To evaluate the extent to which students are able to imagine and mentally simulate the movement that they observed on the interface. Consisted of 3 multiple choice questions, asking students to imagine the position of the pendulum from the graph, and 3 marking questions, asking students to mark points on the graph corresponding to the pendulum's position.

4. Analysis Approach

The goal of our analysis is to pull out interaction patterns from eye and mouse tracking data and explore what it means for a learner working with our interface to develop the thinking skill of RC. For this, we needed to identify patterns in the student interaction that could be markers for integration of MERs. To do so, areas of interest (AOIs) as depicted in Figure 3 were defined, and the eye fixation and mouse click co-ordinates in the respective AOIs were extracted from the eye-tracker. The data was analyzed at multiple levels of abstraction as shown in Figure 4.

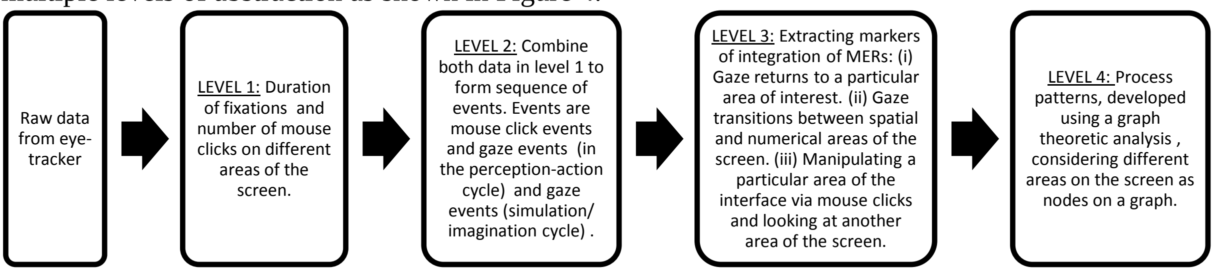


Figure 4: Levels of Analysis

The data obtained from the eye tracker includes eye fixation durations and number of mouse clicks on different areas of the screen (level 1 analysis). In level 2, we determine sequences of fixation events and mouse click events, and classify them into events occurring in the perception-action cycle, and events occurring in the simulation/imagination cycle. The perception-action cycle refers to students manipulating features on the screen (e.g. sliders), playing the simulation, and looking at the dynamic features of the screen (for e.g. the plotting graph). The simulation/imagination cycle (or thinking) happens when the simulation is paused and involves students looking at the static features on the screen (e.g. length/angle values and the graph).

In level 3 analysis, we define markers that signify integration, and abstract out the data further to calculate these markers. An example of a marker is returns, i.e. a learners' eye gaze returning to a particular area of interest after going elsewhere, as this indicates that the learner is retaining a particular feature in memory and returning to it. A second example is eye gaze transitions between a numerical area on the screen (e.g. the equation) and a spatial area on the screen (e.g. the graph) as this specifies integration between numerical and spatial modes. The third example is the learner manipulating a feature on the screen (e.g. pendulum) and looking at another area of the screen (e.g. graph) as this indicates the integration of two representations via the systematic variation offered by control. Once these markers are obtained, we define a goodness measure for these markers by comparing against marker values of experts, or marker values of learners who perform well on the assessment tasks.

The final stage of abstraction is to generate process patterns of how the learners interacted with the interface, using a graph theoretic framework, wherein the AOIs are the nodes and the transitions between the various AOIs are the weights of the branches. The duration of returns, and the sequence in which returns occurred, will also be added to this graph. These graphs will then be compared to the graphs of experts or learners who perform well on the assessment tasks to evaluate learner process. The comparison of graphs is a complex problem, and this is not implemented yet. Thus, results of the analysis at levels 2, 3 and 4 will answer our RQ, "How can eye tracking data analysis give us more insight into the process and mechanism of MER integration?" by allowing us to correlate interaction behaviours such as returns with MER integration (i.e. high performance on the assessment tasks).

5. Indicative results

For lack of space, in this paper we present indicative results, applying our analysis methodology to the data of one student who performed well on the assessment task. This is ongoing work, and we have not completed the level 4 analysis, and correlated the results to assessment task performance, which would give us an answer to our RQ. The data at level 1 of analysis, namely fixations and mouse clicks, is reported elsewhere (Majumdar et al, 2014). Here we report analysis of the fixation data at levels 2 and 3. Figure 4 shows an example event sequence for the learner between two consecutive clicks on the play button and the legend is shown in Figure 7 (also see AOIs in Figure 3). The sequence of events between the play and the pause button are events in the perception/action cycle, while events after the pause button are in the imagination cycle. This sequence shows that the student transitions between spatial and numerical regions both in the action and imagination cycles.

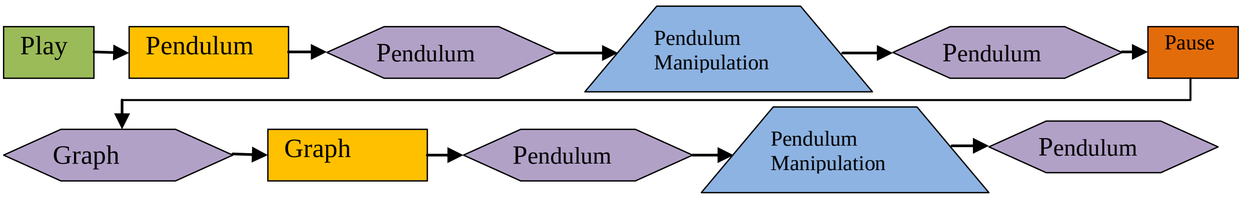


Figure 4: An example of a sequence of events for a good performing student

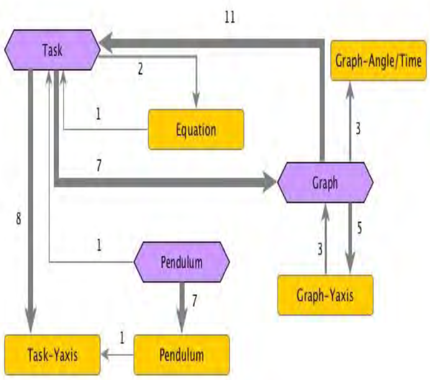


Figure 5: Numerical-Spatial Returns

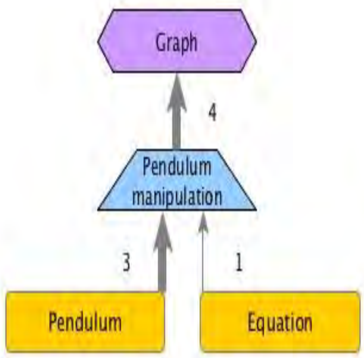


Figure 6: Click-gaze transitions

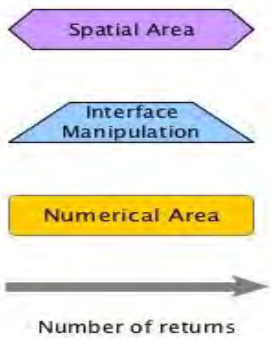


Figure 7: Legend for figures 4, 5 and 6

Next we present two markers of integration at level 3 of the analysis. The first is eye gaze transitions between numerical and spatial areas on the screen (Figure 5) and the second is the transitions between mouse clicks and eye gazes on different areas of the screen (Figure 6). In these figures, the thickness and numbers on the arrow from A to B indicates the number of A-> B-> A transitions made by the student. For instance, Figure 4 shows that this student looks from the spatial area of the graph to the spatial area of the task and returns 11 times. In the final level of analysis, the return data will be combined with duration of each return, to create a rich graph representation of the students' interaction process, which will then be compared to the processes of an expert and a low-performing student.

6. Conclusions and Future Work

In this paper, we presented the design of an embodied computer interface for the development of RC. We evaluated the interface in a pilot study, developed an analysis methodology for extracting process patterns (i.e. how students interacted with our interface) from eye and mouse tracking data and evaluating how these process patterns translate to MER integration. We also presented preliminary results using this analysis. Once complete, our methodology becomes a template for analyzing the process of how learners interact with a new design, using eye and mouse tracking and evaluating whether MER integration occurs using that design.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. Learning and Instruction, 16(3), 183-198.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. Learning and Instruction, 14(3), 325-341.

Chandrasekharan, S. (2009). Building to discover: a common coding model. Cog Sci, 33(6), 1059-1086.

Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. Cognitive processing, 13(1), 161-164.

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the Revolution to Embodiment 25 Years of Cognitive Psychology. Perspectives on psych sci, 8(5), 573-585.

Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. Journal of Mathematical Behavior, 10, 117-160.

Johri, A., & Lohani, V. K. (2011). Framework for improving engineering representational literacy by using pen-based computing. Int. J. Engg Edu, 27(5), 958.

Kirsh, D. (2010). Thinking with external representations. Thinking with External Representations. AI and Society. Springer: London, 25:441–454.

Kirsh, D., &Maglio, P. (1994). On distinguishing epistemic from pragmatic action. Cog Sci, 18(4), 513-549.

Knuth, E. J. (2000). Student understanding of the Cartesian connection: An exploratory study. Journal for Research in Mathematics Education, 500-507.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. Journal of research in science teaching, 34(9), 949-968.

Majumdar, R., Kothiyal, A., Pande,P., Agarwal, H., Ranka, A., Murthy, S., Chandrasekharan, S. (2014). The enactive equation: exploring how multiple external representations are integrated, using a fully controllable interface and eye-tracking. In Technology for Education, 2014 Sixth International Conference on. IEEE.

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). Trends in cognitive sciences, 8(2), 79-86.

Marshall, P. (2007). Do tangible interfaces enhance learning? In Proc 1st int conf on Tangible and embedded interaction (pp. 163-170). ACM.

Pande, P., & Chandrasekaran, S. (2014), A general account of representational competence, Unpublished.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics. The Physics Teacher, 44(1), 18-23.

Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. Comp & Edu, 58(1), 136-153.

Sherin, B. L. (2001). How students understand physics equations. Cognition and instruction, 19(4), 479-541.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. Cogn & Instr, 29(1), 123-145.

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. Learning and Instruction, 16(3), 199-212.

Wilder, A., & Brinkerhoff, J. (2007). Supporting representational competence in high school biology with computer-based biomolecular visualizations. J. Comp. in Math. and Sci. Teaching, 26(1), 5-26.

Wilensky, U. (1999). Netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Using Ontology for Representing Role Change Design in Nursing Service Thinking Education

Wei CHEN [a*], Liang CUI [b], Koji TANAKA [a], Hirotaka NISHIYAMA [a], Noriyuki MATSUDA [c] &Mitsuru IKEDA [ab]

[a] *School of Knowledge Science, JAIST, Japan*
[b] *Research Center for Service Science, JAIST, Japan*
[c] *Faculty of Systems Engineering, Wakayama University, Japan*
*wei.chen@jaist.ac.jp

Abstract: This study aims to develop a learning design representation framework to support the modelling of an in-service education program for fostering the meta-thinking skill of nurses. In this paper, we describe the development of the ontological framework for representing the design intention of the learning goals of the meta-thinking skills and also the relationship of the learning experiences gained from changing leaner's roles.

Keywords: instructional design, ontological engineering, metacognition, nursing service education

1. Introduction

In the social constructivist view of education, the focus of learning transfers from knowing facts and procedures to acquiring knowledge and skills for solving the authentic, contextual and social problems which do not always have a clearly correct or universal answer. The characteristics of these problems are that i) they can be formulated or defined from various perspectives; ii) their solution is not unique, and the criteria for choosing a better solution is implicit or situation-dependent (Munneke, et al., 2007).

Nursing education is one of the forefronts of the professional education areas which aim to foster the abilities for solving the patient related problems. To acquire these problem solving abilities, it is crucial to acquire implicit knowledge and skills based on the experience of reflecting on one's own thinking. Various educational methods, including narrative method and reflective journal have been conducted in the professional education of nurses in order to support the learning of these knowledge and skills (Bulman, 2013).

In this research, we focus on the design of meta-thinking skills education. The terminology "meta-thinking" is a particular concept. The meta-thinking skills are defined as the skill of monitoring and control of thinking which targets to clarify beliefs behind the conflicts of different perspectives. Compared with the concept of Dewey's reflective thinking, the meta-thinking is the narrowly-defined concept that focuses on the careful consideration of belief clarification. We consider that improving the nurse's meta-thinking skills is crucial to build the foundation for the patient-centered medical service. For example, in the realistic setting of medical service, the patients have a tendency to lie to the doctors for protecting the private information. On the contrary, the doctors want to know more information for diagnosis. It is helpful to improve the meta-thinking skills for promoting the quality of medical service with due consideration of the rationality of medicine and the patients' humanity.

For supporting the design of learning goals, learning content and education methods, various theories (Gagné et al., 2004; Keller, 2009) and techniques (Hayashi, 2004; Paquette, 2006) have been studied. For example, Hayashi et al. (2009) proposed an ontology-based theory-aware and standards-compliant authoring system to provide a general knowledge-base for linking educational theory and practice.

However, it is difficult to create a clear model for education which targets the fostering of experientially-acquired skills (especially thinking skills). This study aims to develop a learning design representation framework which supports the modelling of meta-thinking skills education for nursing services. This paper describes the development of an ontological framework for representing the design intention of the learning goals of the meta-thinking skills and also the relationship of the learning experiences gained from changing leaner's roles. The entire ontological framework built through the study will be reported separately.

2. Education Program for Fostering Meta-thinking Skills

An education program named Nursing Service Thinking Method Workshop was designed to foster nurses` meta-thinking skills. The educational objective is that, through building the skill of the thinking representation (identification of thinking logical structure), nurses are expected to 1) understand both of examining their own thinking and considering about other's perspective through identification of thinking logical structure; and 2) to be motivated to continue to learn meta-thinking skills by themselves after the end of the education program (Cui et al, 2014).

As shown in Figure 1, the workshop consists of six parts: Lecture, Case-writing Practice, Case-writing (homework), Case Reviewing, Discussion and Reflective Lecture. It is conducted in 2 day-long sessions spaced approximately one month apart; the workshop is conducted three times per year. It is also necessary for the leaners to attend the entire program of the workshop for three years, because they need to change their role after learning in the workshop for each year.

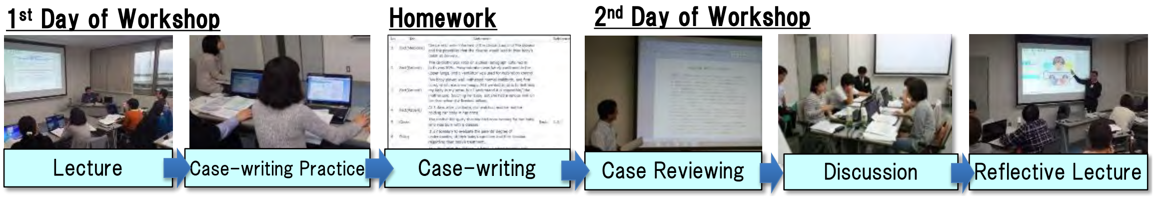


Figure 1. Overview of Nursing Service Thinking Method Workshop

2.1 The Characteristics of the Workshop

a. Using a Reflective Case-writing Supporting Tool named Sizhi (Figure 2) as the core learning material for improving the quality of thinking reflection through the support on the identification of thinking logical structure (Chen et al., 2011).
b. A form that combines the self-thinking reflection experience from case-writing and the group-thinking reflection experience from discussing the case with other learners to help participants recognize the connections between those two kinds of experience and promote the acquisition of meta-thinking skills.
c. A step-by-step learner's role designed to utilize the learning experience from the previous step to trigger the useful learning experience in the successive step to promote the acquisition of meta-thinking skills. The definitions of the leaner's role are:

   Member (1st Year Attendance) attends six sessions of the workshop, acquires fundamental knowledge and skills related to thinking representation logical thinking.

   Leader (2nd Year Attendance) attends the case-writing practice session and the discussion session. In the discussion, the Leader does not take the initiative in presenting a viewpoint but views discussion from the higher perspective and guiding it. Based on the experience of learning in the role of a Member, monitoring and controlling of group thinking skills can be promoted.

Facilitator (3rd Year Attendance) does not attend the workshop but supports the learning of members through revising and giving comments to the cases written by members (Nishiyama et al., 2014).

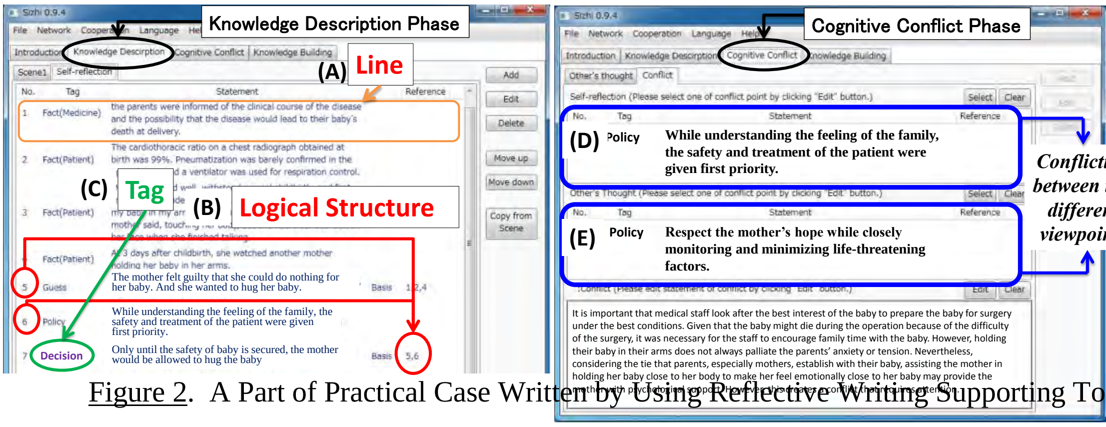


Figure 2. A Part of Practical Case Written by Using Reflective Writing Supporting Tool

Before changing becoming a Leader, a Member needs to attend the workshop three times in one year to acquire fundamental knowledge and skills from their learning experience. Some of the learning experiences Members undergo are implicitly designed to promote further learning as a Leader. This relationship of learning experiences is an important design aspect of the learning of meta-thinking skills.

3. Ontology-base Representation Framework

In order to represent the implicit design intentions mention above, we develop a learning design representation framework that consists of the graphic representation form and the ontology representation form. The graphic representation form can provide a descriptive medium to represent the elements of learning design such as learning goal, learning strategy and etc. The merit of the graphic form is that it can marshal the different elements of learning design in a structured way. The ontology representation form is created for representing the learning design intentions in a further precisely and systematically, which is difficult for the text or graphic representation from to be done.

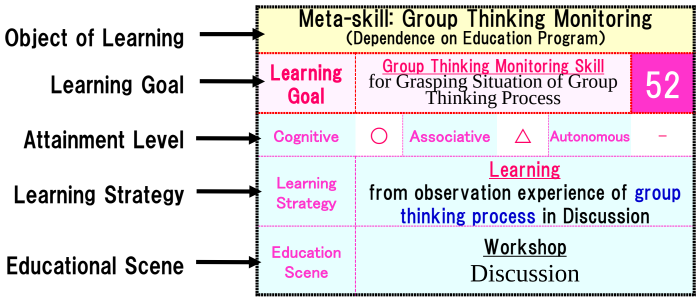


Figure 3. An Example of Learning Unit

An example of the basic form of the graphic representation is shown in Figure 3. The whole of the box in the figure is named Learning Unit. The learning unit is created for integrating the concepts related to the design of the education program, including Object of Learning, Attainment Level of Learning, Learning Goal, and Learning Strategy. The row that represents Learning Goal states the goal which the learning unit aims to achieve. The content in the row that represents Object of Learning refers to the kind of skill, knowledge or attitude. The content in the row that represents Attainment Level indicates the attainment level of learning on the object of learning. For example, the learning unit shown in Figure 3 indicates that the learners are expected to achieve the learning goal "Group Thinking Monitoring Skill for Grasping Situation of Group Thinking Process" to the attainment level of "Cognitive", by the learning strategy learning from observation of group thinking process in discussion. The Group

Thinking Monitoring refers to the realization of group thinking condition through analyzing and evaluating the process of group thinking. The level of Cognitive means knowing fundamental knowledge or knowledge on how to practice skill. Moreover, the learning goal is not independent but a part of a continuous learning process which regards the program dependence meta-thinking skill of "Group Thinking Monitoring" as the Object of Learning. This implies that this skill is intended to be acquired gradually within the program through several steps.

By using the learning unit described above, the learning design of the workshop can be organized and represented in an easily viewable way. In addition, the learning experience gained from what kind of Member's activities influence which part of Leader's learning can be represented. For example, in Figure 4, the learning unit on the left side represents the member's learning experience gained from externalizing the result of logical thinking in Sizhi. The learning unit on right side represents the leaders' learning on the meta-thinking skill through attending and observing the discussion. One of the purposes of the case-writing is to help the participants become aware of the logical structure of thinking by externalizing his/her thinking. This awareness also benefits the discussion, because this awareness is expected to trigger the action of realizing the logical structure of group thinking process.

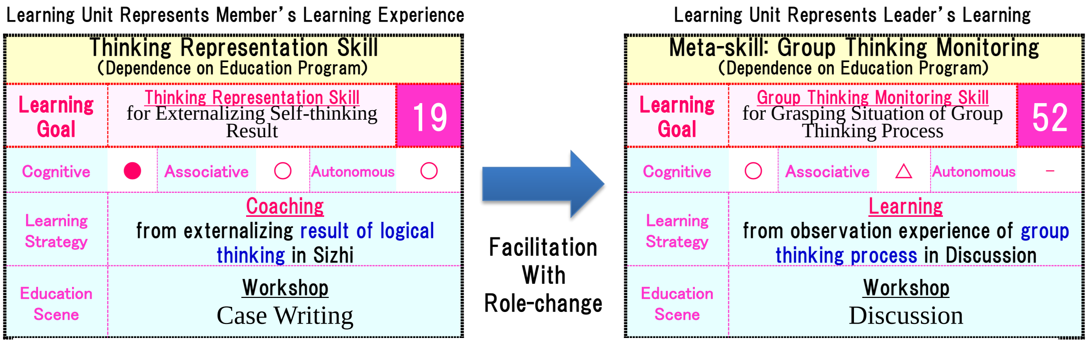


Figure 4. An Example of Relation between Member's Experience and Leader's Learning

In this research, we use the Hozo ontology editor (Hozo Ontology Editor) to build the representation framework (Chen et al, 2014). The most significant feature of the Hozo ontology editor is that it provides a function to support users to build ontologies being aware to distinguish role concept from basic concept (Mizoguchi, et al, 2007). For example, in schools, the teacher role is set up and the person who is assigned the teacher role is considered as a teacher. In this example, the person is basic concept and the teacher role is role concept. The concept of teacher is named role holder. This feature is practical to explore and represent the essence of the concept which is always easily-confused.

The ontology related to the facilitation by experience is shown in Figure 5. A~E of the figure indicates the concepts in a hierarchy. The concepts in upper levels (such as A) represent more general concepts; lower levels (such as E) represent more specific concepts. E shows a specified concept of Facilitation by Experience with Role-change (FER) named FER of Member's Case-writing Experience to Lead's Learning in Discussion. As described in Figure 4, the member's experience of coaching from externalizing result of logical thinking in Sizhi is expected to be helpful for facilitating the leader's learning conducted with the learning strategy of Discovery from observation of group thinking process in Discussion. By using ontology, this influence can be represented. Because the role change in this example is from member to leader, the pro-role is restricted with concept of member [RH] (e1) and Target role is restricted with the concept of leader [RH] (e2). RH means role holder according to the ontology theory that Hozo is based on. In the workshop, the person who is assigned the member role is defined as member [RH] and the person who is assigned the leader role is defined as leader [RH].

In the learning experience (e3), the attributes of member [RH] as the holder (d4) and Identification on logical structure (e5) to self-thinking result (e6) as the outcome are defined.

The meaning of this definition is that the action identification on the logical structure of self-thinking result is involved in the member's learning experience of coaching from externalizing result of logical thinking in Sizhi. Similarly, in the learning strategy (e7), the attributes of leader [RH] as the holder (e8) and Identification on logical structure (e9) to group thinking process (e10) as the action are defined. The meaning of this definition is that the action identification on the logical structure of group thinking process is involved in the leader's learning conducted with the learning strategy of Discovery from observation of group thinking process in Discussion.

The relation same-person (R1) between Pre-role and Target-role represents that the same person plays the different role in the workshop and role-change is conducted from Pre-role to Target-role. The relations sameAs (R2 and R3) mean that the learning experience belongs to the person who is assigned the member role and the learning strategy belongs to the same person who is assigned the leader role.

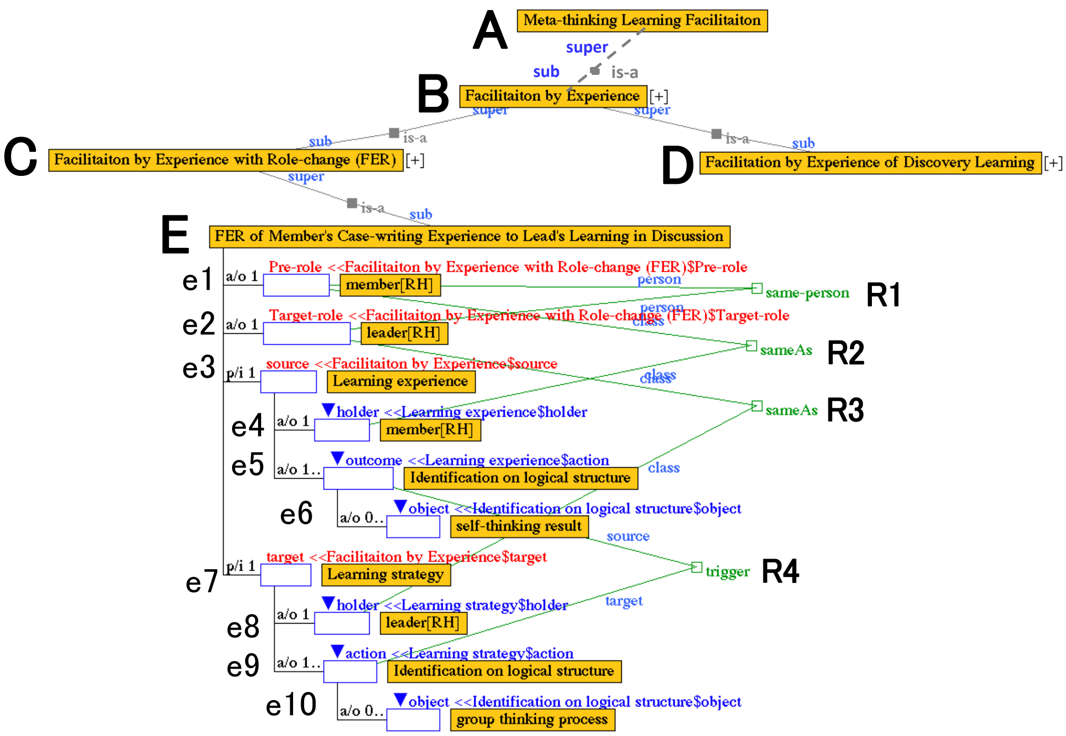


Figure 5. Ontology related to Role Change (Part)

By defining the concepts of learning experience and learning strategy, it is represented that the concepts of Identification on logical structure (e5) and Identification on logical structure (e9) have the same name, but the meaning of these two concepts are different, because they play different role for the concepts which they belongs to. Furthermore, the relation of e5 and e9 is indicated with the relation of triggering (R4). It represents how an outcome of the member's learning experience from the case-writing influence the leader's learning.

As described above, by using ontology, the implicit relation of learning experience is represented. Although the learners will not use the ontology directly, representing the learning design by ontology is useful to encourage the designer to consider about the improvement. For example, in leader's learning, it is necessary to help them to realize the similarities and differences of self-thinking and group thinking and to support their conduct of the action Identification on logical structure in the discussion.

4. Conclusion and Future Work

The meta-thinking skills are difficult to facilitate but crucial for the nurses, because it is important to reexamine and integrate self-thinking, the thinking of patients and the thinking of co-workers and other healthcare professionals through the dialogue with self and other people to achieve high-quality patient-centered medical service. We have been working on this practical research in collaboration with medical institutions. This study is intended to create a foundation for the process of design, implementation and redesign for the educational practice in nursing service. For facilitating meta-thinking skills, besides the design of the role change model described in this paper, we are creating several other models. In this study, the circulating procedures of model creation, ontology construction through segmenting the concepts, and the refinement through practice are conducted repeatedly. The workshop is designed as one turn in three years and the current research is at the stage of completion of former two years. Using this representation framework is helpful to share the design intentions of the meta-thinking skill fostering workshop which is conducted in a long term. We are currently working on the creation of the facilitator's educational model and its application. In addition, we plan to reorganize and report the overview of the entire model after completing other parts.

In summary, the graphic and ontological representation framework is used in the design and creation of textbooks and exercises in the implementation of workshop, in the development of evaluation methods and even for sharing the design intentions between the people who are in charge of different parts of learning design cycle. The feedback from those practical learning design activities is positive. In the future, we will develop a supporting tool for sharing the design intentions to the learners to help them to realize the learning goals. Furthermore, we will also study the supporting tool for assisting the facilitator role learners to improve or customize the workshop according to the demands of their medical institutions.

References

Bulman, C., & Schutz, S.: Reflective practice in nursing, John Wiley & Sons, 163-188 (2013).

Chen, W., Fujii, M., Cui, L., Ikeda, M., Seta, K., & Matsuda, N. (2011). Sizhi: Self-Dialogue Training through Reflective Case-Writing for Medical Service Education. In A. F. Mohd Ayub, B. Chang, K. Leelawong, F.-Y. Yu, T. Hirashima, & G. Biswas (Eds.), (pp. 551–558). Presented at the Workshop Proceedings of the 19th International Conference on Computers in Education, Chiang Mai, Thailand.

Chen, W., Cui, L, Tanaka, K.,Nishiyama, H., Matsuda, N. & Ikeda, M. (2014). A Model of Nursing Service Thinking Skill Education by Shifting Learner Role. Presented at the 26th Annual Conference of the Japanese Society for Artificial Intelligence. (In Japanese)

Cui, L., Tanaka, K., Chen, W., Masuda, N., Ikeda, M. (2014) A Thinking Skill Development Program for Improving the Medical Services Quaility. IEICE Technical Report. Education Technology, vol. 113, no. 377, ET2013-77, pp. 55-60. (In Japanese)

Gagné, R. M., Wager, W. W., Golas, K., & Keller, J. M. (2004). Principles of Instructional Design (5th ed.). Wadsworth Publishing.

Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004). Ontological Engineering: With Examples from the Areas of Knowledge Management, E-commerce and the Semantic Web. Springer.

Hayashi, Y., Ikeda, M., & Mizoguchi, R. (2004). A design environment to articulate design intention of learning contents. International Journal of Continuing Engineering Education and Life Long Learning, 14(3), 276-296.

Hayashi, Y., Bourdeau, J., & Mizoguchi, R. (2009). Using ontological engineering to organize learning/instructional theories and build a theory-aware authoring system. International Journal of Artificial Intelligence in Education, 19(2), 211-252.

Hozo Ontology Editor, http://www.hozo.jp

Keller, J. M. (2009). Motivational design for learning and performance: The ARCS model approach. Springer.

Mizoguchi, R., Sunagawa, E., Kozaki, K., & Kitamura, Y. (2007). The model of roles within an ontology development tool: Hozo. Applied Ontology, 2(2), 159–179.

Nishiyama, H., Tanaka, K., Kanou, H., Cui, L, Matsuda, N. Chen, W. & Ikeda, M. (2014). A Model of Learning Method Reconfiguration Though Reviewing Experiences. Presented at the 26th Annual Conference of the Japanese Society for Artificial Intelligence. (In Japanese)

Paquette, G., Léonard, M., Lundgren-Cayrol, K., Mihaila, S., & Gareau, D. (2006). Learning Design based on Graphical Knowledge-Modelling. Journal of Educational Technology & Society, 9(1).

Self-assessment rubrics as metacognitive scaffolds to improve design thinking

Madhuri MAVINKURVE & Sahana MURTHY
Inter-Disciplinary Program in Educational Technology, Indian Institute of Technology Bombay, India
mavinkurvemk@gmail.com

Abstract: An important goal of engineering education is to develop design thinking competency among students. These competencies include: structuring open problem, information gathering, divergent and convergent thinking, and multiple representations. We have developed Engineering Design Interactive Visualizations (EDIVs), a technology-enhanced learning environment for students' self-learning of engineering design competency. The EDIVs contain activities to promote decision-making, concept integration and overall synthesis, which are necessary cognitive mechanisms to attain design competencies. However, in addition to cognitive strategies, students require metacognitive scaffolds to help them reflectively and mindfully apply these strategies to new contexts. Hence, we included self-assessment rubrics in EDIVs to act as metacognitive scaffolds for the process of design. In this paper, we describe the role and structure of self-assessment rubrics for the engineering design competency of Structure Open Problem, and report results of an experiment where students worked with EDIVs in two conditions – with and without rubrics. We find that the rubrics prompt students to perform formative assessment of their own performance, and correct themselves if necessary, thus leading to improved performance in a new design problem.

Keywords: Design thinking competency, Technology enhanced learning environment, interactive visualization, scaffolding, metacognition, self-assessment

1. Introduction and related work

The development of design thinking ability is an important goal of engineering education. Professional organizations, accreditation bodies (ABET, 2014) as well as educators (Sheppard, 2003) have emphasized that graduating students should be able to design effective solutions for given needs. However, design thinking is complex and teaching design is difficult (Dym, 2005). Efforts to teach engineering design thinking mainly include stand-alone design courses based on a version of project-based learning (Dutson et.al, 1997). These courses have been reported to be beneficial to students, especially in promoting student interest and retention (Wood et.al., 2001). Challenges have been reported too in running such courses, such as extra faculty time, special training, and lack of assessment techniques. Thus design courses are not common in universities, which translate into lack of design ability among students (May & Strong, 2006).

The above problems are compounded in part because of the lack of a unique definition of what comprises engineering design thinking. Plenty of definitions and perspectives abound (Atman et.al, 1997; Crain et.al, 1995). What is common in all approaches is that engineering design is a systematic and thoughtful process, in which "designers generate, evaluate, and specify concepts for devices, systems, or processes" (Dym, 2005). The designed artifacts must satisfy specifications and constraints in order to meet the user's requirements. For example, in electronics system design, in order to design function generators, one needs to consider the type of waveforms, amplitude and frequency range etc. as specifications. Based on these definitions, we take a competency-based approach, and envisage the design thinking process in terms of a set of engineering design competencies that need to be developed and applied while solving design problems. These competencies include structuring open problem, information gathering, divergent and convergent thinking, and multiple representations.

In recent years, the affordances of ICT has led to the development of technology enhanced learning (TEL) environments to promote various thinking skills. There exist numerous TEL systems for modeling ability such as WISE (Linn et. al. 2003) and Co-Lab (van Joolingen et. al., 2005), scientific argumentation (Scheuer et. al, 2009 contains many examples) and designing virtual experiments (Hemlo, Nagarjan & Day, 2002), but fewer for engineering design thinking. We have developed Engineering Design Interactive Visualizations (EDIV), a TEL environment for students' self-learning of engineering design competencies.

In a prior experimental study (Mavinkurve & Murthy, 2012), we have investigated the effectiveness of EDIVs that target Structure Open Problem competency. We have shown that students who learnt with EDIV were able to develop certain sub-competencies such as identifying specifications in an open problem, and were able to apply them to a new design problem. However, students did not satisfactorily demonstrate the development of sub-competencies such as being able to synthesize and write a structured problem statement in the context of the new problem. Subsequent interviews with students revealed that they are often not aware of the cognitive processes that need to be performed at a particular time that would have led to the development of this sub-competency. Unlike experts, novice students have not internalized these cognitive processes. Hence they need to reflectively abstract these processes from the learning context, and mindfully apply them to the new context (Perkins & Salomon, 1992).

In this paper, we focus on the problem of helping students to reflectively and critically apply cognitive processes needed to write a structured statement for a given open design problem. One strategy to help students achieve this is to include metacognitive scaffolds. The inclusion of scaffolds for complex tasks has been recommended to promote students learning of not only conceptual and procedural knowledge, but also flexible thinking skills (Reiser 2004; Etkina et. al., 2010). Scaffolds can promote students' metacognitive thinking, which includes planning, monitoring, evaluating, revising and reflecting (Jacobs & Paris, 1987). Metacognitive thinking is a crucial component for design activities (Davidowitz & Rollnick, 2003). Scaffolds that promote metacognitive thinking help students learn from the experience so that they can apply knowledge and skills in new contexts.

To provide metacognitive scaffolds for design thinking, we added formative assessment rubrics in the EDIVs (Section 2). These rubrics allow students the opportunity of self-assessment, which is a powerful way of implementing formative assessment (Black & Wiliam, 1998). The rubrics provide students feedback on their responses to the EDIV activities so that they can monitor their learning process themselves with respect to the learning goals. At the same time, they focus students' attention on the important cognitive processes needed for accomplishing the complex task at hand. Thus the inclusion of the rubrics are intended to develop students' design competencies. In Section 3, we report results of an experiment where students worked with EDIVs in two conditions – with and without rubrics, and discuss the role of self-assessment rubrics in developing students' design thinking.

2. Learning Environment: Engineering Design Interactive Visualization

The process we followed to develop the leaning environment for various design competencies involved the following steps: i) identify and operationalize specific measurable units of the competency (which we refer to as 'sub-competencies'), ii) analyze the cognitive tasks that need to be performed to attain the sub-competencies, and iii) decide features and activities in the EDIVs that trigger students' cognitive mechanisms to perform these tasks. In this section, we discuss the design of EDIVs that target the competency of 'Structure Open Problem', which is one of the first tasks involved in design thinking (Sheppard, 2003).

The competency of 'Structure open problem' (SOP) is operationalized into four sub-competencies: Student should able to i) identify specifications in open-ended problem (SOP1), ii) use specifications to structure problem (SOP2) , iii) sequence steps of design process to (SOP3) and iv) write structured problem statement (SOP4). To attain each sub-competency above, students need to perform a set of cognitive tasks. For example, to be able to identify relevant specifications needed to structure open problem (SOP1), the set of cognitive tasks to be performed are: identification of all possible specifications, deciding relevant specifications and interpretation of chosen specifications with respect to the concepts. When these tasks for all sub-competencies were analyzed, we found that there are three common cognitive mechanism required to execute these tasks: A. Decision making, B. Concept integration and C. Synthesis.

A. Decision making mechanism. Decision making process is defined as mentally generating possible options for given situation and evaluating options based on set of information (Gresch & Bögeholz, 2012). Decision making is a process that all designers have to engage in throughout the design process (Dym, 2005). It involves an iterative series of divergent-convergent thinking in which students need to generate many options based on the set of information available, evaluate them based on domain knowledge expertise. Decision making is an essential triggering cognitive mechanism in the attainment of SOP1 (identification of specification) and SOP2 (use of specifications), as both these competencies require students to think of multiple options and select appropriate ones. Each EDIV contained activities and features which trigger the above cognitive mechanisms. Decision making can be triggered using question prompts (Ge & Land, 2005) as well as providing opportunity for knowledge integration through experimentation and reflection (Etkina et. al., 2010). To trigger this cognitive mechanism, we added ‘Decision Making Task Questions’ with multiple options and formative feedback. The EDIVs also contained simulative manipulation activity (Chen et.al., 2011) in which students are provided variable manipulation for experimentation, followed by questions with feedback on students’ responses. This provides opportunity for reflection. Fig. 1 shows a Decision Making Task Question activity in which students have to make decisions about the selection of specifications. Fig. 2 shows a simulative manipulation activity related to selection of amplifier circuit to satisfy specifications.

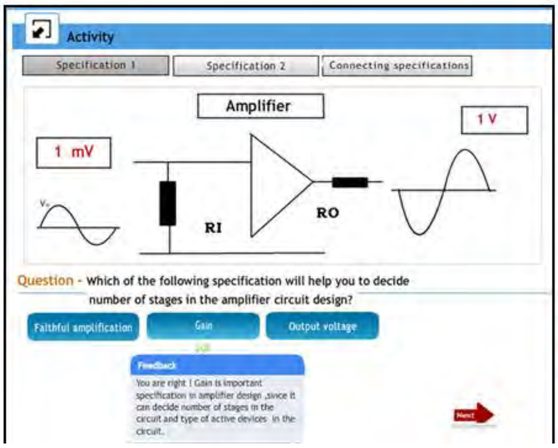


Fig.1. Decision Making Task Question activity

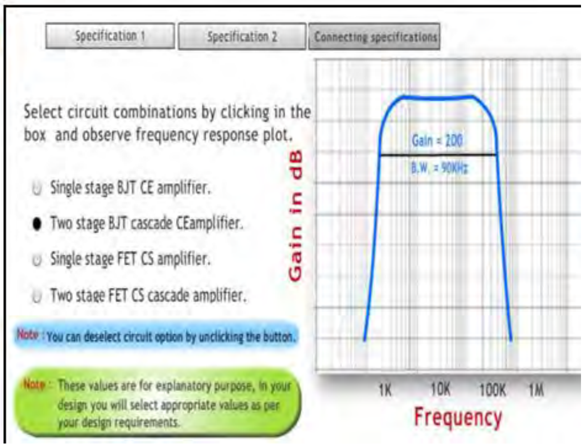


Fig.2. Simulative Manipulation activity

B. Concept Integration mechanism. This is the process of connecting various concepts using information association and knowledge integration (Aurisicchio et.al, 2007). The cognitive mechanism of concept integration expects students to associate different pieces of information based on domain knowledge. This is mainly required for SOP3 (sequence decision steps) wherein students should able to recall and connect appropriate concepts based on domain knowledge. Concept integration also requires knowledge of multiple representations with visual thinking. The third major cognitive mechanism, synthesis forces student to think about entire system. Concept integration primarily expects information association activity, for which the EDIVs contain Concept Enforcement Questions (CEQ) questions In addition, EDIVs contain controlled animation with dynamically linked representations to help students to associate information. To trigger system thinking process, EDIVs contain information agents like design tips and information box along with Decision Making Task Questions and Concept Enforcement Questions.

C. Synthesis mechanism. Synthesis can be defined as integration of all the cognitive mechanisms mentioned above and monitoring of the achievement of these mechanisms (Dym, 2005). Synthesis is required to attain SOP4 (write structured problem), as an entire system-level thinking is required for this sub-competency, including decision making with appropriate concept integration. To trigger system thinking, EDIVs should be able to provide opportunity to monitor learning process. This can be achieved through metacognitive strategies which were added in EDIVs via self-assessment rubrics, based on the scientific abilities rubrics (Etkina et. al., 2006). These rubrics are descriptive rating scales which consist of pre-established criteria to evaluate student’s performance on each design sub-competency. The rubrics for the sub-competencies related to Structure Open Problem competency are shown in Table1. After activities such as Decision Making Task Questions and Concept Enforcement Questions, students are provided the relevant rubric items. Since the rubrics contain descriptors not only of the target performance level, but also of non-ideal performance, they prompt students to carry out formative assessment of their own performance in the activity, and correct themselves if necessary. This helps students not only to monitor their level of achievement of cognitive task, but also plan learning based on expected target level.

Table 1: Rubrics items for various sub-competencies of Structure Open Problem competency

Design sub-competency	Target performance	Needs improvement	Inadequate	Missing
Is able to extract required relevant specifications from given open ended problem	All relevant visible and hidden specifications are identified and interpreted accurately. No irrelevant specifications identified.	An attempt is made to identify specification Most of them identified but few hidden ones missing or needs more interpretation.	An attempt is made but specifications identified are most of them are wrong or irrelevant or incomplete.	No attempt is made to extract specifications
Is able to structure open problem using specifications	Specifications are used to identify interconnections of the system in order to structure problem.	An attempt is made to use specifications but minor specifications are not used, or used incorrectly.	An attempt is made to use specifications but required specifications not used or wrongly applied.	No attempt is made to use specification to structure problem
Is able to sequence the design steps based on specifications	All major and minor design steps are identified and sequenced correctly based on specifications.	Most designs steps are sequenced correctly. Few steps are missing or not sequenced correctly.	Design steps are not sequenced at all or not based on specifications.	No attempt is made to write design steps.
Is able to write structured design problem statement	Problem statement is written clearly including details of specifications and design steps.	Problem statement is written clearly but few minor details are missing.	Problem statement is not written clearly but scattered information is available.	No attempt to write coherent statement.

3. Evaluating effectiveness of self-assessment rubrics in EDIVs

3.1 Method

In a prior study (Mavinkurve & Murthy, 2012) we had shown that features of EDIVs to promote decision making and concept integration (such as decision making task questions, simulative manipulation, concept enforcement questions etc) led to the improvement of design sub-competencies of SOP1, SOP2 and SOP3. Here, we report a two group quasi-experiment to investigate the importance of including self-assessment rubrics as metacognitive prompt in the EDIVs, which targets SOP4 (write structured problem statement) via the cognitive mechanism of synthesis. The two conditions in the experiment were the presence or absence of self-assessment rubrics in the EDIV.

Participants. The study participants were students from 2nd year Electronics Engineering (N=45) major. Students were familiar TEL environments, as well as with the content in the EDIV, as they had learnt it in the theory course. However, they were not exposed to design in this topic.

Procedure. Students were randomly assigned to two groups. The experimental group consisted of 23 participants, control group had 22 participants. The equivalence between the groups was tested on basis of their previous semester's grades and no statistically significant difference was found between them ($t=-0.08$, $p=0.9$). Two sets instructional materials on the topic of amplifier design from electronics domain were developed. The experimental group received an EDIV which contained self-assessment rubrics. The control group received the same EDIV but without the self-assessment rubrics. Students in both groups studied their material for 30 minutes, after which they attempted the post-test. The post-test contained an open design question on a topic related to (but not the same as) the instructional material for which students had to describe (on paper) their design.

Instrument. To assess the development of students' design competencies we used assessment rubrics, similar to the self-assessment rubrics as shown in Table 1. These rubrics were validated prior to the experiment. Inter-rater reliability was found to give 86% agreement between 3 instructors.

3.2 Results

The scores on the post-test are ordinal data, hence we used a Mann-Whitney U-test for analysis. The mean ranks for each sub-competency for the two groups are shown in Table 2. The results show that the mean ranks for the experimental group are higher in each sub-competency. However, the difference was statistically significant only for SOP3. We further analyzed the data by categorizing students based on SOP sub-competency scores. Students who scored 0 or 1 were categorized as 'unsuccessful' on the design task and students who scored 2 or 3 as 'successful' on the design task. For each sub-competency we calculated the number of students from the control and experimental group in the successful and unsuccessful categories respectively. We used the Statistical Attribute Tracking (SAT) diagram (Majumdar & Iyer, 2014) to represent and analyze the data (Fig. 3).

Table 2: Comparing SOP sub-competency scores of experimental and control group

Sub competency	Group	N	Median	Mean Rank	z	p
SOP1	Control	22	2	20.162	1.4	0.15
	Expt	23	2	25.72		
SOP2	Control	22	2	20.48	1.24	0.21
	Expt	23	2	25.41		
SOP3	Control	22	2	19.61	1.9	0.056
	Expt	23	3	26.67		
SOP4	Control	22	1	19.84	1.56	0.11
	Expt	23	3	26.02		

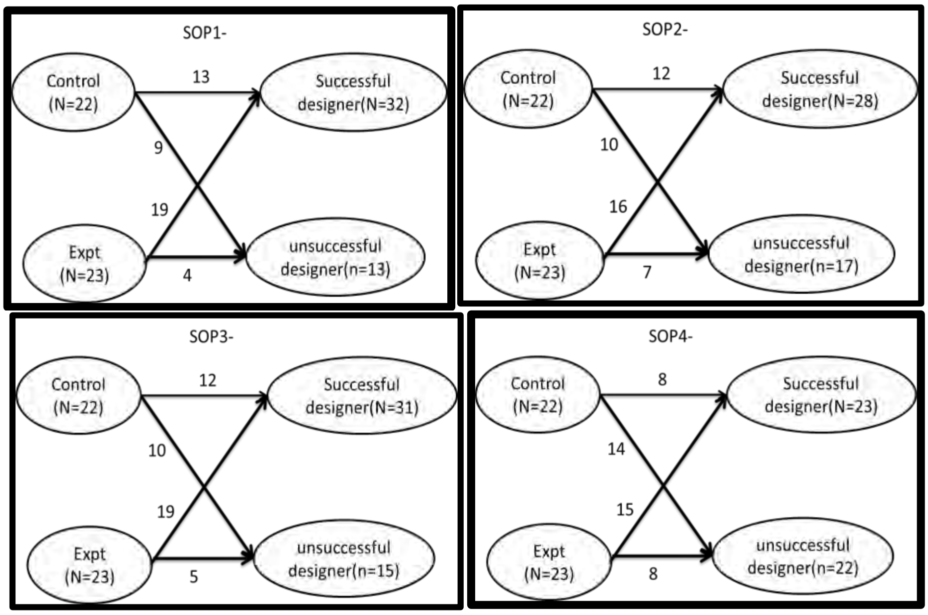


Fig.3. Stratified Attribute Tracking Diagram for successful and unsuccessful design

We found that for the sub-competencies of 'identify specifications in open problem' (SOP1), 'use specifications to structure problem' (SOP2) and sequence steps of design process (SOP3), more number of students fall in successful designer category than the unsuccessful category for both control and experimental groups. Since both groups worked with the basic design thinking activities in EDIVs, this confirms our previous results that EDIVs are useful to develop these sub-competencies. To examine the role of self-assessment rubrics, we compared the number of students in the successful designer category from the control and experimental groups. We found that majority of students in the successful designer category are from experimental group (e.g. 19-expt. group, 13-control group for SOP1). For the sub-competency of 'write structured problem statement' (SOP4), we found that equal number of students lie in successful and unsuccessful categories respectively. But in the successful designer category, majority students were from experimental group (15) compared to control group (8). This indicates that addition of self-assessment rubrics guides students towards successful design.

Following the post-test, we conducted interviews with 5 students from the experimental group. The interview questions focused on which activities students preferred while learning with EDIV and why. Here, we show a quote from a student which indicates students' perceptions of how self-assessment rubrics helped: "*If I know what is wrong in my answer and how I can achieve score 3, it helps me in learning, as most of the time we don't know what is wrong in my answer*".

4. Discussion and Conclusion

The sub-competency of writing structured problem statement from open problem requires students to perform synthesis operation by integrating various decisions and concepts. Attainment of this sub-competency leads to the overall goal of structuring of open problem, which is a key step in the engineering design process. The self-assessment rubrics trigger the process of synthesis by providing students metacognitive scaffolds in the form of the description of the target performance as well as lower levels of performance. They prompts students to carry out formative assessment of their performance, monitor and revise their achievement level and plan their learning based on target level.

Design tasks are open ended and the development of design thinking involves complex cognitive processes. The EDIV activities such as decision making task questions, concept enforcement questions, and simulative manipulation trigger students to perform the cognitive processes involved in design thinking. Self-assessment rubrics provide students the opportunity for thoughtful reflection and improvement of their work in these activities. The rubrics help simplify the complex design tasks by providing transparent criteria of evaluation to students. In future work, we will include self-assessment rubrics as an integral part of EDIVs to help students to engage in system thinking and to monitor the essential cognitive processes of design thinking.

References

ABET (2014). ABET criteria for accrediting engineering programs. http://www.abet.org/eac-criteria-2014-2015/

Atman, C. J., Chimka, J. R., Bursic, K. M., & Nachtmann, H. L. (1999). A comparison of freshman and senior engineering design processes. Design Studies, 20(2), 131-152.

Aurisicchio, M., Ahmed, S., & Wallace, K. M. (2007). Improving design understanding by analyzing questions. In ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 307-317). American Society of Mechanical Engineers.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in education, 5(1), 7-74.

Chen, Y. L., Hong, Y. R., Sung, Y. T., & Chang, K. E. (2011). Efficacy of simulation-based learning of electronics using visualization and manipulation. *Educational Technology & Society*, 14 (2), 269–277.

Crain, R. W., Davis, D. C., Calkins, D. E., & Gentili, K. (1995). Establishing engineering design competencies for freshman/sophomore students. In Proceedings of Frontiers in Education Conference, Vol. 2, 4d2-1.

Davidowitz, B., & Rollnick, M. (2003). Enabling metacognition in the laboratory: A case study of four second year university chemistry students. Research in Science Education, 33(1), 43–69.

Duston, A. J., Todd, R. H., Magleby, S. P., & Sorensen, C. D. (1997). A review of literature on teaching engineering design through project-oriented capstone courses. Journal of Engineering Education, 86(1).

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. Journal of Engineering Education, 94(1), 103-120.

Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., Rosengrant, D. & Warren, A. Developing and assessing student scientific abilities. Physical Review Special Topics, Physics Education Research, 2, 020103.

Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R. and Hmelo-Silver, C. E. (2010). Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories. Journal of the Learning Sciences, 19(1), 54 – 98.

Gresch, H. & S. Bögeholz, (2012). Identifying non-sustainable courses of action: A prerequisite for decision-making in education for sustainable development. Research in Science Education, 1-22.

Hmelo, C. E., Nagarajan, A., & Day, R. S. (2002). It's harder than we thought it would be: A comparative case study of expert-novice experimentation. *Science Education, 86*(2), 219–243.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. Science Education, 87(4), 517–538.

Majumdar, R. & Iyer, S. (2014). Using stratified attribute tracking (SAT) diagrams for learning analytics. IEEE 14th International Conference on Advanced Learning Technologies (ICALT), 7-10 July 2014.

May, E. & Strong, (2006). D. S. Is engineering education delivery what industry requires? The Third CDEN Design Conference; Toronto, Ontario, Canada; 2006.

Mavinkurve, M., Murthy, S. (2012). Interactive Visualizations to teach design skills. 20th International Conference on Computers in Education, ICCE 2012, Singapore.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. International encyclopedia of education, 2.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. Journal of the Learning Sciences, 13, 273–304.

Sheppard, S. D. (2003). A description of engineering: An essential backdrop for interpreting engineering education. Proceedings (CD), Mudd Design Workshop IV, Claremont, Cal.: Harvey Mudd College.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. International Journal of Computer-Supported Collaborative Learning, 5(1), 43–102.

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: research and development of an online learning environment for collaborative scientific discovery learning. Computers in human behavior, 21(4), 671-688.

Wood, K. L., Jensen, D., Bezdek, J., & Otto, K. N. (2001). Reverse engineering and redesign: Courses to incrementally and systematically teach design. Journal of Engineering Education, 90(3), 363-374.

Xun, G. E., & Land, S. M. (2004). A conceptual framework for scaffolding III-structured problem-solving processes using question prompts and peer interactions. Educational Technology R&D, 52(2), 5-22.

The Application of QR Codes in Outdoor Education Activities: Practice and Discussion

Wen-Shian LEE a, Chun-Yen Chang a*, b, c

a *Graduate Institute of Science Education, National Taiwan Normal University, Taiwan*
b *Science Education Center, National Taiwan Normal University, Taiwan*
c *Department of Earth Sciences, National Taiwan Normal University, Taiwan*
Address for correspondence: Prof Chun-Yen Chang, Professor and Director, Science Education Center and Graduate Institute of Science Education and Department of Earth Sciences, National Taiwan Normal University, No. 88, Section 4, Ting-Chou Road, Taipei 11677, Taiwan. Email: changcy@ntnu.edu.tw.

Abstract: We developed a mobile study program based on a cognitive theory of multimedia learning (CTML) and designed a novel outdoor mobile learning system (QR Code Information Greenmap, QRCIG). Learners use smart phones or similar devices equipped with decoding software to decode QR codes embedded in a Green Map that deliver images and voice assistance to learners about relevant content. We observed and measured 120 fifth and sixth grade elementary school students' outdoor activities and analysed the data by using the Technology Acceptance Model (TAM). The results demonstrated that students' behavioral intention to use and attitudes regarding the QRCIG were significantly positive (correlation coefficient = 0.597**). Moreover, according to the analysis, the greater the skill of the students in using the QR code was, the higher their behavioral intention to use and frequency of adoption were (correlation coefficient ranging from 0.341 to 0.852). By conducting analysis using the TAM, we proved that an education program integrated using wireless mobile devices and message image (QR code) is highly applicable in outdoor education.

Keywords：CTML, QR Code, Outdoor education, Green Map

1. Introduction

Using surrounding resources to educate students for multi-learning purposes is a popular outdoor education method today (Elma & Martin, 2010). The Green Map System has been developed collaboratively since 1995, and the system has been extended to more than 65 nations and 845 cities. In this system, simple icons are used to help people understand their surrounding environment and cities (http://www.greenmap.org/greenhouse/en
/about). This system is effective for deploying local resources and provides an alternative channel to outdoor education. However, outdoor environments are more complex than classrooms are. Information technology is a valuable tool for students; therefore, integrating information technology with outdoor education activities is expected to be a future trend (Shavinina, 2009). The mobile learning pattern constructed in technology-based activities involves urgency in occasional learning, an initiative nature in knowledge acquisition, and flexibility in space. Hence, learners must initiate efforts to learn their surroundings by using ubiquitous resources to improve their learning efficiency. This pattern is suitable for facilitating lifelong learning and developing flexibility (Chen, Kao, & Sheu, 2003). Smart phones and tablets have great potential because millions of people use these devices. Moreover, these devices can be used extensively in behavioral investigation, testing, and experimentation (Miller, 2012). Because mobile learning devices have wireless connection capabilities, learning is mobile and can occur at any location, from classrooms to outdoors. Therefore, valuable information can be acquired at any location (Sharples, 2000). As a result, the purpose of this study was to develop an integrated system composed of QR codes, Green Maps (GMs), mobile kits, and wireless technology. By using the integrated system, students enjoyed audio-visual guidance outdoors, and educators maximised the positive effects of texts and materials.

1.1 Multi-learning styles in Outdoor Education: The application of Green Map

In the United Nations Decade of Education for Sustainable Development (2005-2014), UNESCO Education Sector (2005) indicated that many political authorities worldwide are aware of the value of educating citizens on sustainable development. These authorities ask their civil servants to teach climate change and related environmental issues in the school. They have held numerous conferences and seminars to train educators about the project. Therefore, it is expected that the environment can be preserved through implementation of these ideas in national policy and education curricula. In addition, through a project named Carbon School, the European Union (EU) connected scientists, teachers, and students and encouraged them to use obtainable resources in conducting research on climate change and its influence on the earth, and to take appropriate action (Elma & Martin, 2010). This type of inquiry and project-based learning method helped children engage in healthy interaction with the earth and develop correct perspectives about the environment. This method is expected to enable the next generation to be more fervent and innovative regarding environmental protection than the current generation is.

Outdoor education is an interdisciplinary activity comprising multiple interactions. Educators must focus on diversification, targeted content, and methodology. Teachers must have practical experience to become competent in understanding environmental parameters (Pleasants, 2007; Thorburn & Allison, 2010). Many studies have indicated that the effective practice of technology use can stimulate students' motivation, help them write notes, and enforce their independence in self-learning (Rogers et al., 2005; Chen, Kao, Yu, & Sheu, 2004). Moreover, the use of mobile devices can broaden students' experiences in outdoor education (Rogers et al., 2004; Squire & Jan, 2007). Outdoor education does not only involve improving knowledge, but also involves gaining experience through the process (Rickinson, 2001). Therefore, only when educators have a clear vision and teaching materials can they provide outstanding education (May, 2000). This is the main goal of educators today. They are expected to use new technologies for curriculum design and development.

Zuber (1999) and Green and Swanitz (1991) demonstrated that GMs are a suitable material for stimulating students' motivation to learn about green issues. A GM is a map that promotes nature, sustainability, and ecology (please refer to http://www.greenmap.org/greenhouse/files/gms/GreenMapIcon_V3_Chart.pdf). In a GM, the relationship between people and the environment is clearly marked and illustrated, which makes it an effective material for enabling inhabitants and students to engage in environmental exploration (Green Map Activity Guide, 2005). GMs are widely applied in outdoor education, and by the end of 2012, 842 registered Green Map projects were conducted in 68 nations (http://www.greenmap.org/greenhouse/en/node/12842). Therefore, we used GMs as teaching materials for outdoor education and embedded QR Codes into the program, thereby facilitating ubiquitous learning.

1.2 Design of M-learning on Environmental Education

Today, the development of GMs is nearly mature; therefore, students experience little difficulty in using and learning a GM in ordinary maps. We used this advantage to digitalize and apply QR codes to the map, enabling students receiving outdoor education to quickly identify links to information. This application is suited to the current educational goal of governments worldwide, which is to develop the ability of citizens who have basic skills in information technology use (Hiltz & Turoff, 2005; Lim & Kim, 2003). QR codes are two dimensional, the information capacity is 1,000 words, and images and signals can be contained without the limitations of sizes and colours. If the code image is damaged and the damaged area does not exceed 30% of the original image, then the information is retrievable. QR codes have been popular worldwide for years (Lai, Chang, Li, Fan, & Wu, 2012). Accordingly, we adopted a cognitive theory of multimedia learning (CTML) to integrate information into m-learning in outdoor education. Mayer (2001) stated that multimedia information designed based on a mental mechanism provides more meaningful messages to learners than multimedia messages that are not based on a mental mechanism do; therefore, if multimedia tools are involved in a responsive learning environment as a primary tool, then students' understanding is promoted. This wireless learning information system provides appropriate immediate learning at any location. People can use wireless devices equipped with a digital camera to decode QR codes and capture multimedia information at their present location. Quinn (2000) indicated that m-learning, which is defined as learning that occurs through mobile devices, is suitable for education in numerous places.

We used QR codes to design a service that facilitates learning that is not limited by space or time. Learners can access digital information and teaching materials and learn from this method. Asynchronous learning was (Chen et al., 2003) integrated with the Internet and QR codes, enabling learners to acquire necessary information in a short period of time asynchronically. As shown in Figure 1, digital content was linked through decoding and wireless connections and transferred to the mobile devices of learners.

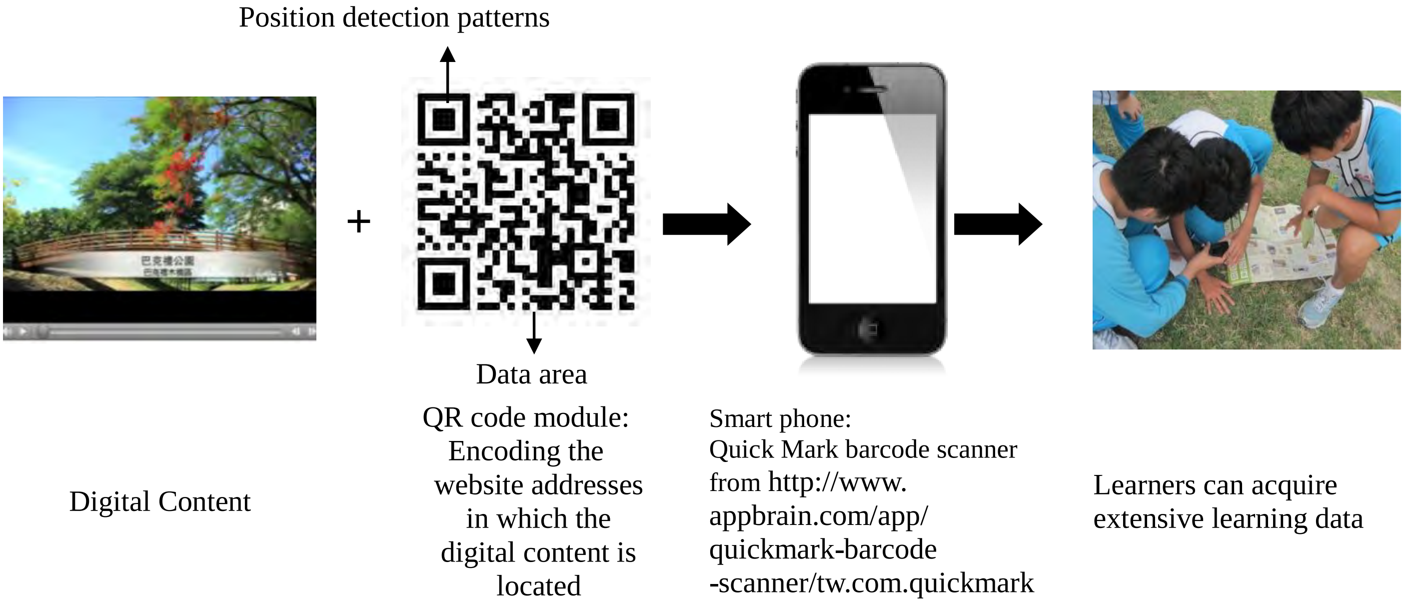


Figure 1. Digital content accessed using mobile devices facilitates achieving the goal of asynchronous learning

As shown in Figure 1, digital information is decoded and connected to the web server, and the existing information is subsequently presented to learners. The official QR code website indicated (http://www.qrcode.com/en/index.html) that the three squares around the four corners are designed to help decoders relocate; therefore, they do not need to scan it precisely. We used QR codes to design a service that is not limited by time or space, providing learners digital information and learning materials and helping them to acquire knowledge. Asynchronous m-learning (Chen et al., 2003) was integrated with the internet to enable users to access fruitful multimedia information. The process is quick and precise, and the acquisition is immediate.

The innovation involved in the present study is the integration of natural science and social science. Integrating a GM with QR codes improves it compared with the original text version. Furthermore, voice and animation were added to the scenic introduction and embedded in the digital platform. This system must be supported by a wireless network and a mobile device with a decoding function (Quick Mark barcode scanner, from http://www.appbrain.com/app/quickmark-barcode-scanner/tw.com.quickmark). When the device decodes a QR code, the service provides animation or voice guidance through the Internet. This service is called the QR Code Information Greenmap (QRCIG). The structure of the QRCIG is illustrated in Figure 2.

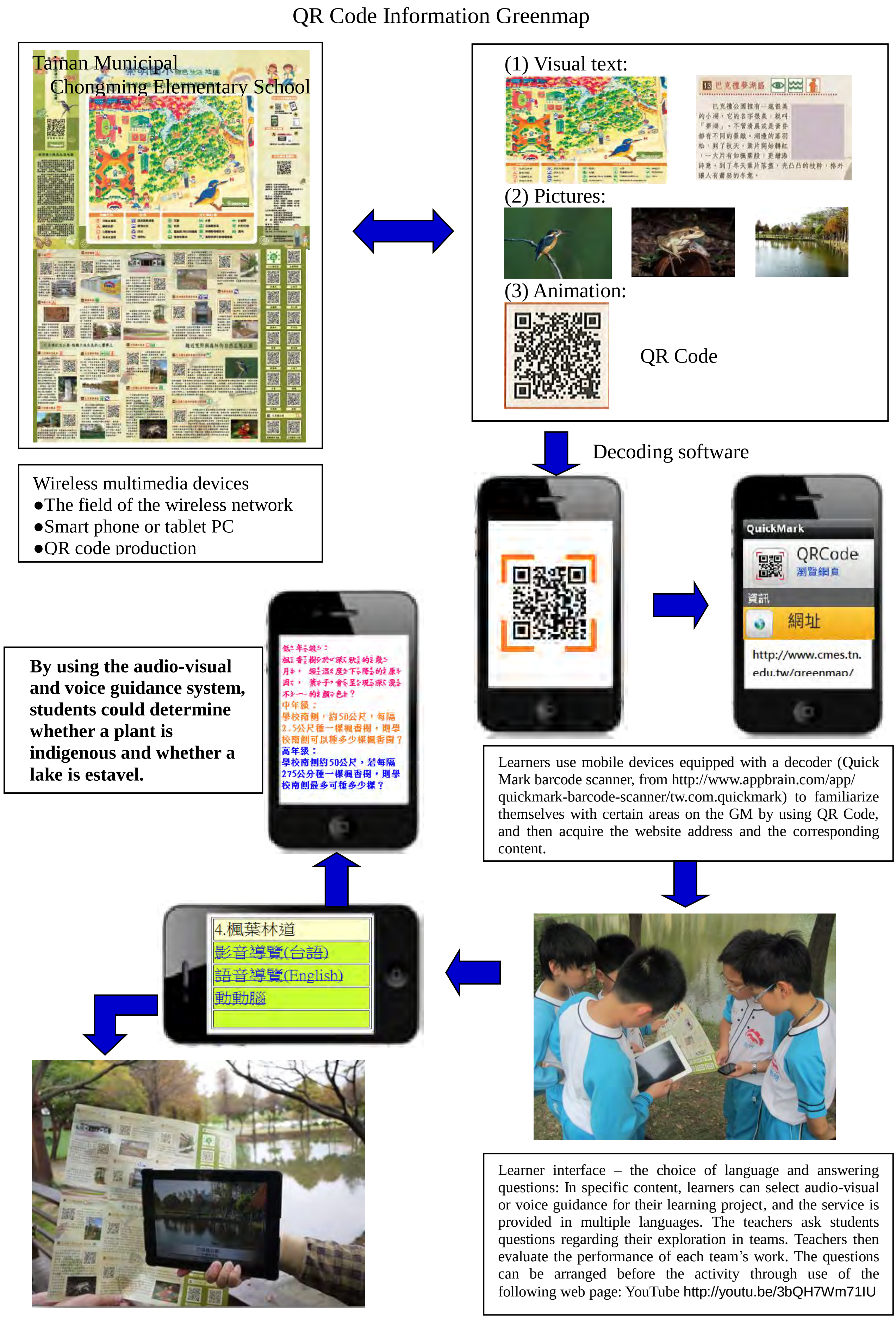


Figure 2. Integrated QR Code Learning System Architecture

Mayer (2001) and Mayer & Moreno (2002) stated that students acquire more information and achieve more positive learning goals when viewing animation presented using voice-overs than when viewing dubbed animation without voice-overs. Based on this previous study, we adopted the modality principle from the CTML to integrate text, images, and animation. The methodology is explained as follows:

A. Text: We used a universal icon system to signify ecological and cultural locations on the map. The locations were then verified by fieldworkers, and descriptions are presented together.
B. Image: Photographs captured during fieldwork are presented as a reference for users when they are reading the text.
C. Animation: Films captured during fieldwork are used to assist students in their learning, which is supplemented by oral explanation. The audio-visual guidance was linked using the QR code platform.

The procedure of the aforementioned system is introduced as follows:

A. School teachers, advisory professors, and association members from societies participated in the procedure, and then designed the system based on the majority's view point, producing a GM with icons.
B. We introduced icons in GM scenic spots by using audio-visual information, making them audio-visual materials.
C. We input the audio-visual materials into Multimedia U-learning platform server and assigned each datum corresponding QR codes.
D. We positioned the encoded QR codes at the corresponding locations on the GM.

The pre-study (Lai et al., 2012) showed that teachers positively evaluated the application of audio-visual guidance in presenting text, images, and voice functions simultaneously. In that study, the coefficient factor modulus between the ease of use and usefulness of the QR codes was 0.749**. This result demonstrates that the high ease of use of the QR code application triggered teachers' perception of the usefulness of the system, and the coefficient factor modulus between the usefulness of the QR code and people's acceptance of using QR codes was as high as 0.872**, which indicates a highly positive correlation between QR code usefulness and QR code acceptance. The analysis also demonstrated that teachers who were skilled at using QR codes tended to have higher behavioral intention to use and acceptance than those who were not. This indicates that QR codes can be accepted by teachers as a tool for use in outdoor education activities. Because the analysis of teachers' perceptions of the program were well conducted and the results proved to be positive, the present study focused on students in outdoor education.

2. Methodology

We adopted a TAM pattern to investigate behavioral intention to use towards using and acceptance of the method. The TAM is a tool developed by Davis (1989) based on behavioral theory (TRA) that is used for interpreting cognition and emotion parameters that affect the use of technology. TAM can be used to inspect people's inclination towards new technology. The pattern was developed to integrate the simplest methods of applying research theory and provides insight into personal exploration. Therefore, researchers can observe the factors affecting people's beliefs, attitudes, and intentions when they are using technology (Davis, Bagozzi, & Warshaw, 1989). Because of this advantage of easiness, we infer that the TAM pattern is the pattern adopted most frequently internationally when researchers are examining people's acceptance of technology (Davis, 1989; Venkatesh & Morris, 2000). Davis (1989) indicated that when users perceive accessing certain information to be simple, they tend to feel positive about the system. When users positively evaluate certain information regarding its cognitive usefulness, their positive attitude towards that information is higher. Furthermore, when users more highly evaluate the cognitive ease of use of certain information, they tend to more positively evaluate its cognitive usefulness.

We administered a questionnaire in which the answer is graded based on a 4-Point Likert scale. The content is divided into seven categories, namely, "personal information," "self-estimation of information accessibility," "preference investigation of the QRCIG and conventional GM," "ease of use of the QRCIG," "usefulness of the QRCIG," "attitude towards the QRCIG," and "behavioral intention to use regarding the QRCIG." The Cronbach's alpha value was 0.925, indicating that the questionnaire is highly reliable. The questionnaire was designed to investigate whether students understand that the function of QR codes could affect the application of the QRCIG in outdoor education and to evaluate the behavioral intention to use of users regarding the QR codes. The questionnaire was also designed to determine whether students perceive the QRCIG as easy to use (ease of use), which enables students to use the system to achieve learning goals and improve their knowledge acquisition (usefulness). Finally, the questionnaire assesses whether it can assist the design of the QRCIG and promote students' intention to use this system in the future, and analyses their adaption of it.

3. Results

This section presents the responses observed in the experiments under different investigations and the corresponding samples. The observations were meticulously analysed and explained. Among the 123 returned questionnaires, 120 copies were valid and 3 were invalid. The section contained three subsections for different explanations. In the first subsection, different backgrounds are used as parameters to discuss the ease of use and usefulness of, as well as the attitudes towards, the QRCIG. In the second subsection, students' preference for either the QRCIG or conventional Green Maps was used as the parameter to discuss differences in preference. In the third subsection, the correlations among information accessibility, ease of use, usefulness, and attitudes regarding the QRCIG were analysed.

3.1 Different backgrounds as the parameter for the discussion of ease of use, usefulness, and attitudes regarding the QRCIG

3.1.1 Descriptive statistics and t test results of the perceived ease of use and usefulness, and the attitudes regarding the QRCIG of students who used QR codes prior to the study

Table 1: Descriptive statistics and t test results of the perceived ease of use and usefulness, and the attitudes, of students who used QR codes prior to the study.

Category	Mean	Standard Deviation	*t* test
Ease of Use	18.30	2.30	1.936*
	17.26	2.50	
Usefulness	18.40	1.99	2.331*
	17.22	2.39	
Attitude	14.38	2.03	2.378*
	13.33	2.00	
Behavioural intention to use	14.33	2.04	1.812
	13.52	2.12	

$^{*}p < .05$

According to questionnaire analysis revealed there were 93 students had been read QR Code icon, it was over 1/4 of all samples. As shown in Table 1, whether students used QR code or related functions prior to the study is an important parameter that significantly affected the perceived ease of use and usefulness, as well as their attitudes, regarding the QRCIG. This indicates that QR code knowledge affects students' perceptions of the ease of use(*t* test =1.936*), usefulness(*t* test =2.331*), and attitudes(*t* test =2.378*) regarding the QRCIG.

3.2 Students' preference for either the QRCIG or conventional green maps as the parameter

Table 2: Frequency distribution of the preference of the QRCIG or conventional green maps.

Item	Which item do I prefer to acquire knowledge?	Which method do I prefer to reach the goal?	Which learning method is better for me?
QRCIG	115(95.8%)	116(96.7%)	111(92.5%)
Conventional Green Map	5(4.2%)	4(3.3%)	9(7.5%)

As shown in Table 2, students received both conventional Green Maps and the QRCIG as the learning methods used to reach the goal of outdoor education, and over 90% preferred the QRCIG as the method for their education. This demonstrates that the design of the QRCIG promotes students' intention to learn.

3.3 Correlations among information competence, ease of use, usefulness, and attitudes towards the QRCIG

Table 3: Comparison of information accessibility, ease of use, application, attitudes, and behavioral intention to use.

Category	Information Accessibility	Ease of Use	Usefulness	Behavioural Intention to Use
Ease of Use	0.361**	--		
Usefulness	0.404**	0.785**	--	
Attitudes	0.341**	0.641**	0.611**	--
Behavioural Intention to Use	0.405**	0.757**	0.712**	0.852**

$^{*}p < .05$, $^{**}p < .01$

As shown in Table 3, the Pearson correlation coefficients of students' information accessibility, ease of use, usefulness, attitudes, and behavioral intention to use were between 0.341 to 0.852, indicating that students are responsive to information accessibility, ease of use, usefulness, attitudes, and behavioral intention to use regarding QR codes at medium to high level.

4. Discussion

We introduced QR-coded learning to an outdoor education m-learning program. The project inspired curriculum design, methodology design, students' perception, and students to learn through various ways. These inspirations helped both teachers and students to practically explore the curriculum and interact positively with the environment. This contribution satisfies the goal of diversification in education. The present study is based on Mayer's proposal (2001) that voiced-over multimedia information is easier for students to understand than dubbed multimedia information is. In addition, we used the TAM for system analysis and found that students are interested in QR-coded mobile learning in outdoor education.

Lai et al. (2012) used QR codes in Green Maps based on the CTML and designed a multimedia learning system. Their results showed that teachers expressed highly positive attitudes towards applying the QRCIG in outdoor education. Therefore, to investigate the inclination of students to use this system in outdoor education, we used information accessibility of student as external variables in an analysis of fifth and sixth grade elementary students' questionnaire feedback after they participated in activities based on the TAM.

The present study demonstrated that, among two learning models, namely, the QRCIG and conventional green maps, more than 90% of the students preferred the QRCIG as the method for attaining the goal of outdoor education. In addition, the m-learning QRCIG was integrated with outdoor education and the parameters were students' backgrounds. "Whether students used QR codes or related function prior to the study" was the main parameter affecting ease of use, usefulness, and attitudes. This parameter is crucial for learning how to use the QRCIG for elementary school students. Moreover, ease of use (Pearson correlation coefficient = 0.361**), usefulness (Pearson correlation coefficient = 0.404**), future behavioral intention to use (Pearson correlation coefficient = 0.405**), and the capability students had before they used the QRCIG were also crucial parameters. These results supported Davis' statement (1989) that, when users' recognition of perceived ease of use of information is high, they tend to hold positive attitudes towards the information system. When the attitude is positive, users tend to hold positive attitudes towards the usefulness of the information.

A TAM pattern derived from information accessibility, ease of use, usefulness, attitudes, and behavioral intention to use perceived by students regarding the QRCIG by using path analysis, as shown in Figure 3. According to Figure 3, the Pearson correlation coefficients indicating students' perceived ease of use and usefulness, attitudes, and behavioral intention to use regarding the QRCIG ranged from 0.341 to 0.852, indicating that these parameters are highly relevant at a medium to high level. This result is consistent with that of Mayer & Moreno (2002), who studied multimedia aids.

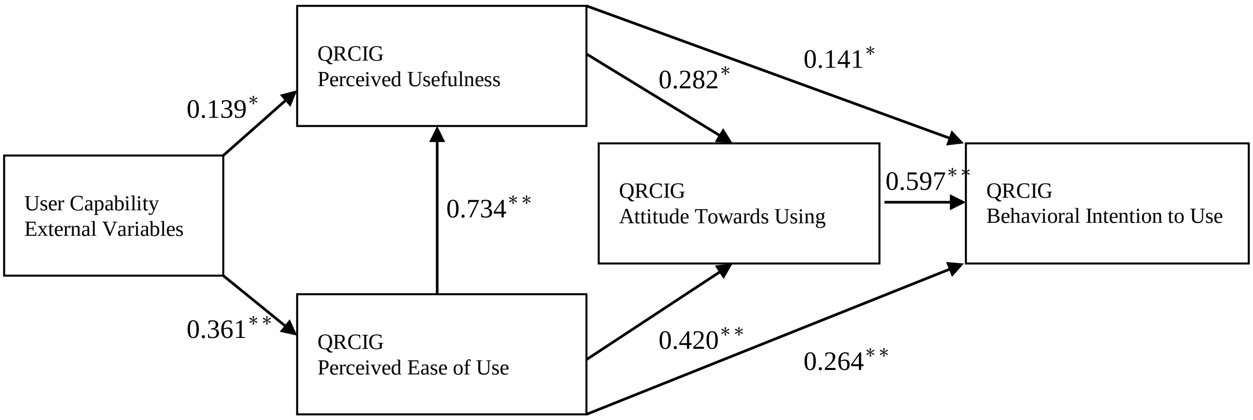


Figure 3. Technology Acceptance Model for the QRCIG

In the TAM, each route carries high significance. The correlation coefficient between perceived ease of use and usefulness of the QRCIG was as high as 0.734**, indicating that if students' perception of ease of use towards the information system is high, then their perception of its usefulness is positive.

Moreover, the correlation coefficient between the attitude towards using and the behavioral intention to use of the QRCIG system was 0.597＊＊, indicating that the higher the attitude towards using of the QRCIG was, the more positive people's perception of the system was. Furthermore, by using path analysis, correlation coefficients are significant in both of perceived usefulness and perceived ease of use of the QRCIG with the attitude towards using, which means that the higher the perceived usefulness of QRCIG and perceived ease of use were, the more willing students' attitude towards using QRCIG was. Several scholars who used the TAM demonstrated that "behavioral intention to use" can be used as a predictive force (Mathieson, 1991; Szajna, 1996). Our results illustrate that the tools used in the present study triggered high perceived ease of use and perceived usefulness among subjects, causing them to hold positive attitudes about their intention to use the system.

5. Conclusion

We designed a learning system integrating information technology and outdoor education; this system inspires novel ideas about curriculum design, pedagogy, and the learning of students. Moreover, the system links teachers and students and encourages students to use various methods to achieve educational goals and to have positive interactions with the environment. The analyses demonstrated that the tools developed in the present study satisfy the principles of recognition and integration. The TAM analysis shows that users' behavioral inclination determines their inclination towards specific information technologies, which indicates that users who participated in the present system attained positive learning outcomes and positively evaluated the system.

References

Chen, Y. S., Kao, T. C. & Sheu, J. P. (2003). A Mobile learning system for scaffolding bird watching learning. Journal of Computer Assisted Learning, 19, 347-359.

Chen, Y. S., Kao, T.C., Yu, G. J., & Sheu, J. P. (2004). A Mobile Butterfly-Watching Learning System for Supporting Independent Learning. *Proceedings of the 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04).*

Davis, F. D.(1989) .Perceived Usefulness, Perceived Ease of Use, and User Information Technology, *MIS Quarterly, 13*(3), 319-339.

Davis, F.D., Bagozzi, R.P., & Warshaw, P.R. (1989), User Acceptance of Computer Technology: A Comparison of Two Theoretical Models, *Management Science*, Vol.35, No.8, pp.982-1003.

Elma, D. & Martin, G. (2010). Evaluation of CarboSchools: Students', teachers' and scientists' opinions on authentic science projects in a European context. Retrieved on February 12, 2012 at http://www.carboeurope.org/education/FinalEvaluationReport.pdf

Green map activity guide (2005). Green Map Japan, Inc. Retrieved April, 1, 2010, from http://www.greenmap.org/greenhouse/en/node/1467

Green, H. & Swanitz, J. (1991). *Coast to coast virtual road trip. 9–11th grade lesson. Schools of California Online Resources for Education (SCORE): connecting California's classrooms to the world.* San Bernardino County: Schools of California Online Resources for Education.

Hiltz, S. R., & Turoff, M. (2005). Education goes digital: The evolution of online learning and the revolution in higher education. *Communications of the ACM*, 48 (10), 59-64.

Lai, H. C., Chang, C. Y., Li, W. S., Fan, Y. Lin., & Wu, Y. T., Lin, M. C. (2012). The Implementation of Mobile Learning in Outdoor Education: Application of QR codes. *British Journal of Educational Technology*.

Lim, D. H., & Kim, H. (2003). Motivation and learner characteristics affecting online learning and learning application. *Journal of Educational Technology Systems*, 31 (4), 423-439.

Mathieson, K. (1991).”Predicting User Intentions：Comparing the Technology Acceptance Model with the Theory of Planned Behavior.” *Information Systems Research*, 2(3), 173-191.

May, T.S. (2000). Elements of success in environmental education through practitioner eyes. *The Journal of Environmental Education*, *31*(3), 4-11.

Mayer, R. E. (2001). Multimedia learning. New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. Learning and Instruction, 12(1), 107-119.

Miller, G. F. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, May 2012; vol. 7, 3: pp. 221-237.

Pleasants, K. (2007). Interdisciplinary teaching through outdoor education. Australian Journal of Outdoor Education, 11(1), 53-55.

Quinn, C. (2000). M-Learning: Mobile, Wireless and In-Your-Pocket Learning. *Line Zine Magazine*. Retrieved May 25, 2010, from http://www.linezine.com/2.1/features/cqmmwiyp.htm

Rickinson, M. (2001). Learners and learning in environmental education: A critical review of the evidence. *Environmental Education Research, 7*(3), 207-320, from http://www.cemus.uu.se/dokument/asc2010/Mark%20Rickinson.pdf

Rogers Y., Fraser D. S. & Fitzpatrick G. (2005). Ubi-learning integrates indoor and outdoor experiences. *Communications of the ACM, 48(1),* 55-59.

Rogers, Yvonne, Price, Sara, Fitzpatrick, Geraldine, Fleck, Rowanne, Harris, Eric, Smith, Hilary, Randell, Cliff,Muller, Henk, O'Malley, Claire, Stanton, Danäe, Thompson, Mark and Weal, Mark J. (2004) Ambient Wood: Designing New Forms of Digital Augmentation for Learning Outdoors. In, *Third International Conference for Interaction Design and Children (IDC 2004),* Maryland, USA, 01 - 03 Jun 2004. ACM Press, 1-9

Sharples, M. (2000).The design of personal mobile technologies for lifelong learning. *Computers and Education, 34*, 177-193.

Shavinina, L. V. (2009). High intellectual and creative educational multimedia technologies for the gifted. In L.Shavinina (Ed.), *International Handbook on Giftedness.* Toronto, ON: Elsevier.

Squire, K. D., & Jan, M. (2007). *Mad City Mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers.* Journal of Science Education and Technology, *16*(16), 5-29.

Szajna, B. (1996).”Empirical Evaluation of the Revised Technology Acceptance Model” *Management Science* 42(1): 85-92.

Thorburn, M. & Allison, P. (2010). Are we ready to go outdoors now? The prospects for outdoor education during a period of curriculum renewal in Scotland. *The Curriculum Journal, 21*(1), 97-108.

Venkatesh, V. & M.G. Morris. (2000). “Why don't men ever stop to ask for directions? gender, social influence, and their role in technology acceptance and usage behavior,” *MIS Quarterly*, vol. 24, no. 1, pp. 115-139.
Zuber, W. (1999). *Green Mapmaking. Green Teacher*, n. 58, pp. 6-11.

The Application of Instructional Media and IRS in Environmental Education - Focus on the Rocky Terrain in Northern Coast of Taiwan

Wen-Mao Chung [a], Te-Shin Tsai [b] & Chun-Yen Chang [c*]
[a] *Graduate Institute of Science Education, National Taiwan Normal University, Taiwan*
[b]*Keelung Municipal Dung Xin Elementary School, Taiwan*
[c] *Science Education Center, National Taiwan Normal University, Taiwan*
*changcy@ntnu.edu.tw

Abstract: Based on the rock and mineral in environmental education, this paper designs instructional media, with the aim to let the users realize the beautiful sceneries on the coast around Taiwan and learn about the rocky terrain at the same time. Using the IRS system, the author tries to make the computer classes for interaction between individual and personal computer, including the lecture-listen teaching model, as well as interaction between teacher and student which could provide a different teaching style. This research shows that this instructional media enables six grade students in primary school learning the relative knowledge about the rocky terrain in northern coast of Taiwan happily and effectively, while the employment of IRS could enable students concentrate more attention to their study and promote their learning interest as well.

Keywords: environmental education, instructional media, IRS, science education

1. Introduction

Taiwan is a seagirt and has graceful and magnificent coastline. These abundant marine resources are precious gifts from nature. Therefore, cultivating students with experience about the marine ecological environment around; promoting their knowledge and strengthen their awareness of protection about marine ecological environment is really significant.

In the respect of instructional media design, it mixes nature, life, science, technology and society together. Thus, students may learn the basic rocky styles, rocky characteristics, and rocky distribution in Taiwan, which could enable students further acquaint the enrichment, diversification and beauty of rocky terrain of northern coastline in Taiwan, and experience a variety of grotesque rocks shaped by various miraculous natural power. Hence, the students would concern about the changes of seashore rocks and ecological environment so as to protect and cherish the rocky mystery which was formed tens of thousands of years. According to the curriculum guidelines for science and life technology (MOE, 2008), fifth- and sixth-graders in primary schools are able to learn about the formation and characteristics of rocks. Matching the available nature and cultural characteristic in the local environment, these two kinds of learning fields integrate closely. The author hopes that through learning and understanding the different formations of seashore rocks and terrain landscape, students will arouse the concern of changes in marine geographic environment, and then show solicitude for local marine geographic ecological environment.

With the development of science and life technology, application of mobile devices in education could be seen more frequently (Ellaway, 2013). Although it brings lots of convenience for teachers and students, in the current situation of education in Taiwan, the popularizing rate of mobile devices are still lower than the computer class undoubtedly. With the updating of IRS, IRS could be operated in desktop in the way of browser. Except the flash interactive multi-media, this research relies more on the advantages of IRS and plus the section of interaction between students and teachers, not individual-machine only, to complement the function of record for students' answering condition. Thus, it makes the interaction between teachers and students more comprehensive and more appropriate for teachers to know the learning condition of students.

2. Literature Review

2.1 Theoretical Basis of Instructional Media in Aided Teaching

Instructional media has been wildly spread for its advantages, such as drive learning motivation (Liu et al, 2011) increase attention (Sun, 2014), protect and update the textbook easily, and so on. The theoretical basis of teaching by media can be divided into 2 parts:

ZPD theory, suggested by Vygotsky, states that people have 2 levels of development, one is practical development level, and the other is potential development level. Practical development level is what Piaget so-called children developmental stage, that is, possess abilities in its period; and potential development level is the ability to deal with some difficulties with the help of parents or partners. The gap between the two is called Zone of proximal development (ZPD). While students are learning something, the teachers just provide with a temporary support to give assistance to students transform from practical development level to a potential development level, which is called scaffold in deed. This scaffold is a kind of instructional strategy or teaching aid. Along with the improvement of student ability, the dominant right of learning will return to students gradually. At this time, even students don't get the help of scaffold could they get knowledge and skill successfully.

In addition, Gardner has put together multiple intelligence theories, which emphasize that everyone has entire intelligence and its exclusiveness. Gardner states that educators should do what they can to dig out the learning styles and trend of students, and then take good advantage of them to forge a most suitable education style. Multimedia integrates with film, cartoon, voice, script, pictures and some other materials and could be operated, listened and viewed for another, which enable good learning tools for different intelligent students.

2.2 Usage of IRS

IRS, Interactive Response System, is a kind of teaching application system used for feeding back information for teachers in classes through mobile devices and electrical tools. It is one of the most important information application apparatuses for improving classroom teaching quality in recent years and wildly used in classroom teaching in Europe and America. Taiwan has carried out quite a few experimental curriculums and interactive rooms in recent years. However, what is the interactive room? The so-called interactive room installs IRS acceptors in the classroom, and the teachers and students have a remote control in hands in order to continue the interactive teaching activities which but also reduces the burden on teachers (Huang et al, 2001). IRS has benefits for teachers and students both as below (Liu et al, 2006):

For students:

(1) The motivation of active participation: provides prompt visualizing feedback for the response of questions from students, and further strengthens the motivation of active discussion.

(2) Enables students to focus on learning content: students should choose the best answers aims to the questions which could help students thinking about the problem purposefully.

(3) Assist students to conduct further conceptual understanding: students should tell the reason for your answer so as to promote students' thought on inner nature.

For teachers：

(1) Assists teachers in diagnosing learning status and offers decision references: the system would help to collect the answers from students; this allows teachers to find out the learning problems of students immediately and allows teachers to make some evaluations of the condition, so that they could clear out the vague concept and enlighten subsequent discussion.

(2) Improves the interaction between teachers and students: it allows all students to possess an equal learning chance by assisting teachers to control the equivalence while communicating.

(3) Improves the frequency of teaching: it allows students to focus on learning points so as to save teaching time, and make class colorful and interesting but in control at the same time.

IRS has both numerous advantages, and disadvantages. For example, IRS is not the main part in class; it can't take the place of teachers. Appropriate leading and in-depth feedback from teachers will result in in-depth study (Wang et al, 2002). IRS has the feature of in-time interaction which can inspire the interest, but may also lead to over excited students. Thus, IRS needs to be guided and managed professionally to keep a proper teaching environment.

3. Methodology

3.1 Objects of Study

This research based on purposive sampling method, and selected 118 sixth-grade students in 2 primary schools from northern Taiwan and 112 students as the valid sample size. It should be implemented in flexible classes by matching science and life technology curriculum after the normal class Rock and Mineral.

3.2 Instructional Media Design

Instructional media was made by flash and matches Nature and Life class in 6th grades; it contains: (1) various kinds of rocks with different characteristics and recognizable peculiarities. (2) changes of rocks effected by weathering, erosion and deposition. (3) present rocky terrain in coastline and flourish students to cherish and protect and appreciate geological environment in coastline. (4) describes district and regional marine rocky terrain characters. Here are the operating instructions of media as follows:

	
This is the title. Put on different buttons will help you with different learning travels.	This page is the distribution of sedimentary rocky coast in northern Taiwan, and displays different coast appearances in north and northeast with live-action.
	
This is the introduction of igneous rock with 3 buttons for different styles; pressing on the mouse will link the student to different pages where you can learn about different kinds of rocks.	This page shows the rocky terrain in northern coast of Taiwan and has marked out volcanic and sedimentary rock by animated mapping. 5 buttons for learning various terrain scenery.

Figure 1: The Interface of the Instructional Media. (1)

This instructional media adopts a pattern of self-directed operation and independent study with a large number of image-text information as well as voice materials, which are from natural landscape nearby. Beside the learning level, this media includes a self-challenge, a puzzle, and a Link Game, as well as choice questions and other forms of interaction, for self-examination after classes.

	
Puzzle fun. Set the pictures into the frames. The color will become deeper and the picture can't be moved once set into the correct position. The picture in top right corner is a reference.	Link Game. The above line shows various kinds of rocks, with the names below.

Figure 2: The Interface of the Instructional Media. (2)

3.3 Application of IRS

IRS, the abbreviation of Interactive Response System, is a kind of system used for interacting with teachers and students during the classes, through computers or mobile decives. It was pointed out that IRS could increase learning interest and add more interaction in classes, which gives play to peer instruction well.

This research employed Socrative 2.0 system, a system has something to do with teaching feedback, assessment test and teaching report and could be applied to mobile decives. Moreover, Socrative 2.0 has broken the limits in different platforms, but appeared in the form of WEB. In other words, Socrative 2.0 system could be easily and conveniently operated only if the appliance has a browser in it, and internet service is available in the environment around. In the research, the teachers conduct a pre-test by IRS first, and then with guidance of it, the teachers will refer to the results after a test to shepherd students continuing discussion and in-time interaction. The students can be guided to self-learning by operate the instructional media actively. A test is conducted to test the students before the class ends. The pattern as below:

(1) To go on a pre-test by IRS.
(2) To go on class discussion and in-time interaction by IRS
(3) To go on self-learning by operate instructional media actively for students.
(4) To go on post-test and collect the attitude by questionnaires.
(5) To organize and analyze the data.

3.4 Data Compilation and Disposition

In the research, questionnaire could be divided into 2 parts. One is for learning efficiency, and it is the single choice aiming at teaching objectives. Compared with the 2 tests before and after class, it is found that after learning by instructional media, the accuracy of every single choice has risen. If the instructional media has been properly used, the learning efficiency would be improved at the same time. The data are shown in Table 1 below:

Table1: Comparison of passing rate before and after class

No.	1	2	3	4	5	6	7	8	9	10	Average
Pre-test passing rates	62.5	92.9	67.9	35.7	96.4	55.4	73.2	35.7	51.8	41.1	60.4
Post-test passing rates	87.5	100	91.1	66.1	100	89.3	94.6	69.6	82.1	66.1	83.4
Ascensional range	25	7.1	23.2	30.4	3.6	33.9	21.4	33.9	30.4	25	23

Although the passing rate has risen, the passing rates after class of question 4 and question 10 are still flat. Question 4 is "what kind of coastline does northern coastline of Taiwan from Jinshan to Beiguan should be?" The correct answer is "sedimentary rocky coastline". Question 10 is "what kind of rock does northern coastline of Taiwan from Jinshan to Beiguan should be?" The correct answer is "Sandstone". While compared with similar question concerning Danshui or Jinshan, the passing rate of reached 91.1%. Most of students didn't where the Beiguan is after inquiry. Because of the time limitation when operate the instructional media, the passing rate of such questions is lower than other questions.

The questionnaire about learning experience adopted Five-point Likert scale. The questionnaire can be subdivided into instructional media and IRS. Table 2 shows the results. In the first part, all students believe that learning by operating media is happy, and the mean could arrive at 4.71. What's more, students also think that it's conformed to their needs to operate instructional media by themselves which is suitable for their tempo and interest, so that they are able to operate several times when they couldn't understand the point. The mean could reach 4.54. In the part of IRS, most students deem that using IRS in class will drive you pay more attention to class because you have to communicate with teachers all the time. The mean reaches 4.77. Owing to this kind of interactive teaching model, you are able to put up your opinion at once and elevate your learning interest since it is full of freshness. The mean reaches 4.45. The data are shown in Table 2 below:

Table 2: The questionnaire results after the end of the course.

	Strongly Agree	Agree	Ordinary	Disagree	Strongly Disagree	**Mean**
1.I feel happy to use instructional media for learning	71.4%	28.6%	0.0%	0.0%	0.0%	4.71
2. Operating the instructional media myself meets my learning demand.	62.5%	30.4%	5.4%	1.8%	0.0%	4.54
3. This approach aimed to upgrade my attitude of learning science	30.4%	41.1%	25.0%	3.6%	0.0%	3.98
4. Using IRS during class make me feel at ease without pressure.	30.4%	33.9%	23.2%	7.1%	5.4%	3.77
5. It's fun to use IRS in class.	55.4%	35.7%	7.1%	1.8%	0.0%	4.45
6. Using IRS in class makes me concentrate better	76.8%	23.2%	0.0%	0.0%	0.0%	4.77
7. I am satisfied with my learning performance.	53.6%	28.6%	12.5%	5.4%	0.0%	4.30

It's worth noting that although using IRS in class is popular among most students, still 1/8 students feel more stressed compared with traditional teaching method. After tracing the following students, it is found that the instantaneousness from IRS puts more stress to students. They may feel anxious once they cannot answer some questions needed thinking or unable to keep pace with teachers. Some students said that the instructional words were not in Chinese, but in English, and they cannot figure out if this language barrier would put more stress on those poor English students. On the other hand, the learning scale of science and life technology would be wilder and the questions would be more difficult if students go to senior class, still 28.6% student are not favor of this subject even instructional media and IRS are used in class, which should be concerned by related educators.

4. Conclusion

While conducting correlated curriculum of environmental education, the proper usage of instructional media may help learners study happily and improve their learning efficiency if the learners are able to operate the media according to their interest and demand. In this research, the usage of IRS is conducive to attention focusing and learning interest inspiration. However, even in an age where mobile devices is prevailing, many schools are unable to provide enough mobiles for using. Fortunately, superexcellent alternative solution—computer rooms are suggested to replace the former two by each school. Utilizing WEB pattern of IRS, such as Socrative software, can allow teachers and students enjoy many benefits of IRS teaching without any mobile device or special equipment. In addition, IRS may as well assist in dealing with data. It is a kind of facilitating device. However, teachers need to focus on students' stress while using IRS, no matter on operation style or language ports. It is suggested to training teachers proficiently before using it in due form. More importantly, teachers are able to grasp the information about those who are unable to keep pace with class or have less interest in study because of the immediacy of IRS, and show their concern and consideration. The usage of IRS makes students feel excited, thus, educators should acknowledge that machine is not the main role in learning but the proper guidance for teachers. Only through proper application from teachers can the instructional media provide the most value.

References

Ministry of Education, (2008) . The Grades 1-9 Curriculum Guidelines. Taipei: Ministry of Education

Ellaway, R. (2013). The informal and hidden curricula of mobile device use in medical education. *Medical teacher*, *36*(1), 89-91.

Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research and Development*, *59*(2), 249-265.

Sun, J. C. Y. (2014). Influence of polling technologies on student engagement: An analysis of student motivation, academic performance, and brainwave data. *Computers & Education*, *72*, 80-89.

Huang, C. W., Liang, J. K., & Wang, H. Y. (2001). EduClick: A computer-supported formative evaluation system with wireless devices in ordinary classroom. *Proceedings of ICCE 2001*, 1462-1469.

Liu, T. C., Chu, C. C., Lin, Y. C., & Wang Y.(2006). A preliminary study on integrating IRS with the lectures of fundamental physics in the first year university courses. In Proceedings of the International Conference on Multimedia and Information and Communication Technologies in Education2006 (m-ICET 2006), Seville, Spain.

Wang, Y., Lai, H. M. , Shi, S. M. (2002). Create interactive classrooms - The Applications of "An-An-An" interactive system in teaching.

Effect of Inquiry Web-Based Learning Competition for Gifted Students in Junior High School

Yen-Hung SHEN[a*], Wen-Gin YANG[b]
[a]Graduate Institute of Science Education, National Taiwan Normal University, Taiwan
[b]Science Education Center, National Taiwan Normal University, Taiwan
*syhdrums@tp.edu.tw, wgy@ntnu.edu.tw

Abstract: The study was designed to explore how to proceed an Subject-Inquiry-Based Learning and enhance their knowledge about weather through joining the Inquiry Web-Based Learning competition held by the Department of Education of Taipei City Government. The purpose was to find out how they completed the assigned tasks and what they learned from the process of this Web-Based Learning competition. The subjects were 30 gifted students in one junior high school. The results show that most of the gifted students completed the tasks by using the database in Central Weather Bureau and Taipei Weather Inquiry-Based Learning Network (TWIN). After the Web-Based Learning competition, they learned how to reflect better from their peers' reflection and gained additional knowledge about and emotional support from their peers' comments. There are some effects such as their attempt to apply the knowledge into practice, change of their beliefs and the gain of additional perspectives.

Keywords: Inquiry, Web-Based Learning, Environmental Education

1. Introduction

Inquiry-Based Learning is a kind of way of learning which is concerned about how students study actively by themselves. It could be practiced widely in all kinds of subjects. Recently, more and more nations have started to care about the students' abilities of inquiry, problem solving and so on. However, the implementation process of Inquiry-Based Learning still meet a lot of restriction and difficulties (Edelson, 1999). In this modern Information Age (or New Media Age), if we could merged Inquiry-Based Learning into teaching through new Web-Based Learning models, we would offer a new beneficial environment for students to study easily.

Web-Based Learning Competition is constantly held by the Department of Education of Taipei City Government. Through the Web-Based Learning Model created by themselves, the government hopes to encourage students studying junior high school or elementary school to use the learning model as a kind of tool to express their ideas and communicate with one another, and increase the ability to evaluate through the reflection and peers' comment. The topic is about the phenomenon of the weather such as climate change, global warming or extreme weather. First, all the participants are asked to form a team including 1 to 3 members and at least one coach who could be teacher or parent. Then, they are able to begin to collect the information such as the temperature or rainfall as much as possible to come to their own conclusion. They are suggested to use the database in Central Weather Bureau, Taipei Weather Inquiry-Based Learning Network (TWIN), or others.

Figure 1. The Front-page of Taipei Weather Inquiry-Based Learning Network (TWIN).

2. Relevant Literature

The technology of computer network is able to not only satisfy the learning motivation, provide the multiple content and create a kind of learning environment full of resources but a online social network. (Linn, 2003) Online social networks (OSNs) are increasingly attracting the attention of academic and educational researchers intrigued by its affordances and richness. OSNs provide powerful means of sharing, organizing, as well as finding contents and contacts. A large-scale measurement presented the study and analysis of the structure of OSNs. According to Chinn *et al.* (2011), the general conceptualization about the nature of epistemic cognition development not only makes its theoretical and empirical boundaries wider, but it also represents one of the main hypotheses of several stage-development models.

Moreover, the website for the Inquiry Web-Based Learning competition is capable of recording the discussion between participants and coaches so that all the participants could watch other's ideas and reflections. Writing reflection enables school teachers to examine the relevance of the training content and improve their teaching practice to meet the constant change of students‘ learning needs (Killeavy & Moloney, 2010). We could regard it as another advantage for all the participants during the process of this competition.

3. Method

3.1 Research Design

The study was designed to explore how to proceed an Subject-Inquiry-Based Learning and enhance their knowledge about weather through joining the Inquiry Web-Based Learning competition held by the Department of Education of Taipei City Government. All of the students were asked to complete all the tasks in different stages in the Inquiry Web-Based Learning Competition including Learning Sheet, searching the useful data and writing their reflections. We collected all the information recorded in the Taipei Weather Inquiry-Based Learning Network and analyze it, and then interview some of them if we need and the teacher.

3.2 Subjects

30 8th grade students were involved in this study led by a science teacher who had some experiences in the Inquiry Web-Based Learning competition. Those students are gifted students studying in the advanced science class. All of them were asked to join the Inquiry Web-Based Learning Competition and finish all the tasks.

3.3 Materials

The process of the Inquiry Web-Based Learning competition is divided into four stages. Each student has to complete the task on time so that they can enter to the next stage. The goals of the task in different stages are as follows: (1) to decide their own research questions; (2) to decide the data which will be chosen, the range, and how they intend to collect the data; (3) to analyze the data and draw the chart or diagram; (4) to finish their own report about the Inquiry-Based Learning. Not only completing the task but peers' comment between different coaches could each team pass into the next stage so that all the participants are capable of collecting more ideas to modify their research.

3.4 Data Collection Procedure

This study adopted the constant comparison approach. All reflections and comments posted onto Taipei Weather Inquiry-Based Learning Network were copied onto a word document. The reflection expressed by the first participant were read carefully in paragraphs, and the way of writing reflection was color coded for easy recognition later. After analyzing all reflections of the first participant, the researcher compiled the codes to further consolidate common themes. The coding process continued in a similar way with the rest of the participants. If a new theme was emerged, it was added to the theme list. After reading through all reflections, common themes were summarized. The themes were further compared and combined. A list of two to four themes was finally generated. The same approach was applied to analyze the comments.

4. Results

4.1 Reflection

Altogether 120 individual reflection were found in Taipei Weather Inquiry-Based Learning Network and 132 codes were labeled. Three main themes emerged from the reflection, which were:(1) elaborating on the content; (2) applying the new information into practice; and (3) changing beliefs.

Reflecting on the topic content was a basic requirement for the individual reflection. The participants elaborated on the content in a little different ways. The most of them often used was that they simply repeated what they had learned from the data downloaded or added certain personal understanding without in-depth explanations (N = 40, 30.3%). Another way was that they elaborated on the content further by adding some new information (N = 34, 25.8%). It was evident that they went to the Internet to search for additional information for better understanding of the contents about the topic that were new to them. In their reflection, they shared the additional information. An additional way of writing reflection was they elaborated on the content by connecting it to previous lessons, reflection, or content learned from other classes (N=10, 7.6%). It seemed that they attempted to integrate the newly learned content into their existing knowledge structure.

Two ways of applying the new learned information into practice was found. One way was that they applied those new information into practice (N=22, 16.7%). The other way was they attempted to explain certain existing weather phenomena by using the new learned information (N=15, 11.4%). Some of them even mentioned that they might share those information with their family members. Another theme emerged from the reflection was the new information inspired them to rethink about their original beliefs and as a result their assumptions started to change (N=11, 8.3%). For example, before joining the Inquiry Web-Based Learning competition, some of them thought that the network was only a tool for sharing information with classmates. They seldom thought that it could be used as a type of learning tool. After finishing the Web-based competition, their opinions started to change and recognized that using technology to study was not that difficult. Besides, some of them also mentioned that they have learned a lot of knowledge about the weather they did not know before.

4.2 Comments

Altogether 40 individual comments were found in Taipei Weather Inquiry-Based Learning

Network and 88 codes were identified. Also, three major themes became obvious in their comments: (1) commenting on the content; (2) expressing encouragement; (3) commenting on the way of writing reflection.

Peers used to further elaborate on the reflection content in their comments (N=32, 36.4%). To some extent, different ways of making comments were noticed. One way was that they picked up certain keywords or points from the reflection and elaborated further by adding new information or explanations. Another way was they tried to offer some useful explanations to the problems mentioned in the reflection or to explain why the problems existed. Part of them might add personal experiences or additional arguments to support the opinions expressed in the reflection (N=25, 28.4%), or disagreed with certain ideas in the reflection by providing with different examples or perspectives (N=8, 9.1%). In some comments, peers also stated what they learned from the reflection (N=6, 6.8%). Some of them indicated that they benefited from the reflection in an indirect way as certain ideas in the reflection stimulated them to search for more information or study further.

Peers also expressed encouragement in their comments (N=10, 11.4%). Generally, peers gave encouragement in two varied situations. One was that they gave positive comments or encouragement when the reflection writer presented good ideas, comprehensive summaries, or constructive suggestions. The other way was that peers gave encouragement when they realized the reflection writer had certain problems or difficulties. In this case, they encouraged them to face the difficulty positively. Sometimes they did not comment on the reflection content, but on the way of writing reflection (N=7, 8.0%). What impressed them most was the responsible way of writing reflection, or the positive attitude towards reflection writing or learning in general.

4.3 Interview of the teacher

According to the reflection of the teacher after the Inquiry Web-Based Learning competition, there are some points mentioned. Among the positive responses, the following reasons appeared: (1) writing reflections through Taipei Weather Inquiry-Based Learning Network could help students study more efficiently; (2) with the aid of Taipei Weather Inquiry-Based Learning Network, students get more useful ideas and encouragement during the process of learning; (3) students are able to learn how to tell and pick up what is the useful information and improve their ability of inquiry.

However, there are some restricting part mentioned by the teacher as follows: (1) some students may not have enough time to finish their task completely; (2) before joining the Inquiry Web-Based Learning competition, all students need to be well-taught to have the ability to search the data and explain it correctly.

5. Conclusion and Suggestion

The study has some implications for using Taipei Weather Inquiry-Based Learning Network. some participants were more experienced or responsible and their reflection was in more depth. The others commonly indicated that they learned how to reflect from these learners. Besides, the result also suggest that involving some higher ability or more experienced learners would show a positive sample to others and hence has the possibility to promote collaborative reflection to a higher level. In this study, all the participants are gifted students in the same class. That is why they could explain the same content or phenomenon from different perspectives, or give varied interpretations so directly. This result also implies that being familiar with one another would increase the likelihood of gaining more benefits from peers in the collaborative reflection process.

Furthermore, in addition to learning content directly from the lesson or the instructor, the participants in this study also shared their understanding with peers. Through the sharing, peers learned additional information, different perspectives, or the way of writing from the reflection. Collaborative studying can also lead to higher level thinking and transformational changes. It involves both cognitive and affective processes.

References

Edelson , D.C., Gordin , D.N & Pea R.D. (1999). Addressing the Challenges of Inquiry-Based Learning Technology and Curriculum Design. *The Journal of The Learning Science*, *8*(3&4), 391-450.

Linn, M. C. (2003). Technology and science education: starting points, research programs, and trends. *International Journal of Science Education*, *25*(6), 727-758.

Chinn, C., Buckland, L., & Samarapungavan, A. (2011). Expanding the dimensions of Epistemic Cognition: Arguments from philosophy and psychology, *Educational Psychologist, 46*, 141-167.

Killeavy, M., & Moloney, A. (2010). Reflection in a social space: Can blogging support reflective practice for beginning teachers? *Teaching and Teacher Education, 26*(4), 1070-1076.

Evaluation the situation somatosensory game digital learning for global warming misconception

Hsin-Chih Lai [a], Chi-Chen Lee[b*]

[a] Department of Engineering & Management of Advanced Technology, CJCU, Taiwan
[b] Department of Information Management, CJCU, Taiwan

* kbarfnh @gmail.com

Abstract: This study designed an innovative learning material based on game-based learning theory and situational somatosensory game digital learning modules (Situational Somatosensory Game Digital Learning modules). It is a funny, interactive and educational tool in the sense of a combination of role-playing game with somatosensory. The main structure is built up by Role-Playing Game masters and Kinect Flexible Action and Articulated Skeleton Toolkit programs. The content of learning modules is developed by the relevant literature and cognitive conflict strategy, the scientific misconception of sea-level rise is designed to be solved during the gamed based self-learning processes. Through a pre-test/post-test questionnaire the descriptive statistics, ANOVA analysis of quantitative data and open questions are discussed. The results shows: 1. the learning (experimental group) effective use of scenarios and games, in addressing the global warming misconception have good study on the effectiveness of, and superior to the control group ($F = 6.31$, $p < .05$); 2.the learning module has over ninety percent of high satisfaction, and can effectively lead to motivation, and overall satisfaction than the control group. Based on the findings, the somatosensory game situational learning led to a strong and highly motivated satisfaction.
Keywords: Global warming, Somatosensory Technology, Role-Playing Game, Misconception

1. Introduction

Scientific knowledge of Global warming mitigation and adaptation is getting more and more important in present world, but through the domestic researches and survey, students and the general public still hold considerable misconception on global warming issues, taking action to provide correct knowledge and learning tools should be the first priority we need to go. To convert misconceptions in non-formal education is not easy, the first step is to inspire the motivation by interesting learning processes but not the unsuitable traditional learning. Therefore, this study tries to combine new technology based on the theoretical, the design and development of a new way of learning science to solve global warming misconception, and further validate its effectiveness and satisfaction.

According to the motivation of this study, research questions are listed below.
Q1: Try in a good theoretical basis and practical framework for the design and development of an innovative way of learning.
Q2: Explore the effectiveness of an innovative way of learning
Q3: Evaluation of the satisfaction of an innovative way of learning.

2. Literature review

Digital game is getting more and more popular recently through combining new technology, it could interest the users to learn by funny, highly interactive and instantaneity, it also can help the users to bring their mind up in how to knowing, using, solving and creating. Traditional pedagogy is not easy to learn because it's highly complexity in class, but digital game can help to solving the problem, learner can get knowledge or technology when they have fun (Trotter, 2004). Digital game is an E-media, it's inoculation for teenagers with casual, relaxed and diversity, thus the teenagers are more willing to learn new knowledge. As a Game-based learning materials, the first priority to develop is how to get achieve a complete efficacy, such as problem solving, role-playing and Situation Simulation, that is to say how creativity and advisability blend into the content of courses and how to make a breakthrough in combining new technology. It's more interested the learner in RPG game genre than the others (Paraskeva, Mysirlaki, and Papagianni, 2010). Generally, the digital games need basic skills, it's not easy to the people that who has creative but unskilled to join the games industry (Blow, 2004).

Kinect

From Nintendo has been launched to the Wii, and the Kinect had been developed by Microsoft, they are the type of NUI. Kinect users can control the game directly by moving the body but needn't to hold the remote controller or wear device. It is easy to apply in many fields including the real life, education learning, entertainment, health care and other fields. Somatosensory interactive games can make learner involve into game situations easier and faster than traditional game with keyboard or mouse. The advantage of kinect game are with high interactivity, simple operation and less susceptible in age restrictions; however the weaknesses is that players need to standing continuously in the game. By study abroad, somatosensory technology have very good results in various fields, including the motivation, learning achievement and learning and memory. (Chao, Huang, Fang, & Chen,2012; Chun-Yen Chang, Yu-Ta Chien, Cheng-Yu Chiang, Ming-Chao Lin & Hsin-Chih Lai ,2013)。

Role-playing game

Role-playing is a game to be satisfied with role recognition , major interface is mainly with text have provided a environment with challenging and similar real-life situation . The most features of the protagonist have different characteristics, attributes or skills that players could accumulate experience points, level up and enhance the value of health points. Raybourn (2006) proposed role-playing game have the most helpful to solving complex problems and conflict mediation in learning. When learners actually play a role have experienced their point of view, the problems and solutions will become clear (Shifroni & Ginat, 1997), At the same time, role-playing is the most favorite game modes (Buchanan, 2004) applied to the assisted learning materials also have good learning effect.

Situated Learning

Situated learning placed the learner in real or simulated situations to learn through interaction between the learner and the situations, so that learner will apply what they have learned in actual life. When the scientific concept is abstract, complex and different with life experiences, it's process will be difficult to change the prior scientific concepts. Kathleen and Deborah (2004) considered the traditional digital learning is to provide knowledge and information, and less interaction with learner. Also Brown, Collins, & Duguid (1989) emphasize the learners construct knowledge by interactive with learning environment. Therefore, if the situational learning theory into the teaching content, it can development of situational learning, explores problems and solves problems. It will be more helpful for improves learning effectiveness. (Guralnick, 2008).

Global warming misconception

Global warming is part of the climate change. It becomes more wider and wider with its complexion than traditional science, the influence including scientific, political, economic and living. Moreover, it is cross-cutting, cross-generational and cross-border to make misconception of learning. Currently, people get many special messages from print media or electronic media in reporting global warming news, the news are not to be digested to accept. (Rye, Rubba, and Wiesenmayer, 1997). A lot of reason of the Scientific misconceptions are caused, Driver (1981) believe the misconception from the behavioral process in children and human interaction, as well as through their own observations and experiences obtained. And Gilbert, and Watts (1983) have proposed three perspectives: 1.children explore the world based on self-view; 2.child interested in doing something special for the scientific explanation does not consider whether the correct natural science; 3.They don't distinguish between different the scientific language and language of everyday life. Head (1986) also made five reasons: 1.daily experience; 2.analogy confusion; 3.fuzzy unknown confusion on the meaning of words; 4.learning with peer; 5.from individual nature of ideas.

3. Situational Somatosensory Game Digital Learning modules

Research framework

Situational somatosensory game digital learning modules which is applied by Game learning and situational learning theory and based on Piaget's theory of cognitive to solve the science misconception. Also it is Combining the role-playing game with kinect to effectively integrate in the learning method, learning tools and learning content (Figure. 1).

- Learning method: game-based, it integrated cognitive conflict strategy into digital material to make using characters as self-exploration and self-discovery. The learning method build cognition and break the original misconception for learners.
- Learning tool: somatosensory technology with role-playing games is used to develop better situational learning environment. Every movement in the game is meaningful to enhance the learner's interest in learning and impressions to reach entertaining.
- Learning content: features of the story to perform tasks and talking, with somatosensory control, learner can play easier and take into the situation for fun and challenge.

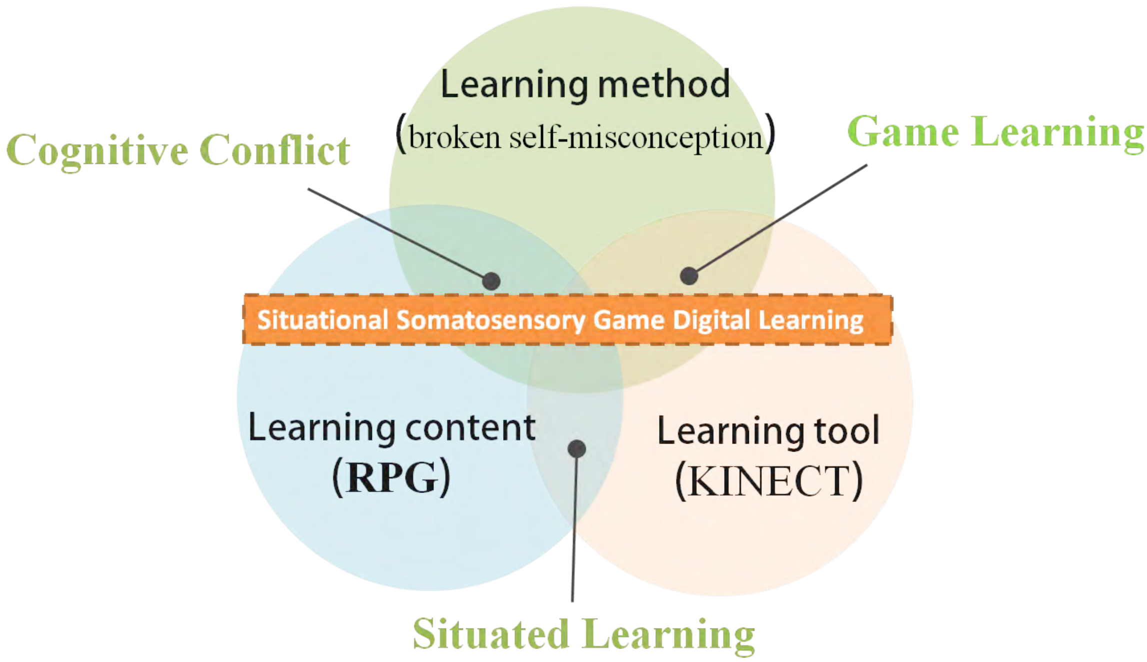


Fig. 1. research framework

4. Global Warming Misconception of material design

Learning strategies and processes

In this study the use of cognitive conflict is in three steps: (1) understand the learners' prior knowledge; (2) provide students with contradictory information; (3) assessment learners' cognitive change between pre-test and post-test. Figure 2 presents the strategies of learning design. First of all, make sure the learners' knowledge about the Antarctic and the Arctic ice melting to cause the sea-level rise of cognition before playing the game; second, play the game, through the game's story and dialogue, presenting a conflict with the results, learners can self-discovery and think for themselves; third, post-test to assess whether there are changes in cognition after the game, assessment points are misconception resolved and cognition established. Finally player will see a movie to consolidate this new concept, "seeing is believing" can help strengthen this perception.

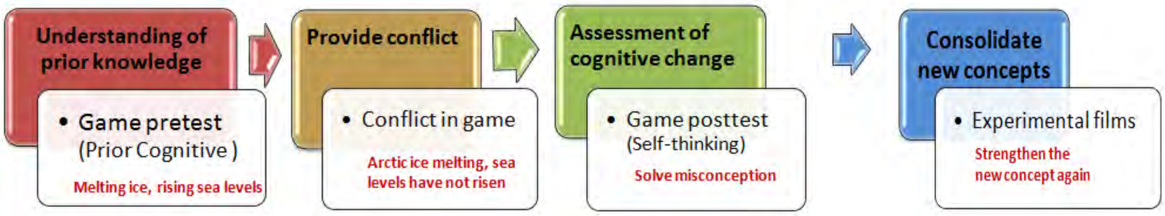


Figure 2. learning strategies

Figure 3. Research Process

Learning material design

The main objective is to establish a scientific knowledge and to solve misconceptions, this study focus on the case of the melting of the Arctic sea ice will cause sea levels rise. Design a self-learning process using the Antarctic, the Arctic ice melting caused by differences in sea-level change situation. There are five stages of the game content (Figure 4), the first step is to accept the task and play the role into the game situation, simultaneously, game will show learners the body operation to control the game; the second stage simulates the real and funny experience of marine environment of Antarctic and Arctic by real photos and virtual screen for learners; the third stage presents differences in structure of the Antarctic and Arctic by plot and dialogue, also guide the learners think about the influence of the land ice and ice land; by situations of life, the fourth stage by the game of saving polar bears and penguins, solving the misconception of "Arctic ice melt will cause sea levels to rise".

Figure 4. Game screens

5. Discussion

Quasi-experimental method is adopt to assessment of the performance, using of the somatosensory system is assigned as the experimental group and the control group is using a conventional keyboard device for learning, the total number of test facilities is 426 (165 in the experimental group and control group 261), through a pre-test/post-test questionnaire the descriptive statistics, ANOVA analysis of quantitative data and open questions are discussed.

According to quantitative data analysis shown that : 1 All of the testing scores shown significant improvement by using the game-based learning processes in solving the global warming misconception (experimental group mean improvement 1.15, control group mean improvement 0.92), also we found the outcome of somatosensory-combined design is better than traditional keyboard ($F = 6.31$, $p <.05$). These results indicate the innovative somatosensory game situational learning can effectively improve the global warming misconception as well as highly confidence in future adaption and promotion; 2. From the investigation, the testers' knowledge of global warming are coming from television media, teachers and the network, it is consistent with domestic research. But the cross analysis from pre/post test scores showed that the source of information got by learners with proper scientific understanding in their prior experiment is come from schools and teachers; 3. By the assessment results, the satisfaction for physical operation is up to 86.3% in experimental group learners, and for the game interesting, memory degrees, smoothness is higher than 90%. Moreover the willing to using the same learning module but different contents is totally agree (100%), that shows the design of this study, said somatosensory game situational learning, lead to a strong and highly motivated satisfaction.

6. Conclusion

This study has successfully developed a low cost Situational somatosensory game digital learning module, it is the first exploration in using somatosensory technology with role-playing games, to solve scientific misconception, than through formal teaching experimental results show that the concept for the global warming misconception resolved, and the establishment of proper scientific knowledge, a good upgrade results. Specific results are concluded below:

1. A successful use of low cost and easy to entry with RPG Maker and a free FAAST to develop the production of the game materials, cost of production has a good advantage in the subsequent promotion of environmental education.
2. Features of role-playing game including the dialogue, the plot and visual presentation to guide the effective thinking and learning in the teaching content. The results show the learners have high satisfaction in innovative learning method that can triggered strong motivation and interest, and to achieve learning objectives and effectiveness for happy learning.
3. Both of the experimental group and control group has good satisfaction higher than 90%, and experimental group were better than control group. Specifically, the willing to using the same learning module in different contents is totally agree (100%), that shows the Situational somatosensory game digital learning modules of this study, led to a strong motivation and highly satisfaction.
4. From this study of pre/post test showed that the source of information got by learners with proper scientific understanding in their prior experiment is come from schools and teachers, so it is recommended that correct conduction of scientific knowledge, and to accept the message of environment and situation are highly correlation between. But the message (learning content) is best through integration and design, in order to establish the correct perception.

References

Blow Jonathan (2004). "Game Development Harder Than You Think", (29). ACM QUEUE.

Brown and Duguid Collins(1989). "Situated Cognition and the Culture of Learning", Educational Researcher, 18, pp.32-42.

Buchanan, K. (2004). "How an educator thinks about computer games", Republic. Retrieved June 13, 2007 from http://www.uwsp.edu/education/kbuchana/

Chao, K. J., Huang, H. W., Fang, W. C., and Chen, N. S. (2013). "Embodied play to Learn: Exploring Kinect-Facilitated Memory Performance", British Journal of Educational Technology, 44(5), pp.151-155.

Chun-Yen Chang, Yu-Ta Chien, Cheng-Yu Chiang, Ming-Chao Lin and Hsin-Chih Lai. (2013). "Embodying gesture-based multimedia to improve learning", British Journal of Educational Technology, 44(1), pp. E5-E9.

Gee, J. P. (2007)." Good Video Games + Good Learning: Collected Essays on Video Games, Learning, and Literacy", New York: Peter Lang Publishing Inc.

Driver, R. (1981). "Pupils' alternative frameworks in science", European Journal of Science Education, 3(1), pp.93-101.

Gilbert K., and Watts K. M. J. (1983). "Concepts, misconception and alternative conceptions: changing perspectives in science education", Studies in Science Education, vol.10, pp.61-98.

Guralnick D.A. (2008). "Putting the education into educational simulations:Pedagogical structures, guidance and feedback", Paper presented at the 11th International Conference on Interactive Computer aided Learning (ICL2008),Villach, Austria, Sep 24~Sep 26 2008

Head, J. (1986). "Research into "Alternative Frameworks" : promise and problems", Research in Technological Education, 4(2), pp.203-211.

Kathleen, I., & Deborah, C. (2004). "Scenario-based e-learning design", Performance Improvement, 43(1), pp.16-22.

Limón M. (2001). "On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal", Learning and Instruction, vol.11, pp.357-380.

Mintzes J., Wandersee, J. H. and Novak, J. D. (2005) . " Teaching science for understanding - a human constructivist view", Elsevier, pp.349.

Paraskeva, F., Mysirlaki, S., and Papagianni, A. (2010). "Multiplayer online games as educational tools: facing new challenges in learning", Computers & Education, 54(2), pp.498-505.

Pfundt, H., and Duit, R. (1991). "Biblography : Students' alternative framework and science education", Kiel Univ., Germany.

Raybourn, E. M. (2006). "Applying simulation experience design methods to creating serious game-based adaptive training systems. Interacting with Computers", 19, pp.206-214.

Rye, J. A., Rubba, P. A., and Wiesenmayer , R. L. (1997). "An investigation of middle school students' alternative conceptions of global warming", International Journal of Science Education, 19(5), pp.527-551.

Shifroni, E., & Ginat, D. (1997). "Simulation Game for Teaching Communication Protocols", Proceedings of the 28th SIGCSE Technical Symposium on Computer Science Education (SIGCSE 1997). San Jose, California, pp.184-188.

The Environmental Education of Migration Birds Using a Near Time Web-based Design

Chow-Jeng WONG[a*], Chen-Jeih Pan[b], Yi-Jong TSAI[c]

[a]School of Physics, Universiti Sains Malaysia, Malaysia
[b]Institude of Space Science, National Central University, Taiwan
[c]Kenting National Park Taiwan
*wongcj.usm@gmail.com

Abstract: Raptor Migration is one of the most famous ecosystem mystery phenomena around the world. Both Malaysia and Taiwan are located on the flight path of the migration tracks of raptors in East Asia. For the past years, researchers associated with the Kenting National Park used a Doppler Weather Radar (Central Weather Bureau) to investigate the migration tracks of raptors, few episodes achieved very successful results. This study is to build up an information system which uses satellite images and Doppler Weather Radars to identify the raptor signal from the radar echoes. It is also going to build up a data flowchart, data base and Geographic Information System (Google Earth) for a near-time display. This system will provide the migration information for academic research, surface surveying and also for tourists and amateur bird watchers. Based on the near time track skill improvement, this study designed a web-based migration bird education for the public, we used web skill to demonstrate the near time migration track of raptors, the watching locations, as well as the e-book to introduce species of raptors.

Keywords: Raptor Migration, Web-Based Learning, Environmental Education

1. Introduction

Raptor Migration is one of the most famous ecosystem mystery phenomena around the world. Both Malaysia and Taiwan are located on the flight path of the migration tracks of raptors, especially where the Asian Monsoon directly affects the weather, Taiwan is a very important site for studying migrating raptors in East Asia. For the past years, to using a Doppler Weather Radar to investigate the migration tracks of raptors is a new skill to improving the track watch, Taiwan and Malaysia star the international project originated from the use of meteorological radars in monitoring the conditions of clouds and rains with radar.

The Doppler weather radar at Kenting, first set up in 2001, is especially useful in monitoring severe weather systems like typhoons that may inflict terrible damage to the southern part of Taiwan. It occasionally also detects echoes reflected from flocks of migratory birds. That gave the bird-loving KTNP Headquarters the idea of tracking birds with radars. The first attempts failed as the data received was too complicated and required human interpretation, which was time-consuming and error-prone. After the NCU scientists put in place an information system which uses an all-sky camera to identify radar echoes and conduct surface surveys, and also to build up data flows, a database and Geographic Information System (Google Earth) to form a real-time display system. Weather, sea and man-made noise is removed automatically and only the raptor information is displayed. It has met with great success in collaborations with KTNP in tracking birds by radar. Now, not only can noise in the radar be identified and eliminated, but the directions and routes as well as the times and locations of landing can be accurately monitored and forecast. This is an important advance in both bird conservation and bird-watching activities.

Usually, it is a matter of “Luck” for bird-watchers to enjoy following raptors in spring as the birds usually come in from the sea. Unlike autumn, in spring the raptors stay overnight in the Kenting area before leaving Taiwan to continue their flights. People then can enjoy watching the raptors land in the evening as well as rising in the morning. Even the most experienced officers of KTNP cannot guarantee where and when one will see the birds in spring. With the help of radar, it is easier to know where and when the raptors are going to land in spring. On the other hand, we know where they are leaving for in autumn.

Analysis of radar echo data shows that it is now known that these birds have already decided where they are going to land when they are about 100 km from their destination,. Researchers can tell from the radar data the birds' migration habits, flight altitude, speed, and other elements. They can even recognize when there is a coup among a flock of birds. However, more care must be taken in providing the information and promoting bird conservation; otherwise the project will help not just bird watchers but also hunters, and cause irreversible harm to the ecology.

Based on the near time track skill improvement, this study designed a web-based migration bird education for the public, we used web skill to demonstrate the near time migration track of raptors, the watching locations, as well as the e-book to introduce species of raptors.

2. Design of the Migration Raptor Tracking System for Education

The educational purpose of this information system is to attract people who is interested in nature environment and bird watching. It is very important to show the target of the system on the homepage (figure 1). Users can get the information what they want to know by selecting the functions that provide from the website, we use Java scripts to make the Raptor Migration Information System running faster and a few part functions to design a new type of inquiring function to help the users find their request.

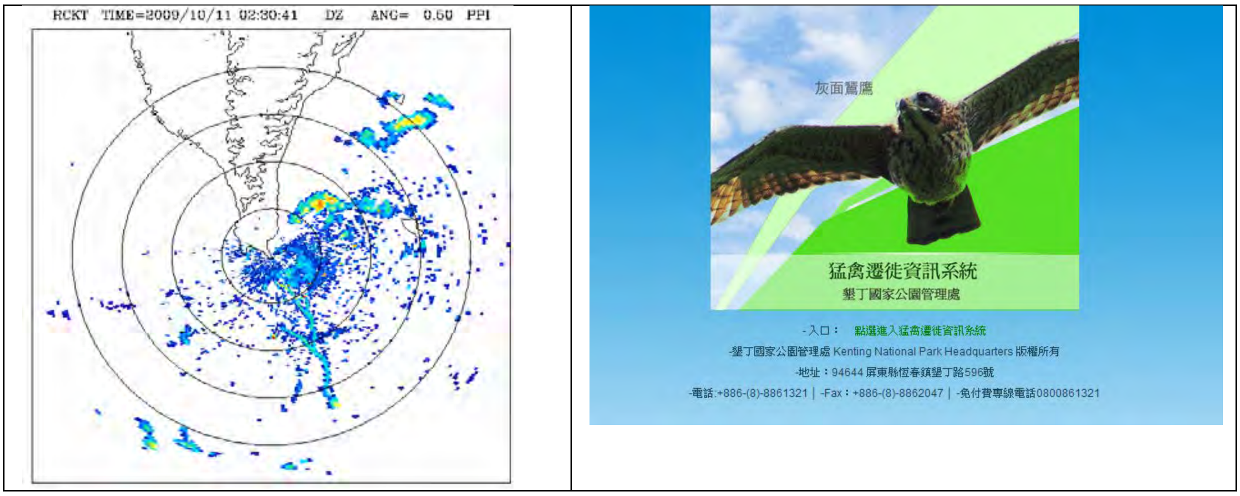


Figure 1. The welcome page of the Migration Raptor Tracking System

On the welcome page (Figure 2), the information includes project introduction and Features. We display animation about the successfully identify of raptor migrate case in recent years, help the users with knowing how the Raptor Migration Information System works easily. Renew display photo function, linked the upload photo in backend database and display photos at Homepage. An timely automatic statistics dynamic charts and bar charts regarding to the count of raptor migrate in Spring and Fall at bottom are also shown. The right sidebar is about the Sharing function, Visitor numbers, Member login, Raptor-expert message board, Raptor migrate of the season and related link.

We also design the Green Map (GM) of Raptor in Kenting for presenting the most species of raptors in the local area (Figure 3), GM is a map that promotes nature, sustainability, and ecology (please refer to http://www.greenmap.org/greenhouse/files/gms/GreenMapIcon_V3Chart.pdf), it is usually used to demonstrate the relationship between people and the environment. From 2005 to present the GMs are widely applied in outdoor education, even in Malaysia and Taiwan, a lots of GM are used to introducing schools, national parks, cities…etc., therefore, we used GMs as materials for raptors. Our design also embedded QR Codes into the program (Lai et al, 2013), it is a linked function provided the raptor information in Kenting area and video player by click the different raptor pictures.

Figure 2. The raptor information of the Migration Raptor Tracking System

Figure 3. the Green Map (GM) of Raptor in Kenting

An e-book (as known as electronic book) is a book publication in digital form, consisting of text, images, or both, readable on computers or other electronic devices. (http://en.wikipedia.org/wiki/E-book). It can be viewed in screen associated with computers, tablet PC, hands on smartphones. Now not only adults read an e-book, but also applied in children's education, it is over 50% of Americans by 2014 had a dedicated device, either an e-reader or a tablet, that is a great potential for the education in using e-book. By introducing the near time migration message could be involved in the e-book, a database of the successfully identify of raptor migrate cases are loading in the e-book system.

Added a database of the successfully identify of raptor migrate case(Figure 4), the identify system is one of the most features functions. By identify the signal on radar while the raptor migrate and display on website in time, there are many successfully identify cases are sorted in database, users can known the raptors's processes and habits when they were migrating by this cases.

Figure 4. The e-book of the Migration Raptor Tracking System

3. Discussion

From 2009 to 2010, camera field experiments were conducted to improve the radar echoes validation. We find the all-sky camera is a very useful instrument in conjunction with the weather radar to distinguish the raptors and clouds. Since 2010 we focus on the radar image recognition with more accuracy to create a complete information platform of the birds. The satellite image appending tests were carried out in spring and autumn to improve those parameters deduced by the weather radar.

The study has also worked an international cooperation in migration raptors education, there are 5 countries, hundreds of people were using the website to understand the raptors. It was also promoting in an international workshop hold in Penang, Malaysia (Figure 5), the participants were all experienced the system and made clear understanding with the migration raptors. This study will continue to serve in Spring and Autumn, and develop more convenient path as well as tools to help people in environmental education.

Figure 5. The international education path in using the Migration Raptor Tracking System

References

Green map activity guide (2005). Green Map Japan, Inc. Retrieved April, 1, 2010, from http://www.greenmap.org/greenhouse/en/node/1467

E-book, from http://en.wikipedia.org/ wiki/E-book

Lai, H. C., Chang, C. Y., Li, W. S., Fan, Y. Lin., & Wu, Y. T., Lin, M. C. (2012). The Implementation of Mobile Learning in Outdoor Education: Application of QR codes. *British Journal of Educational Technology*.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Development of Sign Language Training Machine using Depth Sensor

Yuichiro MORI[a*], Akie FUKUHARA[a] & Shogo HAYASHIDA[b]
[a]*Department of Information Science, Kochi University, Japan*
[b]*Graduate School of Information Science, Kochi University, Japan*
*ymori@is.kochi-u.ac.jp

Abstract: In this paper, we describe a development of Japanese sign language training machine using depth sensor that targets a beginner of Japanese sign language user and supports a learner by feeding back mistake information. This system consists of two parts, one is a input device of sign language and the other is a personal computer for study. That input device consists of flex sensors and accelerometer installed on gloves, and a depth sensor. The depth sensor (Xtion) is used to acquire the position of the hand and head, because this information is very important to analyze a sign language. Correct or incorrect determinations of input sign language motions are very important to achieve the sign language study accurately. Therefore, this thesis especially describes the correction judgment of sign language. For the internal processing, we defined fundamental motions and forms of a sign language, and modeled sign language motions by combination of fundamentals. These fundamentals are defined by some information of the shape of the hand, the inclination, the position and the direction of the movement. The number of sign language words necessary for communication with a hearing impaired is about 2,000 words. As the first step, we selected 231 words that an elementary conversation is possible and constructed the system. As a result, we are able to achieve correction judgments that have high accuracy. On the other hand, the problem in a present method is clarified, too.

Keywords: Japanese Sign language, Hearing-impaired person, Xtion, Data glove

1. Introduction

The United Nations adopted 'Convention on the Rights of Persons with Disabilities' in the 61st general meeting in December 2006. In this convention, there is the following item.

> *Accepting and facilitating the use of sign languages, Braille, augmentative and alternative communication, and all other accessible means, modes and formats of communication of their choice by persons with disabilities in official interactions;*

So, it is a policy of aiming at the environment that can easily take communication even if sign language is selected and used as a communication tools. This convention is ratified by 138 countries in October 2013, keeps increasing now, and the spread activity of sign language is advanced all over the world. In Japan, 'Basic Act for Persons with Disabilities' was revised on in 2011, and sign language was taken formally as a language (Cabinet Office Japan, 2013). To make the environment where it is easy to live for a hearing-impaired or speech-disabled person, a sign language broadcasting and a sign language window, etc. are set up now. However, sign language cannot be used in other places, consideration to a handicapped person is not complete. To make that environment, it is necessary to increase a place where sign language can be used more than now. And to do that, an able-bodied person should study sign language and well understand them. There are three typical communication way between hearing impaired and able-bodied person, it is writing communication, lip speaking and sign language. However sign language requires much training, a lot of people are using it recently because the information carrying capacity is very high.

There are roughly separately two kinds of study methods of studying sign language. One is a image/video teaching material such as a book or DVD, the other is a sign language school. The book and DVD have a weak point that learner's mistake cannot be pointed out. Therefore, there is uneasiness that the learner is learned wrong motion and does not pass in practice. On the other hand, that weak point does not exist in a sign language school. However, the cost is high, and some restrictions of time and place are very large. Therefore, we developed the sign language training machine that targets beginners of sign language, and there are no restrictions of time and place, furthermore it supports the learner by feeding back mistake information.

2. Framework of Sign Language Training Machine

2.1 Outline

As shown in Figure. 1, this system consists of two parts, one is a input device of sign language and the other is a personal computer for study. This input device consists of flex sensors and accelerometer installed on gloves, and a depth sensor. The depth sensor (Xtion) is used to acquire the position of the hand and head. In our laboratory, we used a depth sensor from Kitagawa et al. (2013), and used both hands version of data glove from Matsushita (2014).

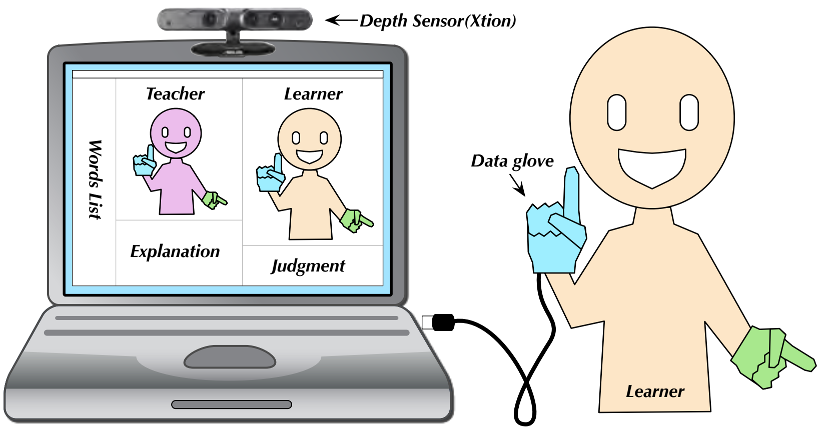


Figure 1. Configuration image of Sign Language Training Machine

Information of correct sign languages are registered in the database of this training machine as a teacher data. In the process of study, this system compares sign language information acquired from the learner with teacher's data, and judges the correction. When the learner makes a mistake, the reason is pointed out from the system, and the learner advances study again considering the mistake pointed out.

2.2 Sign language input device

That input device consists of flex sensors and accelerometer installed on gloves, and the depth sensor that is used to acquire the position of the hand and head. In present performance of a depth sensor, Xtion cannot acquire the shape and motion information of hands and fingers. Then, we use a data glove together, and acquire detailed shape, movement, and inclination of the hand.

2.2.1 Data glove

The role of the data glove in this system is acquisition of the information of hand's shape, movement, and inclination. As shown in Figure. 2, the flex sensor is installed in each finger, and we can observe the curved condition as a change in resistance. So, we convert this change into the change of voltage by principle of partial pressure ratio. We use PIC (Peripheral Interface Controller) to treat information from each sensor. Concretely, PIC digitalizes the voltage with A/D converter, and its data is sent to PC with USB communication. Similarly, the acceleration sensor also outputs the acceleration as a change in the voltage, it is converted with A/D converter in PIC, and it transmits to PC. There are a lot of signals that should convert A/D, one group is 5 signals from flex sensor of each finger, the other group is 3 signals (xyz-axes) from acceleration sensor of the hand. In total of both hands, they become 16 signals. Because the number of A/D channels of PIC is not enough, analog switches are connected outside, and data is acquired by switching the sensors. These switches are called a multiplexer (MUX), and we show MUXes in Figure. 2.

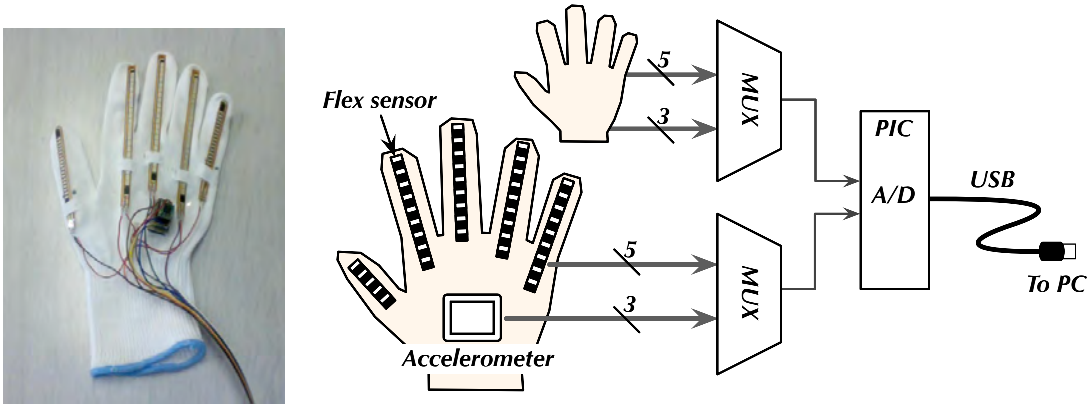


Figure 2. Data glove (left) and the peripheral circuit (right)

2.2.2 Depth sensor (Xtion)

This system uses depth sensor (Xtion) as a sensor to acquire the location information of learner's hand. By using skeleton tracking function (Figure 3) of Xtion, we acquire the three-dimension location information. As a depth sensor, we selected Xtion, which has the function equal with Kinect. Because Xtion is smaller than Kinect, and AC adaptor is not needed (USB Bus Power). Moreover, Xtion uses OpenNI, we can widely select a environment.

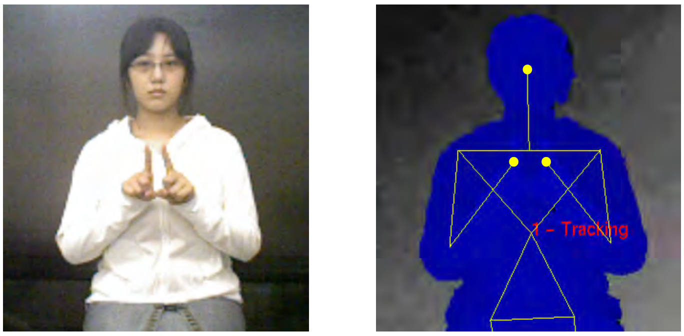


Figure 3. Real RGB image (left) and its skeleton tracking image (right)

3. Targeted words of sign language learning

The sign language word of about 2,000 words are necessary to take communication freely with a hearing impaired person, as shown in Table 1. Therefore, a practicable sign language training machine should be able to study more than 2,000 words finally. As the first step of this study, we selected 231 words that elementary conversation is possible and constructed the system. This 231 words are necessary for the level 5~7 of Sign Language Proficiency Test in Japan (Sign Language Proficiency Test Association, 2014).

Table 1. Necessary word number classified by Sign Language Proficiency Test

Level	*No. of words*	*Practicality*
1 ~ 2	about 2,000 words	A free communication is possible.
3 ~ 4	about 1,000 words	A daily conversation is possible.
5 ~ 7	about 200 words	An easy conversation is possible.

4. Correct or incorrect determination of sign language motion

In this system, sign language recognition is not a purpose. In the process of sign language study, the user perform sign language word requested by the system, and the system judges the motion which is correct or incorrect, and at the same time points out a mistake point. Seemingly, the system seems to recognizing a sign language. However, this system's processing is very simple compared with recognition. Extracting process of sign language word from time-series data can be simplified, because the user begins performance of sign language after the system directs the beginning point. Moreover, because system side knows what word motion of the input, there is no necessity for recognizing what it is, and it only has to judge the correction compared with a correct answer. If a recognition processing is requested, we need very complex processing (Toyokura et al., 2006 and Matsuo et al., 2013). However, processing can be simplified like these.

In the determination process, we should think about acceptable errors compared with the right motions. Recognition is not a purpose in this system as previously described. If this system aims at recognition, there is a necessity for considering about a tolerance of individual variation (habit) and error margin. It is a correction judgment ability that this system needs when training. Therefore, for the training of correct motion, we dared to reduce the tolerance of error margin and individual variation. So, these tolerance settings are very tight for recognition.

4.1 Determination process

In a judgment of a sign language motion, we resolve the sign language motion to fundamental motions, and are using the modeling method, which is modeling a sign language motion by combining the fundamentals. Moreover, it is possible to correspond to an increase of words in the future, because this method can express a new word by combining the fundamentals. In Table 2, we show various basic patterns of the hand used to define the fundamental motion.

Table 2. Patterns for fundamental motion definition

Basic pattern	*No. of patterns*	*Sensor*
Shape	24	Flex
Position	6	Depth
Tilt	6	Acceleration
Relocation	53	Depth & Acceleration
Shake	2	Acceleration

Even if which enumerated information in this table is missed, an accurate judgment is difficult. It became possible to acquire those information easily by not using a picture processing and using a depth sensor. Especially, the acquisition of the position of hands and head shown in Figure 4 are very difficult by using only existing sensors.

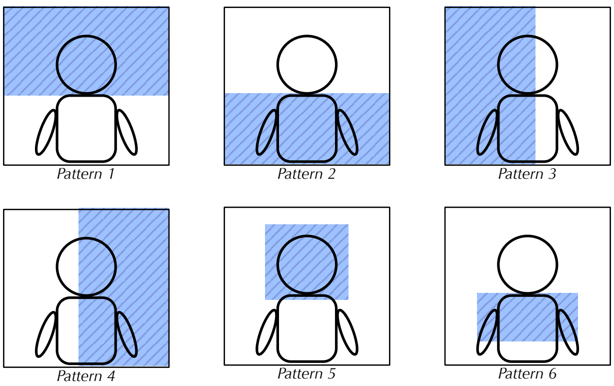


Figure 4. Position relation between hands and head

This system has teacher's database expressed by combining fundamentals, and correction of learner's motion is judged by using it. In most determination process, only initial state and end state of a sign language motion are modeled by fundamentals, and correction judgment processing is done by using it. Because the system knows the correct answer that should be judged, the correction judgment is enough by this method.

5. Learning system

At the learning part of this system, we implemented a basic function, as shown in Figure 5. The teacher's demonstration video and explanation of the sign language can be browsed. These are not different from video learning materials. Afterwards, the learner performs the sign language motion, and the evaluation is obtained. This is the feature of this system. When making a mistake, the learner confirms which part of motion is wrong by playback own video. Of course, an incorrect part is pointed out by the system. The learner can remember correct sign language motion by repeating this process. The study result is preserved by the system, and the learning scheme is made based on it. The environment of this system is shown below.

- CPU : Intel® Core™ i7-2670QM 2.2GHz
- Memory : 8 GB
- OS : Windows 7 64bit
- IDE : Microsoft Visual Studio C# 2010
- Framework : OpenNI & NITE ver.1.5.2

Figure 5. Screen shot of training machine

6. Evaluation

The determination accuracy of motion should be high to achieve correct sign language study by using this system. In other words, there is not a meaning if feedbacks of this system are not based on correct information. So that, we first executed evaluation of the determination capability.

As examinee's condition, we assumed a user of this system. It was a sign language beginner and if lecture was received, it was assumed level to be able to do correct motion. In other words, it is not a person who can do sign language with strange habits.

6.1 *Determination capability*

We experimented on all of 231 words this system targeted. Examinees were five of sign language beginners. First, the examinee selected a word with manual operation, confirmed the sign language motion, and afterwards, judged the motion by the system. Those experiments of each words were done after enough practice, and all motions were correct in check with eyes. And so, the Table 3 shows the determination rate when the examinee does correct motion. The average determination rate of the system became 88.7%, and shows a high performance. It can be said that this result has enough capability for use to learning system. On the other hand, the feature of the motion that failed in the judgment became clear, too.

Table 3. Determination rate

Examinee	***A***	***B***	***C***	***D***	***E***	***Ave.***
Determination rate(%)	90.5	90.5	90.0	86.1	86.1	88.7
No. of faults	21	23	23	32	32	26.2

This system has high determination accuracy. However, it is necessary to solve the following problems to improve accuracy from the result of experiment.

- Malfunction of skeleton tracking because of hands overlapping
- Shortage of definition of basic motion pattern "Hand's vibration (i.e. shake)"

The first is a typical problem of skeleton tracking by depth sensor. When hands overlap, it does not see from depth sensor, and malfunction occurs. It is necessary to prepare processing when losing. In addition, the middleware side should not lose even if it overlaps, too.

The next is a problem of definition of the fundamental motion based on information from acceleration sensor. In this part, we define the fundamental motion of a hand's inclination and vibration that is small shake motion. In these motions, there are a lot of individual variations more than our assumption. In this problem, it causes the policy of exclusion of individual variation and reduction of acceptable error.

7. Conclusion

In judgment of sign language motion, the position of hand and head is very important to say nothing of the shape and motion of hand. In an existing method, the location information was acquired by recognizing hand and head position by image processing. However, the processing is not easy. At the present day, that information can be easily acquired by using a depth sensor according to skeleton model with human body. And, we can acquire the motion and shape information of hand and finger accurately by using data glove. Accurate fundamental motions of sign language can be defined by combining these technologies. As a result, high determination accuracy could be achieved.

On the other hand, the problem in determination processing became clear, too. At the stage of practice of sign language, the learner should exclude an individual feature (habit), and do basic correct motion. We are aiming at a more accurate determination based on this policy now. However, when this system is enhanced to recognition processing, the individual variation problem cannot be avoided. Therefore, we should carefully consider these problems as the next one.

References

Cabinet Office Japan (2013). Annual Report on Government Measures for Persons with Disabilities 2013, Cabinet Office, Government of Japan.

Kitagawa, M. & Mori, Y. (2013). Development of the sign language training machine using the Kinect, *Technical Reports on Information and Computer Science from Kochi.* Vol. 5, No. 6.

Matsuo, T., Yamada, Y., Shirai, Y., & Shimada, N. (2013). Feature Extraction and State Decomposition for Image-based Japanese Sign Language Words Recognition Using HMM, *The Transactions of Human Interface Society 15*(1-4), 85-94.

Matsushita, T. (2014). Development of sign language training machine - About correspondence to the both hands sign language -. Bachelor thesis 2013, Dept. of Information Science, Kochi University.

Sign Language Proficiency Test Association (2014). Retrieved Aug. 7, 2014, from http://www.shuwaken.org/

Toyokura,Y., Nankaku, Y., Goto, T., Tokuda, K., & Kitamura, T. (2006). Approach to Japanese Sign Language Word Recognition using Basic Motion HMM. *Proceedings of IEICE Engineering Science Society 2006*, A-4-5, 72.

Significance and Possibility of E-Learning for Choreographic Skills in Contemporary Dance

Bin UMINO[a*], Asako SOGA[b] & Motoko HIRAYAMA[c]

[a]*Faculty of Sociology, Toyo University, Japan*
[b]*Faculty of Science and Technology, Ryukoku University, Japan*
[c]*Faculty of Health and Sport Sciences, University of Tsukuba, Japan*
*umino@toyo.jp

Abstract: In this paper, we introduce a learning methodology for choreographic skills in contemporary dance using the "Body-part Motion Synthesis System (BMSS)," which actualizes "analytic-synthetic choreography." The software allows students to select body-part motion clips of basic dance movements and compose a short dance motif, which is previewed using 3D animation. Experiments on learning choreographic skills with the software were conducted in order to evaluate its usability. Data were collected from 26 students using a questionnaire. From the results of the experiments, we verified that the software's usability is acceptable and that our e-learning method for contemporary dance is effective for learning choreographic skills.

Keywords: E-learning, choreographic skills, motion data, contemporary dance

1. Introduction

Although computational methods for dance creation have advanced significantly over the past two decades, there are only a few dance teachers who utilize computers to teach dance skills in their classes. Using computers in dance education has always been contentious because dance has the distinction of being both somatic and artistic. Somatic skills essentially need to be learned through physical practice, and it is currently difficult for computer devices to support the experience of physical action unless the teacher has access to costly hardware. Artistic skills need to be learned through an individual's creativity, but dance people generally believe that it would be almost impossible for a computer to inspire creative feelings or intuitions in a purposeful manner.

However, some researchers in dance education have reported the impact of computer technology on dance learning (e.g. Smith-Autard, 2009; Leijena, Admiraala, Wildschut and Simons, 2008). In particular, 3D animation and motion capturing technology have expanded the capability of e-learning for dance. Karkou, Bakogianni and Kavakli (2008) developed a web-based learning environment using 3D animation for traditional dances in the U.K. and Greece. Moreover, both Matsumoto, Miura and Kaiga (2011) and Shibata, Tamamoto, Kaiga and Yokoyama (2012) developed e-learning tools that could display 3D animation of model performances made from motion capture (MoCap) data for traditional dances in Japan.

The authors have been working on dance education and creation using 3D motion data that were captured from performances by professional dancers. This paper introduces a specific teaching methodology for choreographic skills in contemporary dance utilizing e-learning software that we have been developing in order to actualize "analytic-synthetic choreography," which is described in Section 2. The methodology targets undergraduate and postgraduate students who are studying contemporary dance choreography. The software is called "Body-part Motion Synthesis System (BMSS)," and it is described in Section 3. Experiments on learning choreographic skills with the software were conducted in order to evaluate its usability. Data were collected from 26 students in Japan and the U.S. using a questionnaire, which is described in Section 4.

2. Pedagogy of Choreography

2.1 Top-down and Bottom-up Approaches

In general, a choreographic process is carried out in a top-down approach. A choreographic work is composed of many artistic elements other than dance movements, such as narrative, music, sound effects, lighting, scenography and visual effects, costumes and stage make-up. Professional choreographers attempt to integrate all of these elements into their work coherently according to their own choreographic planning. The planning usually precedes the creative process of dance movements and theatrical audio-visual elements. Choreographers sometimes choose a specific narrative or a given musical score as a point of departure for their creation, but after that the process is likely to be a top-down approach.

Choreographic skills are normally taught to students at universities as a top-down approach, just as a professional choreographer normally works. They are required to formulate their own choreographic planning and implement it in their choreographic works. Most teachers of choreography place heavy emphasis on a holistic concept and the consistent structure of choreographic works at higher educational levels. Beyond that, the concepts that the teachers suggest tend to be expressionistic ones using narratives, emotions, or feelings to set up the choreography.

However, the authors believe that bottom-up and non-expressionistic approaches would also be vital and effective ways to master choreographic skills in contemporary dance. The students need to have the experience of composing dance movements without narrative, emotions, music, and all other audio-visual elements except for the dancers' bodies. This is because the pursuit of novel movements with the intention of independence from concrete references in the world characterizes the contemporary dance scene of the day. Contemporary dance can be described as an artistic dance without any common or standard choreographic vocabulary. A number of legendary twentieth-century choreographers, such as Rudolf Laban (1879-1958), Merce Cunningham (1919-2009), and the early William Forsythe (1949-) sought to produce their original dance movements without narrative, emotions, and music. Such an abstract way of dance composition should be taught as part of learning choreographic skills at universities lest the students' works be produced from only derivative material.

2.2 Analytic-synthetic Choreography

It is more difficult, but an exciting challenge, for choreographers to adopt the bottom-up and non-expressionistic approach than the top-down and expressionistic one. Computer technology sometimes helps them to pursue the abstract way. Cunningham famously pioneered the use of the choreographic software Life Forms in the late 1980s, while Forsythe used moving picture processing and hypertext technology to explain his unique algorithmic method of creating dance movements in the late 1990s (Forsythe, 1999).

The authors found that 3D animation made from MoCap data could be a powerful e-learning tool for learning choreographic skills in contemporary dance in the bottom-up approach. The basic idea of the method is to segmentalize dance movements performed by professional dancers into short elemental motions and then synthesize these segments as building blocks to create new movements. Both the elemental motions and the synthesized movements can be simulated easily and instantly as 3D animation. We call this method "analytic-synthetic choreography."

The segmentalization of dance movements is transacted in a double way. First, the dance movements digitized by a motion capture system are separated into basic whole-body movements along the time axis, such as stamping forward, sliding aside, bending the knees, and toppling in an off-balanced way. Second, the whole-body movements are articulated to extract basic body-part motions, such as contraction of the breast region, rotating the head, shaking the left leg, and crossing the arms in front of the body. Single elemental motion of the whole body or a body-part lasts a few seconds.

Meanwhile, the synthesis of dance movements is transacted in a triple way. First, you can select several whole-body movements and combine them in a row on the time axis to create a short dance sequence. For example, if you select three movements which are denoted by A, B, and C, then you can create ABC, BAC, ACBA, CBCAAB, and so on. Second, you can blend one whole-body movement into another whole-body movement. For example, if you select the sliding aside movement, then you can add the bending the knees movement to it at any timing you decide. The movements are overlapped on the time axis unlike in the first way. Third, you can replace part of a whole-body movement by a body-part motion. For example, if you select the stamping forward movement, then you can replace the head by rotating movement and the arms by crossing in front of the body. The movements are overlapped in this way as well.

The authors developed e-learning tools that actualized the first synthetic way for learning classical ballet (Soga, Umino, Yasuda and Yokoi, 2007; Umino, Longstaff and Soga, 2009) and hip-hop dance (Soga, Tsuda and Umino, 2014). In addition to these tools, the authors have developed an e-learning tool for contemporary dance that actualizes the second and third synthetic ways of creation, as described in Section 3.

The learning methodology based on analytic-synthetic choreography consists of three phases: (i) students create short dance motifs using the software and simulate them as 3D animation on the display of a computer; (ii) students try to perform the motifs by mimicking the 3D animation by operating their own bodies; and (iii) students create and perform their own short dance sequences by combining, arranging, and sophisticating the motifs in a studio. The second and third phases are essential for students to discover and develop their own creative process by themselves. Experiments were conducted according to the methodology, as described in Section 4.

3. Body-part Motion Synthesis System

The software BMSS has been developed to actualize analytic-synthetic choreography (Kohno, Soga and Shiba, 2010). In this research, 40 basic motions were selected meticulously as elemental motions so that dance students could use the system easily within a short span of time. Each motion's potential for synthesis was analyzed, and the 40 motions were categorized into 3 main groups: Base motions, Blend motions, and Body-part motions. Base motions consist of whole-body movements. Blend motions consist of whole-body movements that are able to be blended together with a Base motion. This group mainly consists of hip movements like jumping and twisting motions. Body-part motions consist of movements that involve only specific body parts or limbs. This group is further categorized into five sub-groups: Body, Neck, L-Leg (left leg only), Shoulders, and Arms.

The system creates movements by the synthesis of a Base motion, Blend motions, and Body-part motions. It has two modes: Blend mode corresponds to the second synthetic way mentioned in 2.2; Replace mode corresponds to the third one. In the Blend mode, the selected Blend motion of the whole body is blended with a Base motion in a way that the vector of the Blend motion for each joint in every frame is added to that of the Base motion. For example, when a user selects sliding aside as a Base motion and bending the knees as a Blend motion, a sliding motion with bending knees is created. In the Replace mode, the system replaces motions of specific body parts with different Body-part motions.

There are currently two versions of BMSS. The old version runs on a notebook PC with keyboard input, and the latest version runs on a tablet with touch input. The sets of 40 motions are slightly different, and the old version does not implement the Blend mode. Despite the differences, they are underpinned by the same learning methodology explained above. Table 1 shows the number of motions and examples of motion codes in each category.

Table 1: Number of motions and examples.

<table>
<tr><th colspan="2" rowspan="2">Category</th><th colspan="4">Number</th><th rowspan="2">Example</th></tr>
<tr><th colspan="2">Notebook PC</th><th colspan="2">Tablet</th></tr>
<tr><td colspan="2">Base</td><td colspan="2">15</td><td colspan="2">10</td><td>Stamp, Slide, OffBalance, Soutenu</td></tr>
<tr><td colspan="2">Blend</td><td colspan="2">0</td><td colspan="2">6</td><td>BendDown, JumpUp, JumpPivot</td></tr>
<tr><td rowspan="5">Body-part</td><td>Body</td><td rowspan="2">10</td><td rowspan="5">25</td><td>5</td><td rowspan="5">24</td><td>Contract, BustRoll, SideSwg, Wave</td></tr>
<tr><td>Neck</td><td>3</td><td>Down, Side, Round</td></tr>
<tr><td>L-Leg</td><td rowspan="2">5</td><td>5</td><td>Bend, InOut, LegShake, FootRound</td></tr>
<tr><td>Shoulders</td><td>3</td><td>UpDown, Shake, Roll</td></tr>
<tr><td>Arms</td><td>10</td><td>8</td><td>RaiseUp, PushFront, CrossFront</td></tr>
<tr><td colspan="2">Total</td><td colspan="2">40</td><td colspan="2">40</td><td></td></tr>
</table>

Figure 1 shows the GUI of the tablet version of BMSS. It employs only one window, which consists of a virtual environment and GUI components. The GUI components such as buttons are layered over the virtual world. The codes of all 40 motions are listed as buttons. First, you select a Base motion and the system displays it by 3D animation with a virtual dancer. Then, you select Blend motions or Body-part motions. The Base motion and the selected motions are synthesized, and the result will be displayed instantly. The system also has functions to support dance creation such as changing the viewpoint and saving the synthesized movements (Soga and Matsumoto, 2013).

Figure 1. GUI of BMSS for Tablet.

4. Experiment

4.1 Method of Experiment

Experiments were conducted using both the notebook PC version and the tablet version of BMSS. The experiments aimed to evaluate the usability of the software. From the point of view of learning choreographic skills, usability is the most important aspect of the e-learning tool. Usability is defined in ISO 9241 as “The effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments.” The users of the software are assumed to be students who study contemporary dance at universities. The goal of using BMSS is to learn skills in contemporary dance based on analytic-synthetic choreography. To evaluate the usability of BMSS, five questions were prepared preliminarily:

- Can dance students master the system to create dance motifs in a short span of time?
- Can dance students be satisfied with the system for learning dance movements?
- Can dance students be satisfied with the system for creating new dance movements?
- Can dance students discover new dance movements they did not know before?
- Can dance students discover new choreographic skills they did not know before?

The 26 examinees consisted of 18 students who majored in contemporary dance at University of Tsukuba in Japan and 8 students who belonged to the Department of Dance, University of California, Irvine, in the U.S. Here, 16 out of the 26 examinees were postgraduate students. They studied not only contemporary dance but also other genres of dance for 4 to 40 years. 8 out of the 18 Japanese students used the notebook PC version, and the remaining students used the tablet version.

The examinees gathered in a gymnastic studio and formed groups, each of which consisted of four or five students. After brief instructions on the software, the experiments were implemented in three phases: (i) the examinees tried to create short dance motifs on the display; (ii) they tried to perform the selected motifs by operating their own bodies; and (iii) they were requested to create their own dance pieces by selecting and connecting the short motifs. In the third phase, they were also requested not to mimic the motifs but to arrange and sophisticate them as their original choreographic works. After rehearsal, their short dance pieces were performed by themselves in front of a video camera one by one. At the end of the experiments, they were requested to complete a questionnaire. The experiment took about 90 minutes for each group, including the instructions.

4.2 Results and Discussion

Although the examinees were given only short and simple instructions lasting about 10 minutes, every examinee could master the system readily and smoothly with both the notebook PC version and the tablet version. They all completed the creation of their original short dance pieces within 90 minutes. The duration of the pieces varied from 13 to 53 seconds, with an average of 25 seconds. The number of motifs they connected varied from 5 to 10, with an average of 7.

In the questionnaire, the examinees were asked to answer about the effectiveness of the system for "understanding of dance movements" and "creation of contemporary dance" (Umino, Soga and Hirayama, 2014). The examinees were requested to choose from four options, and Table 2 shows the number of replies for each option.

Table 2: Students' ratings of BMSS.

	Understanding of movement		Creation of dance	
	Notebook PC	Tablet	Notebook PC	Tablet
effective enough	3	5	6	14
effective if modified	5	10	0	4
not so effective	0	0	0	0
don't know	0	3	2	0

In terms of understanding movements, 23 out of the 26 examinees chose either "effective enough" or "effective if modified." Three Japanese students with the tablet version chose "don't know" and commented similarly that watching 3D animation without practice was insufficient to understand the movements. However, no one chose "not so effective." The results suggest that the examinees were satisfied with the system as a useful tool for learning movements. In terms of contemporary dance creation, 20 examinees chose "effective enough." Two Japanese students with the notebook PC version chose "don't know," but all of the students with the tablet version chose either "effective enough" or "effective if modified." These results suggest that the examinees were satisfied with the system as a useful tool for creating new movements.

The free descriptions by the 26 examinees in the questionnaires were analyzed in detail. First, to the open question "Please describe in detail any choreography sequences that you made using BMSS that were unexpected and outside of your usual movement vocabulary," all 26 examinees described what they found as new movements of contemporary dance for themselves. Several typical responses (JP denotes Japanese, US denotes American) were elicited: "When I created movements which looked interesting on the display, I noticed which body-part I usually tend to ignore" (JP); "Jumping from being in a deep lunge/one knee on the floor was unexpected" (US); "Having the foot shake at the end of the leg touch was way outside of my movement vocabulary" (US).

Second, to the open question "What did you learn about creating contemporary dance choreography through this experiment?", all of the examinees described something they had learned. Furthermore, 22 out of the 26 examinees reported that they learned a novel choreographic method of contemporary dance. Several typical responses were gathered: "I could think of dance movements as combinations of simple movements" (JP); "I realized that combination and permutation of movements gave me original choreographies infinitely" (JP); "I learned that creating movements by layering different elements or body parts can be very effective" (US). The responses proved that the examinees learned the idea of analytic-synthetic choreography, although it had not been explained explicitly during the experiments. Although four examinees did not describe a new choreographic method expressly, they nevertheless pulled off creation using the analytic-synthetic way unconsciously.

After analyzing the free descriptions by the students in the questionnaires, it is reasonable to conclude that they could discover both new movements and new choreographic skills through the learning methodology based on analytic-synthetic choreography. Finally, we received positive answers to all five of the questions proposed in 4.1. A comparison between the notebook PC version and the tablet version was clarified in another paper (Umino, Soga and Hirayama, 2013).

5. Conclusion

Body-part Motion Synthesis System (BMSS) has been developed for learning choreographic skills in contemporary dance. The software is designed to actualize analytic-synthetic choreography. As a result of experiments, we verified that the system's usability is acceptable in terms of easy operability, students' satisfaction, discovery of new movements, and learning a new choreographic method.

At the beginning of this paper, the authors mentioned the difficulty of learning somatic and artistic skills with a computer. Regarding the somatic skills of dance, students will have learned new movements through the second and third phases of the proposed methodology. Regarding the artistic skills of dance, students will have learned analytic-synthetic choreography by themselves through this methodology. In conclusion, the methodology is effective for learning both somatic skills and artistic skills of contemporary dance.

Through the experiments, we received a lot of responses about how the examinees wanted to improve BMSS. In future work, we intend to improve it after reviewing these responses. For example, we plan to augment the user interface so that the system can promote learning more effectively.

Acknowledgements

We would like to thank Kanagawa Institute of Technology for permission to use a studio and a motion capture system. This work was partly supported by JSPS KAKENHI Grant Number 22300038.

References

Forsythe, W. (1999). Improvisation technologies: A tool for the analytical dance eye. (CD-ROM) Karlsruhe: Zentrum für Kunst und Medientechnologie.

Karkou, V., Bakogianni, S., & Kavakli, E. (2008). Traditional dance, pedagogy and technology: An overview of the WebDANCE project. *Research in Dance Education*, *9*(2), 163-186.

Kohno, Y., Soga,A., & Shiba, M. (2010). A system for motion synthesis of human body parts using a touch panel. *Proceedings of the 9th ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Application in Industry*, 145-146.

Leijen, Ä., Admiraal, W., Wildschut, L., & Simons, P. R. (2008). Students' perspectives on e-learning and the use of a virtual learning environment in dance education. *Research in Dance Education*, *9*(2), 147-162.

Matsumoto, N., Miura, T., & Kaiga, T. (2011). Research on application digital content teaching devise for learning Japanese folk dance in Akita: Trial production and application of 3DCG animation devise for dance education dating motion capture. *Journal of Japanese Society for Dance Research*, *34*, 1-10. (in Japanese)

Shibata, T., Tamamoto, H., Kaiga, T., & Yokoyama, H. (2012). Realtime Dance Learning Assistant System by 3D Human Motion Capturing. *Transactions of the Virtual Reality Society of Japan, 17*(4), 353-360. (in Japanese)

Smith-Autard, J. (2009). Creativity in dance education through use of interactive technology resources. In J. Butterworth & L. Wildschut (Eds), *Contemporary Choreography: A Critical Reader* (pp. 90-105). New York, NY: Routledge.

Soga, A. & Matsumoto, S. (2013). A motion synthesis system for dance using a tablet. *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, 178.

Soga, A., Tsuda, K., & Umino, B. (2014). A system for generating choreography on demand using dance motion. *Proceedings of NICOGRAPH International 2014*, 111-114.

Soga, A., Umino, B., Yasuda, T., & Yokoi, S. (2007). Automatic composition and simulation system for ballet sequences. *The Visual Computer*, *23*(5), 309-316.

Umino, B., Longstaff, J. S., & Soga, A. (2009). Feasibility study for ballet e-learning: Automatic composition system for ballet enchaînement with online 3D motion data archive. *Research in Dance Education*, *10*(1), 96-107.

Umino, B., Soga, A., & Hirayama, M. (2013). Choreographic support for contemporary dance using a Body-Part Motion Synthesis System: Comparison of learning effect between laptops and tablets. *IPSJ SIG Technical Report*, *2013-DCC-5*(19), 1-6. (in Japanese)

Umino, B., Soga, A., & Hirayama, M. (2014). Feasibility study for contemporary dance e-learning: An interactive creation support system using 3D motion data. *Proceedings of International Conference on Cyberworlds 2014* (in press).

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Design of an Environment for Motor-skill Development based on Real-time Feedback

Keita YAMADA[a*], Kenji MATSUURA[b]
[a]*Graduate school of advanced technology and science, The University of Tokushima, Japan*
[b]*Center for Administration of Information Technology, The University of Tokushima, Japan*
*ma2@tokushima-u.ac.jp

Abstract: Advancements of information technology and sensing technology enables to support motor-skill development in real-time. The authors propose a supporting model, which improves the form of motion in real-time. Our proposal shows the difference between the present form and the ideal form. In addition, it offers the expected form in future and the result of the performance. It is a reason why the authors consider the process of human motor-skill. The target skill is a sort of closed skill. The retrival in the form database is improved using the k-means algorithm. At last, the authors touch upon an experiment of our proposed model.

Keywords: skill development, motor skill, real-time, motion capture

1. Introduction

Advances in hardware and software technology make it possible to obtain digital data of human motion by video and sensors. In addition, there are several studies that support acquisition of motor skills using the digital data (Hamagami, et al., 2012) (Kosaka, et al., 2011). The field of these studies is called motor-skill development. In this research field, a motion of a human is called "motor-skill". A motor-skill is a physical-activity, which humans can master with a specific education or a training.

The supporting environment of motor-skill requires significant functionalities of "Monitoring", "Analysis" and "Feedback” (Schmidt, 1975). “Monitoring" function provides conversion of a human-motion to data. “Analysis" function is to calculate the difference between the monitored data as a learner and the model data as an instructor. “Feedback" function reports the straightforward information denoted by the difference to the learner using natural language and graphics. In addition, Zimmerman (2008) reported a model of the timing when learners acquire motor-skill in self-learning.

Moreover, motor-skill can be classified according to some criteria. We can classify the supporting-model using the timing of intervention. Intervention by the system is carried out either asynchronously or synchronously. In asynchronous supporting-model, we monitor and record full action of motion from beginning to the end using video and sensors at first. Secondly, we calculate the characteristic difference between the monitored motion and the other one for comparison. Finally, we advise the learner in an appropriate manner due to the result. Such an asynchronous supporting-model has an advantage of low-cost in terms of the system device and the development. Regarding synchronous supporting-model, we also monitor and record motor-action of a target skill using video and sensors at real-time. Secondly, we calculate the difference in parallel to monitoring while the learner performs continuously. At last, we advise the learner during the ongoing performance of the motor-skill. The synchronous supporting-model has a contradistinctive advantage, which the learner can obtain the advice at real-time. Therefore, the learner can correct the behavior using the advice on real-time.

In recent years, the improvement in the throughput of a computer enables high-speed information processing. Therefore, “synchronous supporting" becomes possible. In this research, the authors discuss the system design for "synchronous supporting”.

2. Requirements for Motor-skill Development

Fitts et al. (1967) say that the motor-skill development consists of three process; i.e. cognitive, union, and automatic ones. In a cognitive layer, a learner of motor-skill understands the goal of the target motion and its strategy for the goal. Nextly, in a union layer, the learner practices repeatedly so that s/he can perform the coordinated movement smoothly. In the overtimes practice, the learner needs the feedback in order to proceed effectively. Finally, the learner may get to the automatic layer. In the automatic layer, the learner can perform the skill smoothly and perform without consciousness.

Schmidt (1991) supposes that human performs a skill based on a process. The process consists of “stimulus identification”, “response selection”, and “response program”. In the “stimulus identification”, human recognizes an environment human exists and trajectory of limbs. The trajectory means a trace of own motion. In the “response selection”, human decides internal models within a brain and a physical manner. In the “response program”, human makes a program to perform an action. The action is a body representation of an internal model.

In motor-skill development, a learner in a union layer of Fitts needs to recognize a gap between an ideal motion and a selected internal model. The authors assume that real-time feedback makes easy to recognize the gap. Since the learner performs the skill based on the strategy, the action of a learner in a moment should attain the strategy. In order to attain the strategy, the learner needs to grasp the difference between an ideal motion and a present motion. Moreover, the learner needs to grasp the influence of the present motion on the future. That is, our real-time feedback should include a future’s motion predicted from the present motion.

Gentile (1972) classifies motor-skill with common criteria. Therefore, we can apply the supporting method according to the classification. For example, a motor-skill is classified as Open-skill or Closed-skill. In the Closed-skill, a performer should perform fixed motion. Therefore, the learner of Closed-skill needs feedback after the end of motion. The feedback should represent a difference between a target motion and the performer’s motion. This feedback enables the performer to obtain routinized and fixed motion. In the Open-skill, a performer should determine the contents of motion along with what the motion changes with time and situations. Therefore the learner should obtain a capability to perceive a situation and obtain a capability to judge contents of motion. We assume that real-time feedback should be considered of the skill’s classification.

In order to show predicted motion, it is necessary to calculate subsequence of future’s motions using present motion data. There are some researches to predict time series data. For example, Autoregressive analysis(AR) and Moving average (MA) are fundamental methods as a prediction with time series data. However, there are few related works about prediction of multi-dimensional data for a human-motion. Moreover, if the authors create the differential equation about the motion using the physical model, it may be possible to predict the next motion by solving the present state analytically. Since it is necessary to create a differential equation for every skill, the authors cannot use general-purpose supporting system. Then, the authors collect a learner's data and calculate the similarity of the stored data and input data. The authors predict future’s motion using data with the highest similarity. Moreover, the authors predict the performance result using similar motion's data. The authors think similar motion has similar performance result. Since AR, or MA and a differential equation predict only motion, these methods cannot predict performance result.

Predicting future’s motion is a motion recognition problem. In other words we classify current data into the past data. There are some researches about motion recognition. For example, many techniques, such as a Fourier transform, discriminant analysis, principal component analysis, and a support vector machine, are proposed. Moreover, Yamada, et al. (2013) recognize a motion by HMM. However, this method recognizes the motion from all sequential data. Therefore, we cannot recognize the motion using sub-sequence of motion data. It means that we cannot predict the future's motion while a performer conducts a motor-skill. Then, the authors calculate similarity about all data of posture and presume the present state. If the present state can be specified, the remainder of the motion is a future motion. A problem is that there is much computational complexity in order to calculate similarity with all data of posture. In this research, the authors create a key to a dataset, and reduce computational complexity.

3. Supporting Scenario

3.1 Preparations

In our proposal, some preparations are necessary before learning. Figure 1 expresses the outline of the preparation that a learner should do until the learning. It also shows the flow of the preparation of the system side. In step (1)(on left side), it is checked whether a skill is "Closed-skill" or not. Otherwise, a learner divides into sub-skill to make "Closed-Skill". In step (2), a learner collects the data of motion and the result of performance about a skill. In process (I)(on right side), our system saves the data of motion and the result of performance into the database. Step (1), (2), and process(I) show preparations. Step (3) and process (Ⅱ) shows practices receiving actual feedback.

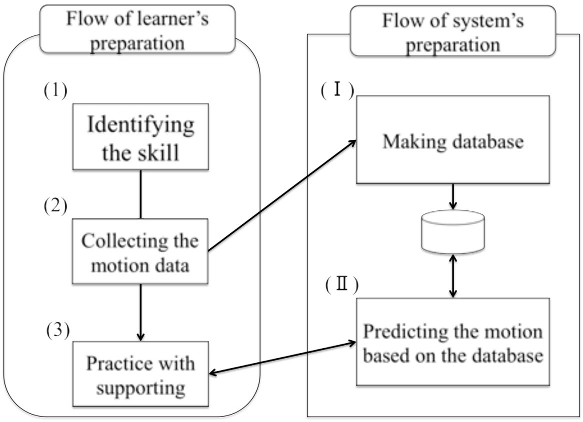


Figure 1. Flow of preparations

3.2 Contents of real-time feedback

Figure 2 shows the outline of supporting. At first, in (a), the system gets a learner's data of posture. In (b), the system finds the similar data of the posture from a database. In (c), the system shows the predicted performance result to the learner as a feedback. By this feedback, the learner can update selected motion. Furthermore, the system shows the sequence of the posture that may happen to the next of the present posture as a feedback. By this feedback, the learner can recognize a gap of own action and past action. In (d), the learner updates own selection. The learner corrects own motion by repeating (a)~ (d).

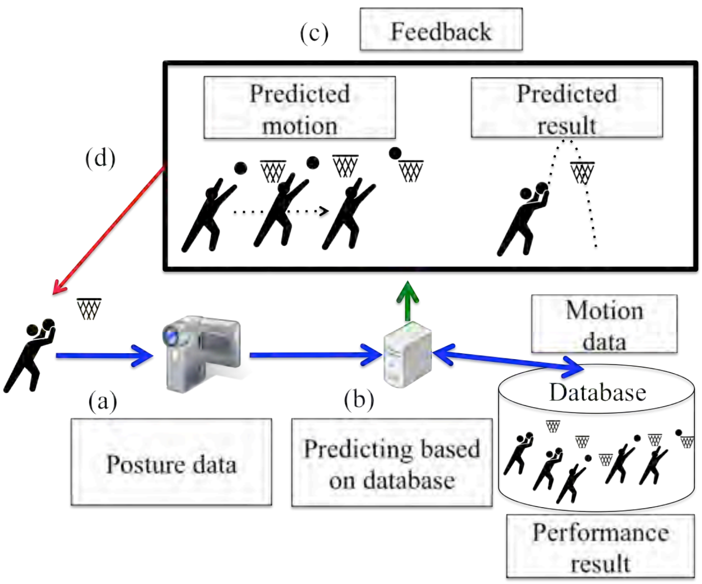


Figure 2. Flow of supporting system.

4. Development of Supporting System

Our system obtains data of the posture using a motion capture. The motion capture is the optical type in our system, while there are optical type, mechanical type, magnetic type, and so on. By using motion capture, posture's data of time-series can be obtained by the three-dimensional data of limbs. Since it is related to performance, the authors don't obtain the information about a tool using motion capture. When a performance is related to the locus of a motion, the authors may obtain the information of a tool using motion capture. Our supporting system predicts motion using accumulated data. Furthermore, a data structure that enables high-speed recognition is required. In our system, an input is multi-dimensional data describing the posture. Moreover, the similar data is required as an output. A simple algorithm that fills these inputs and outputs is to calculate similarity with all posture's data in a database. The similarity between postures is sum of Euclid distance about three-dimensional coordinate in each limbs. Therefore computational complexity is 3 x (number of limbs) x (number of postures) x (number of forms in database). However, depending on the number of the stored data, it is not a realistic method. Then, by segmenting the data using the k-means algorithm, the authors make other data structure, which can search faster than the simple algorithm. Segmenting a motion with the posture, or its acceleration, the authors create the average posture. When our system has posture data as an input, our system calculate similarity with this average posture.

The authors use the k-means algorithm for segmentation. A k-means algorithm is a kind of non-hierarchical clustering. A given dataset is classified into k clusters using the average of a cluster(hereinafter centroid). The k-means clustering consists of following procedure.

1. A cluster's label is assigned at random to each data.
2. The centroids are created using an arithmetic average.
3. Calculating a distance between a data and each centroid of the new cluster. A cluster that has minimum distance is assigned to the data.
4. Procedure 2~3 are repeated. In "procedure 3", when no cluster updates, a loop becomes an end. Furthermore, the loop is an end when the criteria of updating are cleared.

The reason of using k-means algorithm is that a motion of motor-skill doesn't have clear hierarchy.

The authors describe an algorithm applying a k-means algorithm to segment the motion. Firstly, if a target skill is continuous skill, the learner saves a motion of one cycle to the database. The next procedure is following.

i. A data is divided into k pieces at equal intervals in order of a time series, and a cluster's label 0~(k-1) is assigned to each data. The reason why the first cluster is not random is for the data to change according to time series.
ii. The centroid is calculated for every data. The centroid is an arithmetic average of the coordinates of each limb.
iii. Our system calculates the similarity between the centroid and each data. Our system assigns the cluster's label using the minimum Euclid distance.
iv. "Procedure ii~iii" is repeated until updating is lost.

Thus, the centroid is created with a cluster. The learner or the instructor using this system determines the number of clusters. For example, they measure the acceleration of the limbs and determine the number of clusters on the basis of the number of peaks. In addition, if the number of typical postures is decided, they use it. Moreover, the reason for using the Euclid distance, the authors think that it is easy for the learner to understand an absolute distance between parts.

The authors show a flow that the system searches for a similar data. The searching procedure is as follows.

A) The posture's data is input into the system.
B) The system calculates the Euclid distance with each centroid that exists in the database.
C) The system determines the cluster with the shortest distance.

The computational complexity is 3 x (limbs) x (number of centroids) x (number of forms in database) by using the k-means method. Compared with a simple method, it has succeeded in reducing the computational complexity. Since the centroid has the index set, the our system feed back the posture data with the next index as predicted data.

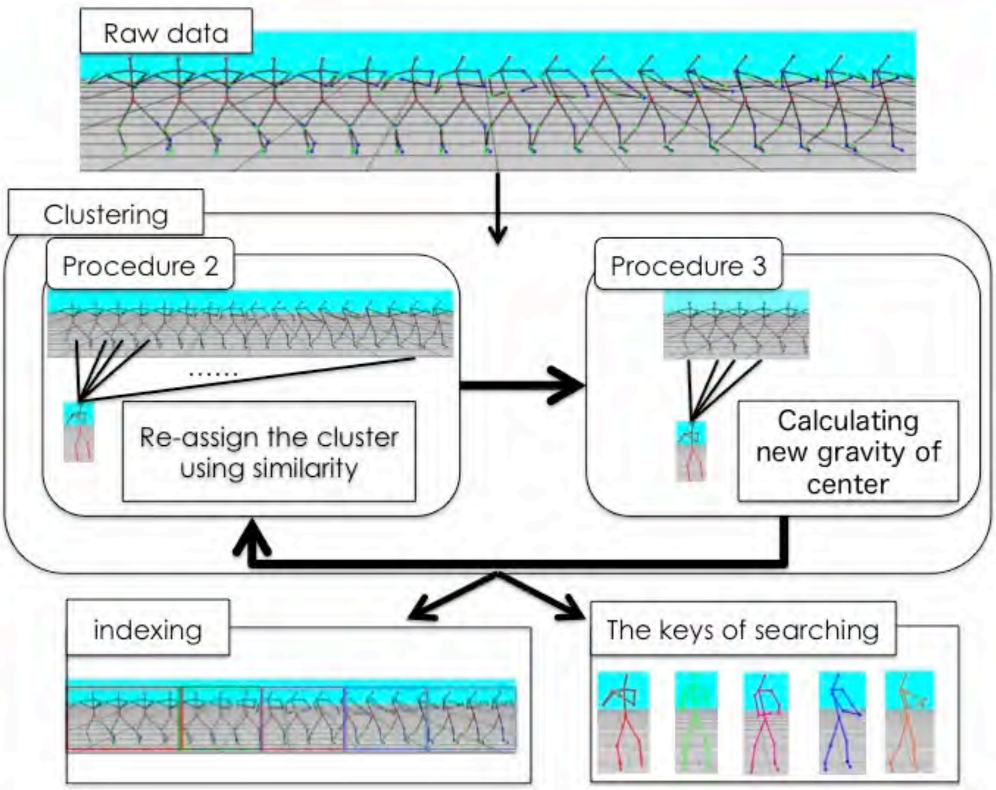


Figure 3. Flow of clustering the motion data

Figure 3 illustrates the flow of creating the keys with proposal algorithm. "Raw data" shows three-dimensional data of a motion and it is the form of the batting of baseball. Since this motion has "establishing", "backswing", "beginning to shake", "impact", and "follow-through", these five states are set to the keys. "Clustering" in Figure 3 describes "procedure2" and "procedure3". "Procedure2" in Figure 3 describes calculating the similarity for all data and updating the cluster. In "procedure3", the new centroid is created based on the updating a data. "The keys of searching" expresses the centroids in searching. "indexing" shows assignment of cluster's label to each posture's data.

5. Evaluation Method

The authors conduct an experiment for verifying the prototype system. The authors propose the system that performs real-time feedback for closed skill. By using this system, recognition of the gap with "response selection" and "response program" becomes easy. This system makes efficiently for a learner to acquire a closed skill. The authors think that the learner using the system can acquire the motion much earlier compared with no supporting learner. Then, the authors create a control group and an experimental group. Next, a learner practices respectively. The authors measure an effect without supporting of the system after one practice. Figure 4 express a flow of an experiment. In "Pre-test", the learner performs the target motion without supporting of the system. The authors record the form to verify the effect of the support. In "First learning", the learner looks at animation of an ideal motion to know what is an ideal motion. In "First learning", the learner doesn't perform the skill actually. In ”Learning with support”, the learner practices with an actual movement. In an experimental group, the learner practices based on the proposed supporting system. In the control group, the learner practices without the supporting system. In "Post-test", the learner performs the target motion without the supporting system. The authors record these forms and calculate a difference between the forms in "Post-test", and the motion in "Pre-test".

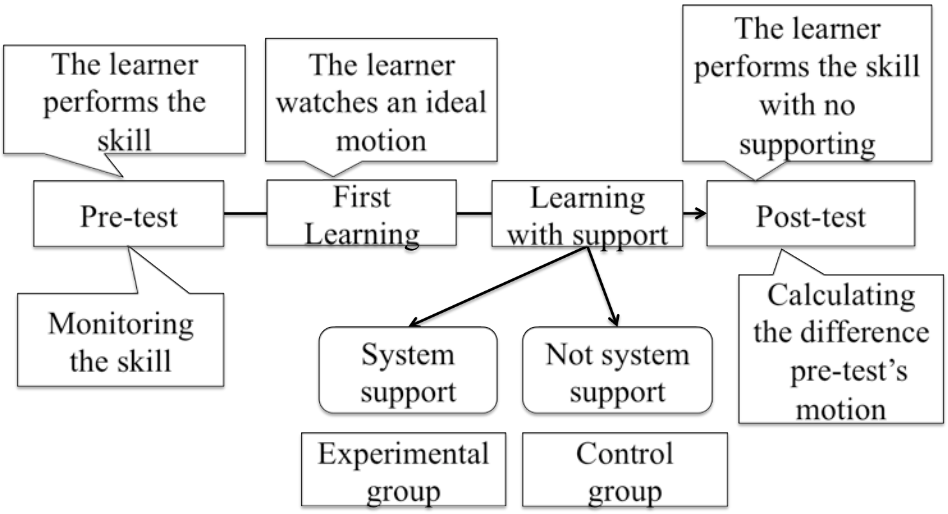


Figure 4. Flow of experiment

6. Concluding Remarks

The authors propose the environment for motor-skill development based on the real-time feedback using sensing technology, such as motion capture. From the information on the present form, the system predicts the next form and performance. Furthermore, the system feed back a predicted result to a learner. The authors think the learner who is in a union layer acquires the skill earlier. About prediction, the system calculates the similarity with the past learner's data. Furthermore, the system outputs the performance result with the largest similarity of forms. As a technique on the computer, our system segments the motion data using k-means algorithm. The computational complexity is reduced by the clustering. In order to verify this support model, the authors should apply the concrete skill.

References

Fitts, P.M., & Possner, M.I. (1967). Human Performance. Oxford, England: Brooks and Cole.

Gentile, A. M. (1972). A Working Model of Skill Acquisition with Application to Teaching. *Quest*, 17, 3-23.

Hamagami, K., Matsuura, K., Kanenishi, K., & Gotoda, N. (2012). Support on Repeating Skill Development - Modulating from monitored data to a target. *Workshop Proceedings of the 20th International Conference on Computers in Education*, 491-498.

Kosaka, M., Matsuura, K., Gotoda, N., Ogata, O., Tanaka, T., Kanenishi, K., & Yano, Y. (2011). System Design to Improve Running-form with Motion-capture, *Proceedings of ICISS2011*, 260-264.

Schmidt, R.A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, Vol. 82, No. 4, 225-260.

Schmidt, R.A. (1991). Motor Learning and Performance: From Principles to Practice. *Champaign*, IL. :Human Kinetics.

Yamada, K., Matsuura, K., Hamagami, K. & Inui, H. (2013). Motor skill development using motion recognition based on an HMM. *17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems* - KES2013, 37-45.

Zimmerman, B.J. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166-183.

Content Management System to Support Improvement in Quality of Fitness Testing of Athletes

Yuji KOBAYASHI[a*] & Naka GOTODA[b]
[a]*Graduate School of Instructional Systems, Kumamoto University, Japan*
[b]*Department of Sports Sciences, Japan Institute of Sports Sciences, Japan*
*yuji-k@kumamoto-u.ac.jp

Abstract: Fitness testing is conducted to evaluate athletes' physical fitness and provide scientific evidence and data that will help improve their performances. The validity and reliability of the data need to be guaranteed, and this depends on the proficiency of staff in charge of testing. However, there are several problems in staff training. In this study, we propose a content management system to support staff training improvement for quality assurance of athletes' fitness testing.

Keywords: Quality assurance, fitness testing, elite sports, skill proficiency, CMS, blended learning

1. Introduction

The Japan Institute of Sports Sciences (JISS) conducts surveys and measurements for several factors of physical fitness that help determine sports performance. Furthermore, JISS (2014) provides knowledge and data to help athletes improve their performances by indicating their fitness strengths and weaknesses. Thus, test results are used to set appropriate individual training intensity. Moreover, test results can help implement athletes' training programs efficiently and effectively. Whether a current training program is working well can also be confirmed from these test results, and the athlete's progress can be monitored. Additionally, data on elite athletes and averaged data on many athletes can become criteria for talent identification and transition. Therefore, the validity and reliability of the testing should be guaranteed.

The Australian Institute of Sport (AIS, 2014) conducts a quality assurance program for fitness testing, called the National Sport Science Quality Assurance Program. Since JISS was established in 2001, it has also standardized fitness testing by creating a unified manual. However, in implementing its staff development program, JISS is faced with various problems and constraints.

In this study, we present an overview and the problems of the JISS staff development program. Furthermore, we propose a content management system (CMS) to support the staff training program and solve its problems.

2. Staff Training for Proficiency at JISS

2.1 Overview of staff training for proficiency at JISS

At JISS, staff proficiency training comprises workshops and seminars on fitness testing, self-learning and self-practice, and proficiency checking (Figure 1).

2.2 Workshops and seminars on fitness testing

At JISS, workshops and seminars on fitness testing are held for new staff, called “rookie staff,” during the second week of April. In these workshops, the rookie staff attend lectures on basic knowledge and cultivate an appropriate attitude for fitness testing of athletes. For acquisition of skills related to fitness testing, first, the rookie staff observe demonstrations and then practice fitness testing using the actual measurement equipment.

2.3 Self-directed learning and self-practice

After these workshops and seminars, the rookie staff manage their own training through self-learning and self-practice. For self-learning, they study the manual created by expert JISS staff. Based on testing procedures in the manual, they practice fitness testing on each other to develop proficiency. For each fitness test, they use a checklist to confirm their proficiency (Table 1).

2.4 Skill proficiency check-up

After the rookie staff complete self-learning and self-practice, a “skill proficiency check-up” is administered to confirm their skill and knowledge. This check-up is conducted under the same conditions as actual testing of athletes. Its flow contains preparation of the test, the actual testing of the athlete, feedback to the athlete, and finally, storing the equipment and materials. Throughout this flow, several evaluators judge the individual rookies’ skill levels using a check-sheet especially designed for the evaluation and also associated with the self-learning and self-training checklist. During and after the check-up, individuals are questioned about their knowledge. Evaluators perform comprehensive evaluation of whether each rookie’s knowledge and skill achieves the evaluation criteria.

The rookie staff receive their individual evaluation results, and those who are successful are then qualified to perform fitness testing of athletes. On the basis of feedback, unsuccessful staff study and practice to improve their skills and then re-take the check-up.

Workshops and seminars about fitness testing
- Checking and acquiring basic knowledge
- Observing and experiencing fitness testing of athletes

↓

Self-learning and self-practice
- Acquisition of knowledge of fitness testing
- Mastery and proficiency of testing skills using the manual and check sheets

↓

Skill proficiency check-up
- Confirmation of understanding of fitness testing of athlete
- Confirmation of skill proficiency

Figure 1. Flow of measurement training.

Figure 2. Biodex System 4.

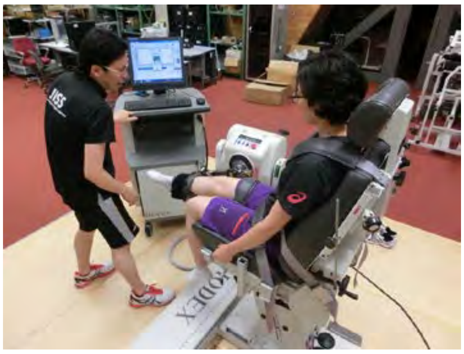


Figure 3. Scene of Measurement.

Table 1: Example of checklist (example of equipment: Biodex System 4).

Procedure	Evaluation point
PC start-up	Following the manual's instructions
Database connection	Selecting the correct database
Subject registration and selection	Following the manual's instructions Registering using the correct method
Protocol selection	Selecting the correct protocol
Explanation prior to measurement	Explaining clearly and correctly
Seating subject on Biodex System 4	Confirming the subject's physical condition (e.g., confirming the subject's injury) Following the manual's instructions
Setting center of rotation	Confirming and setting correct center of rotation
Setting range of motion	Following the manual's instructions (e.g., determining the position of maximal extension)
Warm-up instructions	Implementing warm-up adequately and sufficiently
Explanation of measurement flow	Explaining clearly and correctly
Consideration of subject	Conversing with subject appropriately
Confirmation of measurement completion	Checking continues, and completing data collection Creating a measurement report in accordance with the manual's procedure
Explanation of measurement result	Explaining results based on correct knowledge.
Total time for the procedure	Completing testing within 30 min from measurement preparation to explanation of result (Measurement time, approximately 20 min)

3. Problems in Improving Testing Skill

This section provides a description, along with some specific examples, of the problems in self-learning and self-practice. First, Figure 1 displays the conceptual flow of self-learning and self-practice in some detail. For "Acquisition of knowledge of fitness testing," if practice has been inadequate, rookie staff are likely to remember only the fragmentary procedure shown in Table 1. In other words, although the rookie staff memorize each procedure, they do not understand the subtleties of the "evaluation point" and the linkages between procedures.

Furthermore, because the manual and checklist are intended to summarize the procedure, not for all athletes but only for the "standard" athlete, rookie staff must learn to adjust when testing individual athletes. Therefore, when the correct value is not obtained from the actual measurement, the staff must review and readjust the measurement procedure, and then re-measure. Of course, only one correct measurement is desirable. However, it is crucial that staff recognize an inaccurate or inadequate measurement, diagnose what is wrong with the procedure, make adjustments, and conduct re-measurement.

For example, during the measurement of thigh muscles' isokinetic strength with a dynamometer (Biodex System 4, New York, USA) (Figures 2 and 3), the graph of measured force is displayed as Figure 4. The peak value is adopted as an individual record. However, this device displays the waveform of measured force and records its peak value as correct data even if correct measurement is not conducted. In Figure 5, the athlete could not perform as strongly as compared with the athlete in Figure 4. In other words, the staff's explanation and the athlete's practice and/or warm-up are insufficient. Additionally, compared with the athlete's leg shown in Figure 6, that shown in Figure 7 is not fixed firmly. This may result in impulses or noise being measured in the waveform of force even if the measurement process itself is correct, as in Figure 8. Therefore, it is important for the staff to recognize these waveforms and incorrect points.

For acquisition of the ability to diagnose ‘what is wrong’, the rookie staff should practice on the actual equipment. Nevertheless the time and equipment for practice are limited in the current training program. In fact, even the practice equipment is occasionally used for actual fitness testing, further limiting practice opportunities.

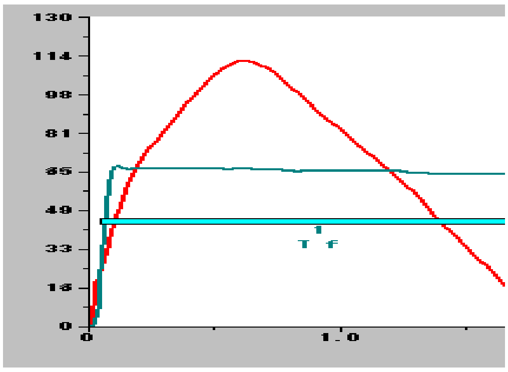


Figure 4. Example of correct waveform of force

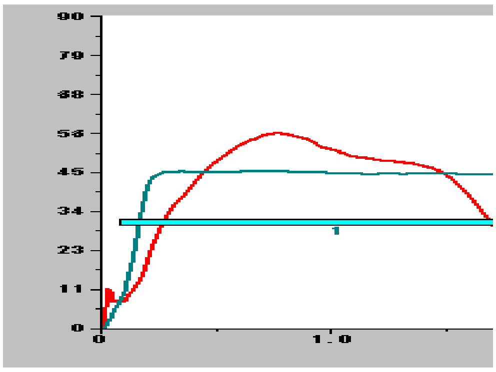


Figure 5. Example of wrong waveform of force

Figure 6. Leg fixed firmly

Figure 7. Leg not fixed firmly

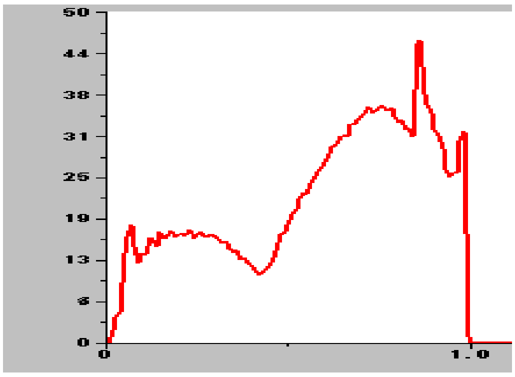


Figure 8. Example of wrong waveform including noise or impulse

4. CMS to Support Improvement of Testing Skill

4.1 Proposal of blended learning program and CMS to support the program

As explained, the current rookie staff training program depends on self-learning and self-practice. However, limited time and equipment availability cause some program inefficiency.

Thus, we suggest a blended learning program and CMS for self-learning, as depicted in Figure 9. CMS services can compensate for the lack of face-to-face instruction in the program. Because JISS is the domestic representative of athletes' fitness testing, this model can be expected to create a ripple effect on various relevant organizations.

4.2 Online Video Textbook: Delivery of practical instructional video materials

Because the checklist was created from the evaluators' standpoint, rookie staff have often faced difficulty in understanding its content. Furthermore, opportunities to observe demonstrations by skilled staff are limited.

Therefore, we propose to create video teaching materials that contain explanations and demonstrations by skilled staff, and to deliver it in the CMS. These materials will help rookie staff efficiently understand the subtleties of the "evaluation point" and the linkages between the fragmentary procedures. Moreover, even without equipment, opportunities for improved self-learning are provided. During their limited free time, rookie staff can learn fitness testing.

4.3 Online Test Sheet: Introduction of practice and test materials for preparation and review

As mentioned, practice time with the testing equipment is limited. Delivery of practice and test materials for preparation and review in CMS will support the currently limited learning opportunities with actual equipment and face-to-face instruction by skilled staff. The new preparation materials will help rookie staff learn the basic knowledge required for fitness testing in advance. Additionally, preparation materials will enable rookie staff to review the required skill proficiencies objectively.

4.4 Applied Q & A: Problematic case collection system

In actual fitness testing, responding flexibly to various athletes is essential. Therefore, accumulating experience of incidental failures and problematic cases is very important. Problematic cases, insolvable through use of the manual alone, should be collected for consideration of solutions and for practice problems.

Accordingly, we propose the introduction of a support system based on problem-based learning (PBL) in CMS. PBL with CMS for practical staff training in several domains, for instance, in nursing, (Majima, So, & Seta, 2006) has been proposed; however, in our domain, it has not been proposed. If skilled staff participate assertively in the collection activity, this system will help educate rookie staff and support skilled staff in charge of actual fitness testing.

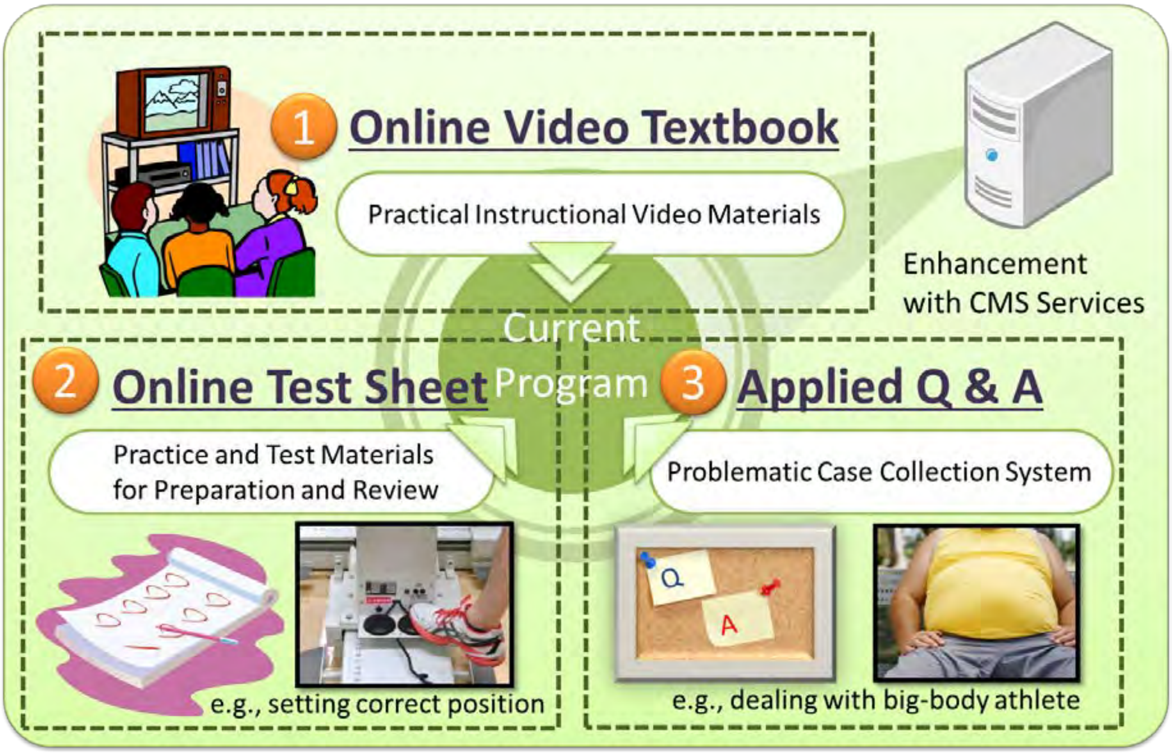


Figure 9. Proposal of blended learning program with CMS Services

4.5 Flow of Support Scenario with CMS

Supportive CMS services for effectively and efficiently promoting skill proficiency are not independent, but mutual. In fact, Figure 10 illustrates an example of support by the services shown in Figure 9. In correcting the test in Figure 8 of section 3, in many cases, rookie staff cannot understand that the problem involves fixing the leg firmly from the beginning, in addition to the evaluation points and procedures. Therefore, before recognizing the problem, but after watching the video of the first step as an initial learning strategy, rookie staff can confirm the key point at the second step.

After that, the next learning step shifts to PBL. Input by expert staff triggers reflective learning activities, such as re-watching and re-testing. Repeated self-learning can help rookie staff deeply understand the relationship between the evaluation point and the procedures. Thus, applied Q & A provides linkages with the online video textbook and online test sheet. Throughout the flow of testing, rookie staff can enhance their understanding of the evaluation point and linkage of procedures, as in Table 1, to solve current problems.

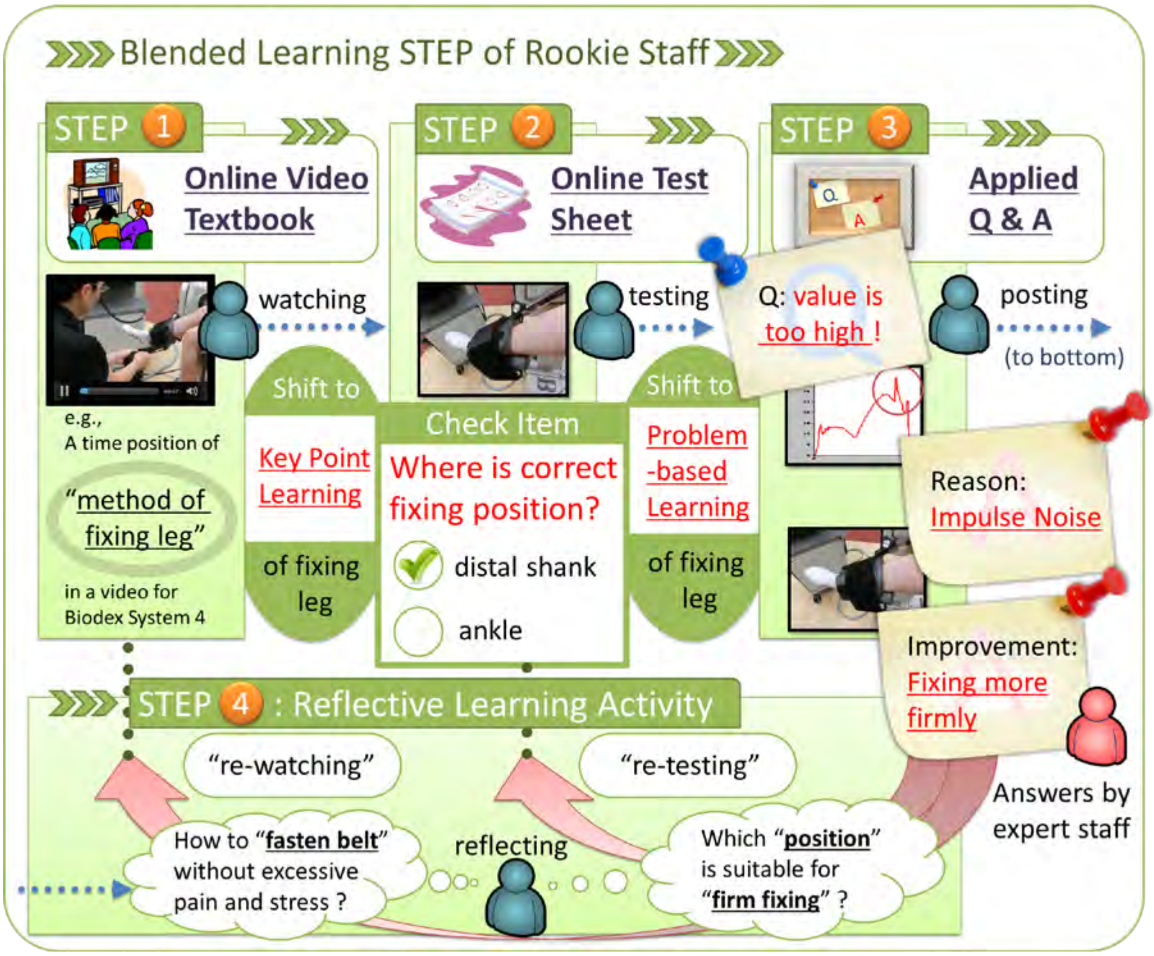


Figure 10. Example of flow regarding self-learning with CMS services

5. System Configuration

5.1 Overview of server and client system through web interface

Figure 11 illustrates the overall system configuration between a CMS server and the client PCs. The server provides HTTP Service by Apache. All staffs can utilize CMS on the web interface through a browser. Essentially, however, the server is accessible on a closed network via private LAN or Internet with VPN. This is because, presumably, limited client PCs, permitted to handle personal data, can only connect to the server. CMS is developed with an open source named WordPress as the fundamental environment. WordPress is a typical and widespread CMS used for blogging in social network services. Therefore, many people, including rookie staff, are familiar with the front-end GUI, with the exception of the system control GUI such as administrator pages. Moreover, official or other developers have released various plug-ins and provide an extension method with a programmable plug-in that has several PHP frameworks. The following proposal functions are implemented as the module.

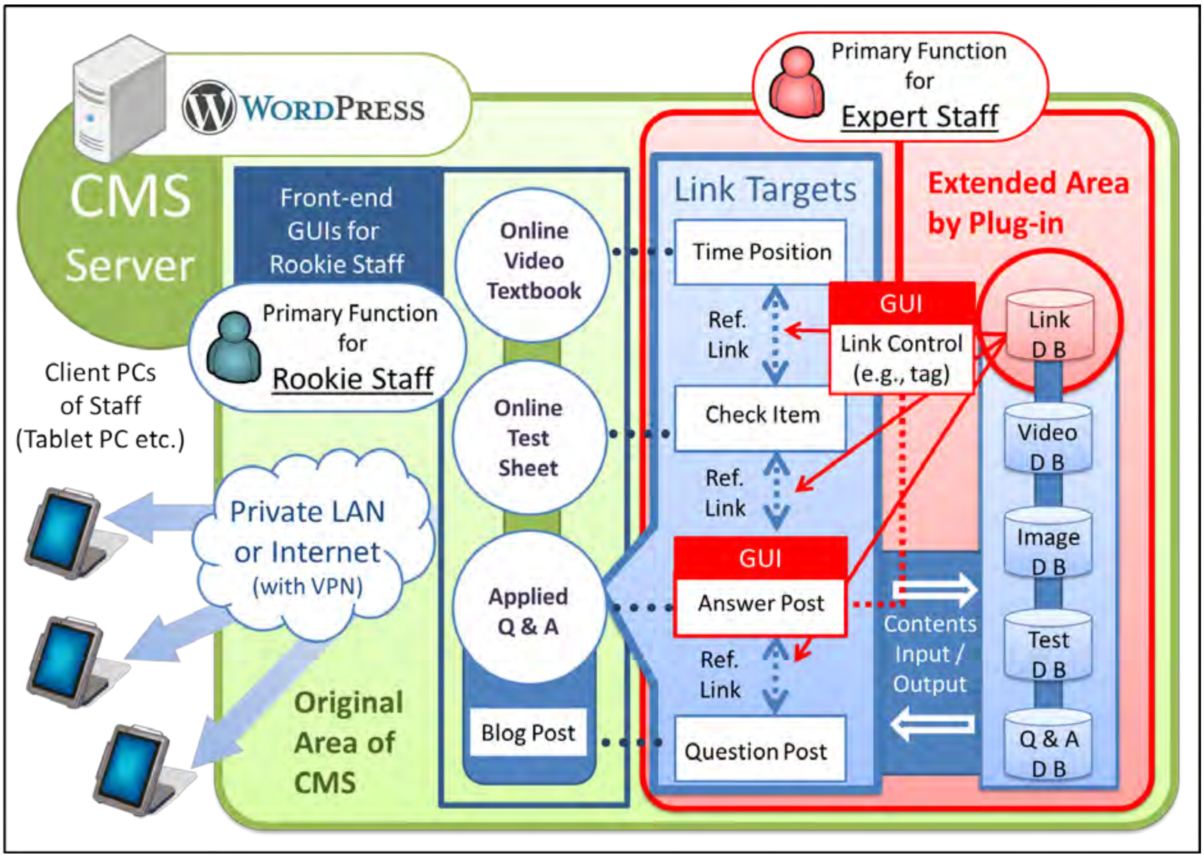


Figure 11. Overall system configuration including server and clients

5.2 Primary function for rookie staff

From the viewpoint of information literacy and for user-friendliness, all action by rookie staff would be within the front-end GUI. The almost internal architecture of CMS services without relational linkages between services is based on the original function of CMS, similar to a blog post. Both video play and test services are also constructed and customized by distributed common plug-in. On this basis, the function enables rookie staff to concentrate on self-learning and self-practice.

5.3 Primary function for expert staff

In contrast, expert staff at JISS generally have a sufficient level of literacy, but are extremely busy at work. Therefore, expert staff members have a limited amount of time to produce learning materials, and they mainly use an exclusive front-end GUI, similar to the administrator page from short-term efficiency for one-to-many rookie staffs. We will develop the GUI based on the extension module and an additional relational database of MySQL. Using the GUI, expert staff can add reference linkages as reflective learning materials, in addition to answering posted questions. Expert staff will have indirect roles as "learning navigators" during the rookie staff's reflective stage in self-learning, as illustrated in the lower part of Figure 9.

Moreover, these linkages are built effectively with a short code framework. Once a certain rule of short code is preliminarily defined, expert staff can easily create a flexible reference linkage in a text form like WYSIWYG. Speaking more concretely, the staff can create a simple, logical program with a placeholder and a conditional placeholder with a short code framework. For example, an expert staff member can control an automatic display of recommended linkages that enable deep learning, along with the individual results of online test sheets.

6. Summary

By introducing CMS to support the staff training program for athletes' fitness testing, rookie staff can be expected to learn quickly and improve their proficiency effectively and efficiently. Furthermore, a problematic case collection system, based on PBL, can be helpful to both rookie and expert staff. Finally, this system can expand to support the solution of problems occurring in actual fitness testing.

Acknowledgements

This study was a part of the research project "Creation of Research and Diffusion Center of Instructional Systems: Development and Diffusion of an Intensive Certificate Package to Train Education Specialists for Supporting Recurrent Adult Graduate Students," supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Furthermore, this study was partly supported by JSPS KAKENHI Grant Numbers 25750097 and 25871206. The authors would like to thank Enago (www.enago.jp) for the English language review.

References

Australian Institute of Sport. (2014). Quality assurance. Retrieved from: http://www.ausport.gov.au/ais/sssm/quality_assurance

Japan Institute of Sports Sciences. (2014). Research project of sports science and medicine. Retrieved from: http://www.jpnsport.go.jp/jiss/gaiyou/jigyou/tsc/tabid/254/Default.aspx

Majima, Y., So, Y., & Seta, K. (2006). Framework for Problem-Solving Based Learning in Nursing Domain: A Practical Study. *Proceedings of the 2006 Conference on Learning by Effective Utilization of Technologies: Facilitating Intercultural Understanding*, pp. 625-628.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Development of a Learning Environment for Novices' Erhu Playings

Fumitaka KIKUKAWA[a*], **Masato SOGA**[b**] **& Hirokazu TAKI**[b]
[a]*Graduate school of Systems Engineering, Wakayama University, Japan*
[b]*Faculty of Systems Engineering, Wakayama University, Japan*
*s145062@gmail.com
**soga@sys.wakayama-u.ac.jp

Abstract: In this paper, we describe the development of a learning environment for novices' erhu playings. In case of self-learning, the novices who want to learn playing bowed string instruments give up because they get into a wrong habit and don't understand how to play. Therefore, our goal is to develop a learning environment for novices. We focus on the erhu in bowed string instruments. The system we developed diagnoses finger positions and bow strokes by using magnetic position sensors and gives awareness of errors by showing finger positions and bow strokes on virtual 3D space. In case of bow strokes, the system shows how to improve bow strokes by using arrows. Moreover, the system shows how and where the errors occur on score. Accordingly, the novices can recognize and improve their errors by using this system. In future works, we will conduct evaluation experiments to disclose learning effects by using this system.

Keywords: Bowed string instruments, skill, learning environment, learning support, electro-magnetic field sensing device

1. Introduction

In this paper, we describe the development of a learning environment for novices' erhu playings. Because of the following backgrounds, we aimed to develop a learning environment for playing instruments.

Usually, to learn skills for playing instruments is achieved by means of taking a lesson, practicing by self-learning, referring to a study-aid book, and so on. In case of self-learning, the novices give up learning playing instruments because they get into a wrong habit, don't understand how to play instruments and don't make progress. Recently, there are many studies about analyses of skills and learning environments because of progress of sensors and computers performances. In case of the piano, there is an analysis of finger movements of a pianist (Rahman et al., 2010). Moreover, there is a learning environment regarding the piano (Takegawa et al., 2012). In case of the guitar, there are visual methods for the retrieval of the guitarist fingering as well (Burns et al., 2006). Furthermore, there is a learning environment of guitarist fingering (Tobise et al., 2013). In case of the drums, there is a motion analysis for emotional performance of snare drums as well (Miura et al., 2012). In addition, there is a learning environment of playing the drums (Iwami et al., 2007).

Particularly, the bowed string instruments (e.g. the violin, the cello, and the erhu) have many parameters to be adjusted. The parameters are the finger positions, the pressures by pressing strings, the gesture, the speed, the acceleration, and the angle of bow strokes. Accordingly, it is difficult for the novices to play with accurate pitches. Therefore, if we develop a learning environment for playing the bowed string instruments, these problems will be solved. As related works, there are many studies about analysis of playing the violin. Rasamimanana et al. (2005) analyzed the gesture of bow strokes by using augmented violin bow. Moreover, they analyzed bow and arm movements and bow pressure on strings by using 3D optical motion capture system and a custom pressure sensor (Rasamimanana et al., 2007). Furthermore, Carrillo (2013) characterized bow strokes by using audio and a motion capture system. In addition, Maestre et al. (2007) analyzed bow strokes by using electro-magnetic field sensing device. On the other hand, there are studies about learning environment of playing the violin. Wang et al. (2012) developed a real-time pitch training system for violin learners. However, this system does not diagnose learner's gesture parameters and pressure parameters, but diagnoses only sound pitches. Therefore, the novices cannot identify the cause of pitch errors. Thus, we aimed to develop learning environment for novices to understand why pitch errors were caused. Moreover, Ng (2009) developed 3D motion analysis and visualization system. This system visualize only bow stokes. Therefore, we aimed to develop learning environment for novices to understand bow strokes and finger position improvement simultaneously.

Although there are many studies about musical skills, there is no study about the erhu except for our studies. As our previous works, we describe following three studies. We have aimed to develop a learning environment for playing the bowed string instruments and have developed learning environment. Firstly, we have analyzed novices' parameters during playing a bowed string instrument (Soga et al., 2010). Specifically, we analyzed the erhu playing parameters by novices. As the results of the analyses, novices need supports of finger positions, bow motions, bow speeds, bow accelerations, and bow angles. The sounds depend on these parameters. A novice doesn't understand the reasons why the sounds aren't accurate because there are many parameters to be adjusted. Therefore, a novice needs a learning support environment for controlling each parameter. Secondly, we tried to assist novices' finger positions. We developed a learning environment regarding finger position on strings (Kikukawa et al., 2013). By conducting evaluation experiment, we found the system is useful for a novice to learn finger position on strings. Thirdly, we tried to assist novices' controls of bow strokes. We developed a gesture learning environment for novices' erhu bow strokes (Kikukawa et al., 2014). In this paper, we describe integration of previous two systems and improvement of the systems. By integrating previous two systems, the new system visualizes improvements of finger positions and bow strokes simultaneously and gives the novice learner feedback of diagnosed results of finger positions and bow strokes simultaneously for novices' understanding.

2. Instruments Choice

2.1 The Erhu

We have chosen the erhu in bowed string instruments. Figure 1 shows an erhu. There is a reason that we have chosen the erhu in bowed string instruments. It is a two-stringed instrument. Because of this feature, a learner only needs to judge which string s/he touches in those two strings with a bow. In case of the violin, a learner needs to judge which string s/he touches in four strings with a bow. Therefore, his/her hand and finger motions are very complicated when s/he learns the violin. It is comparatively easy for novices to learn how to play the erhu. Moreover, the erhu has the parameters that other bowed string instruments have (e.g. the finger positions, the gestures of bow strokes, the speed of bow strokes, the acceleration of bow strokes, the angle of bow strokes, and so on) in spite of only two strings. Therefore, to develop learning environment of playing the erhu is comparatively easy and can apply to other bowed string instruments.

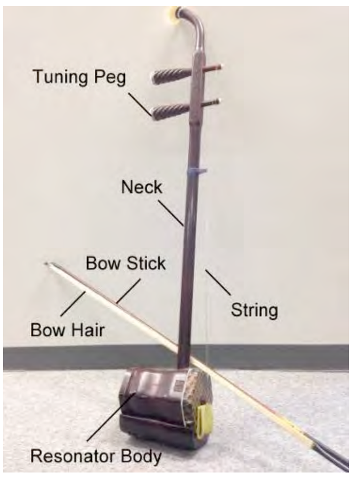


Figure 1. The erhu.

2.2 How to Play

Figure 2 shows how to play the erhu. How to play the erhu is as follows. Left fingers press strings with pitches as Figure 2(a). Then, it needs to press accurate finger positions. A pitch error reflects a finger position error directly because the erhu has no border between pitches. On the other hand, right hand holds a bow from underneath as Figure 2(b). Then, right index finger is attached to the wood part and right middle finger and ring finger control hair tension of the bow. A learner plays sounds by moving the bow between right and left, by the frictions between the bow and the strings. Loudness and expression of sounds depend on pressures between the bow and the strings, moving speed of the bow, and how to move the bow.

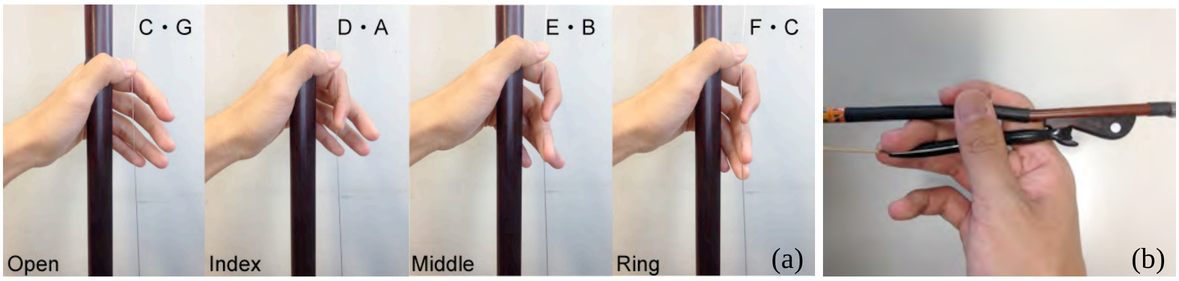


Figure 2. How to play the erhu.

2.3 Bad Example of Erhu Bow Strokes

Figure 3 shows bad examples of erhu bow strokes. The bow strokes should not leave the resonator body as Figure 3(a). Moreover, the gesture of erhu bow strokes should be parallel to the ground and parallel to the surface of the body. Therefore, Figure 3(b) and Figure 3(c) are bad examples. Based on the above, we aim to develop the system that makes the novice learner being aware of these bad examples and learning accurate gesture of erhu bow strokes.

Figure 3.Bad examples of erhu bow strokes.

3. System Design

3.1 Electro-magnetic Field Sensing Device LIBERTY

LIBERTY is an electro-magnetic field sensing device developed by Polhemus Inc. A transmitter and up to 16 receivers are connected to a main unit for use. Figure 4(a) shows LIBERTY. Figure 4(b) shows transmitter of LIBERTY. Figure 4(c) shows receivers of LIBERTY. A transmitter generates electro-magnetic fields which have three directions by passing an electric current in turn through coils around three axes. Receivers also have coils around three axes. When the transmitter generates an electro-magnetic field, induced electric currents are generated on the coils in the receiver. Then, position and orientation of the receiver are calculated by this amperage. These measured data are transmitted to a connected PC with ASCII or binary data format. LIBERTY is connected with PC through USB ports or serial ports. The accuracies are 0.03 inch RMS for X, Y or Z position and 0.15 degrees RMS for sensor orientation. The reason why we choose LIBERTY as a motion tracking sensor is that we need to use real-time accurate 3D tracking data for real-time visualization and real-time diagnosis of finer positions and bow strokes.

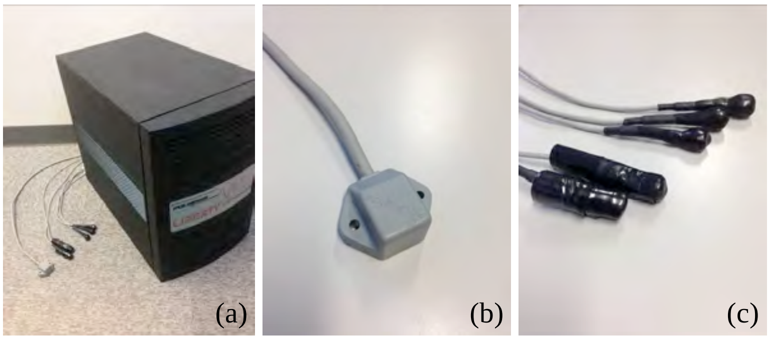


Figure 4. Electro-magnetic field sensing device LIBERTY.

3.2 *System Composition*

Figure 5 shows system composition. The system consists of a PC, an electro-magnetic field sensing device LIBERTY and an erhu. The erhu is equipped with a transmitter and receivers of LIBERTY. Figure 6(a) shows the erhu equipped with a transmitter and receivers of LIBERTY. Figure 6(b) shows left hand equipped with receivers. We equipped the erhu with a transmitter and receivers for measuring gesture of erhu bow strokes as Figure 6(a). Moreover, we equipped learner's index finger, middle finger, and ring finger of left hand with receivers for measuring learner's finger position on string as Figure 6(b). The system diagnoses learner’s gesture of erhu bow strokes and finger positions by using data measured by the receivers. The learner learns playing skills about bow strokes and finger positions by using this information.

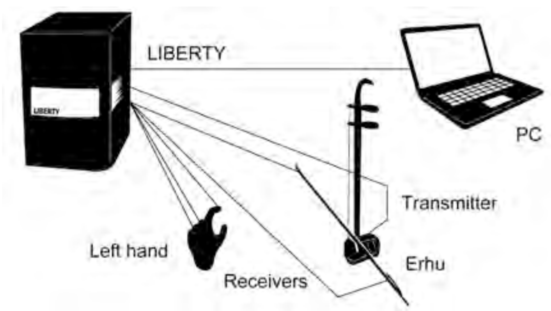


Figure 5. System composition.

Figure 6. The erhu and left hand equipped with LIBERTY.

3.3 *Visualization Window*

Figure 7 shows visualization window. It shows bow strokes and finger positions with virtual erhu model in the virtual 3D space. These bow strokes and finger positions represent learner’s bow strokes and finger positions measured by electro-magnetic field sensing device on real-time. The learner's playing is diagnosed by the system. Viewpoint in the virtual 3D space is changeable freely. The default viewpoint is matched with the real learner's viewpoint while playing the erhu.

The results of diagnosis about bow strokes are represented as follows. During accurate bow strokes, they are colored green as Figure 7. Figure 9 shows visualization window while playing bad examples. If bow strokes leave the resonator body (over 0.5 inch), they are colored black as Figure 9(a). If bow strokes are not parallel to the ground (over 10 degrees), they are colored red or blue as Figure 9(b) and Figure 9(c). If bow strokes are not parallel to the surface of the body (over 10 degrees), they are colored pink or orange as Figure 9(d) and Figure 9(e). Viewpoint in the virtual 3D space is changeable freely. Each feedback shows 3D arrows for indicating how to improve bow strokes. The learner can recognize bow strokes errors by watching visualization window.

The window shows totally six colored small spheres on the virtual erhu model's strings as Figure 7. Three spheres of them represent learner's left finger positions. The positions of index finger, middle finger, and ring finger are colored light red, light green, and light blue, respectively. The other three spheres represent correct finger positions. The correct positions of index finger, middle finger, and ring finger are colored red, green, and blue, respectively. The correct positions can be switched to display/non-display. If the learner's finger position is accurate, the light red sphere, the light green sphere, and the light blue sphere change the sky blue sphere. The learner can recognize finger positions errors by watching visualization window.

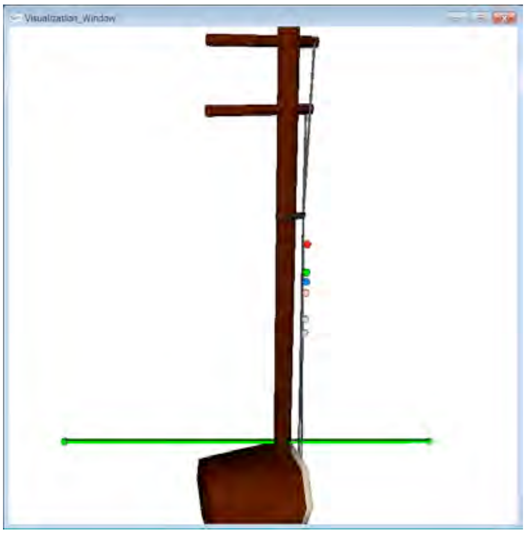


Figure 7. Visualization window.

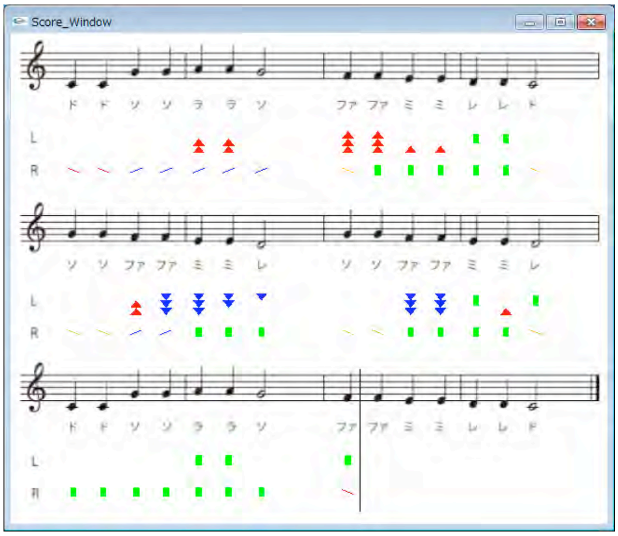


Figure 8. Score window.

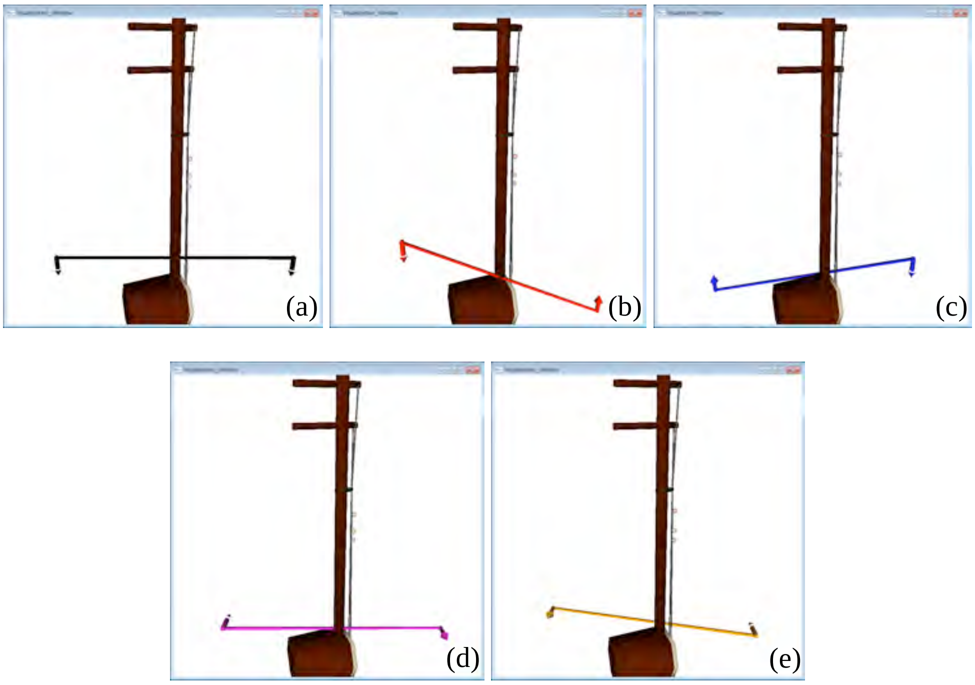


Figure 9. Visualization window while playing bad examples.

3.4 Score Window

Figure 8 shows score window. It shows score of the music, current diagnosing point on the score, and results of diagnosing. The current diagnosing point is shown by a black vertical line, and it moves along the score. A learner must play the erhu by synchronizing the current playing note with the current note on the black line. The results of diagnosis are indicated under the notes. 'L' in Figure 8 is results of diagnosis left finger positions. The red triangles represent that learner's finger position is upper than correct finger position. The blue triangles represent that learner's finger position is lower than correct finger position. The number of red/blue triangle represents the size of difference between learner's finger position and correct finger position. The green squares represent that learner's finger position is accurate. Moreover, 'R' in Figure 8 is results of diagnosis right hand bow strokes. The green squares represent that learner's bow stroke is accurate. The other colored figures that correspond to the representation of bow strokes in visualization window. The learner can recognize where s/he made errors of bow strokes and finger positions and how to improve his/her errors.

4. Conclusion and Future Works

In this paper, we designed and developed a learning environment for novices' erhu playings. The system diagnoses bow strokes and finger positions by using electro-magnetic field sensing device LIBERTY. The system has two windows, visualization window for visualization of bow strokes and finger positions on real-time, score window for showing the results of diagnosing bow strokes and finger positions along score. The novices can recognize and improve bow strokes and finger positions errors by using these functions. Therefore, we have achieved the aim.

In the future, we will conduct quantitative evaluation experiments for verifying learning effect during using the system by the data of electro-magnetic field sensing device.

Acknowledgements

We would like to thank all the people who cooperate with our studies.

References

Burns, A.M., & Wanderley, M.M. (2006). Visual Methods for the Retrieval of Guitarist Fingering. *Proceedings of NIME 2006*, 196-199.

Carrillo, A.P. (2013). Characterization of Bowing Strokes in Violin Playing in Terms of Controls and Sound: Differences between Bouncing and On-string Bow Strokes. *Proceedings of Meetings on Acoustics*, *19*, 1-8.

Iwami, N., & Miura, M. (2007). A Support System for Basic Practice of Playing The Drums. *Proceedings of ICMC 2007*, 364-367.

Kikukawa, F., Ishihara, S., Soga, M., & Taki, H. (2013). Development of A Learning Environment for Playing Erhu by Diagnosis and Advice regarding Finger Position on Strings. *Proceedings of NIME2013*, 271-276.

Kikukawa, F., Soga, M., & Taki, H. (2014). Development of a gesture learning environment for novices' erhu bow strokes. *Proceedings of KES2014*, (in press).

Maestre, E., Bonada, J., Blaauw, M., Perez, A., & Guaus, E. (2007). Acquisition of Violin Instrumental Gestures Using a Commercial EMF Tracking Device. Proceedings of ICMC 2007, 386-393.

Miura, M., Mito, Y., & Kawakami, H. (2012). Motion Analysis for Emotional Performance of Snare Drums Using a Motion Averaging Method. *Proceedings of ICMPC 2012*, 682-685.

Ng, K. (2009). 3D Motion Data Analysis and Visualisation for Technology-Enhanced Learning and Heritage Preservation. *Proceedings of AIKED '09*, 384-389.

Rahman, Md.M., Mitobe, K., & Suzuki, M. (2010). Analysis of Finger Movements of a Pianist Using Magnetic Motion Capture System with Six Dimensional Position Sensors. *TVRSJ*, *15*(2), 243-250.

Rasamimanana, N.H., Bernardin, D., Wanderley, M.M., & Bevilacqua, F. (2007). String Bowing Gestures at Varying Bow Stroke Frequencies: A Case Study. *Proceedings of GW2007*, 216-226.

Rasamimanana, N.H., Fléty, E., & Bevilacqua, F. (2005). Gesture analysis of violin bow strokes. *Proceedings of GW2005*, 145-155.

Soga, M., Ishihara, S., Ota, S., & Taki, H. (2010). Analysis of erhu playing by novice player - First step to skill learning environment for erhu playing -. *Proceedings of ICISS2010*, 695-700.

Takegawa, T., Terada, T., & Tsukamoto, M. (2012). A Piano Learning Support System considering Rhythm. *Proceedings of ICMC 2012*, 325-332.

Tobise, H., Takegawa, Y., Terada, T., & Tsukamoto, M. (2013). Construction of a System for Recognizing Touch of Strings for Guitar. *Proceedings of NIME 2013*, 261-266.

Wang, J., Wang, S., Chen, W., Chang, K., & Chen, H. (2012). Real-Time Pitch Training System for Violin Learners. *Proceedings of ICME Workshops*, 163-168.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Analysis and Feedback of Baseball Pitching Form with use of Kinect

Yasuhisa TAMURA [a*], Taro MARUYAMA [a] & Takeshi SHIMA [b]
[a] *Department of Information and Communication Sciences, Sophia University, Japan*
[b] *Department of Health and Physical Education, Sophia University, Japan*
*ytamura@sophia.ac.jp

Abstract: This paper introduces an automatic analysis and feedback system for amateur baseball pitching learners with use of Microsoft Kinect. Compared with conventional 3-D motion capture systems, Kinect has advantages of reasonable cost and easiness of application system development for physical exercises. The authors developed a similar system for flying disc throw, which is disclosed in Tamura, Yamaoka et al. (2013) and Tamura, Uehara et al. (2013). In this paper a target motion is moved to baseball pitching, so focused body parts and judgment criteria were changed, although utilizing the same Kinect platform. The proposing system acquires postures and motions of amateur baseball pitchers, and judges them in 2 criteria: (1) maximum angle of elbow, and (2) hand twist, which were decided by a training expert. It also displays feedback messages to improve their motions. As a result, novice testees of the target group showed significant improvement of their pitching motions.

Keywords: Baseball, pitching movement, Kinect, analysis, feedback

1. Introduction

In a field of sports science research, motion analysis of human body has become popular in the last decade. Barris and Button (2008) surveyed vision-based motion analysis researches for sports. Moeslund et al. (2006) surveyed vision-based human motion capture / analysis systems. Miles et al. (2012) surveyed applications of Virtual Reality environments for ball sports. There are wide variety of equipments adopted in these researches: GPS sensor, acceleration sensor, muscle sensor, HMD (Head Mound Display) etc. Among them, the major equipments are so called “motion capture systems”, that measure many points of human body in three-dimensional space in real time manner. However, the major commercial motion capture systems are extremely expensive, costing several hundreds of thousand dollars. Additionally, they require a dedicated equipped room, multiple cameras, special lighting facility and dedicated “tracking suits” to specify a tracking points of human body. Furthermore, myriad steps are necessary to set up and acquire data including the activity called “calibration”, which adjusts the 3-D points of marking sensors on the tracking suit. As a result, this kind of analysis is infrequently performed other than specialized researches or specific studies of limited top athletes.

Among them, Microsoft Corporation released a device called Kinect in 2010. It provides a simple and inexpensive way to perform 3-D analysis of a human body movement. First, the device itself costs only U.S.$110, which is far cheaper than conventional motion capture systems. Second, Kinect is capable of capturing data easily. It does not need any tracking suits, complex set-up, and troublesome operation procedure for data acquisition. Third, Microsoft has publicly released a software development kit (SDK) that includes necessary software libraries for data acquisition using Kinect. Application system developers are able to write customized Windows applications with use of this library in the C# or C++ languages. The proposed research in this paper has 3 major points below:
(1) Utilizes Kinect
(2) Captures 3-D motion and give feedback to sports learners
(3) Target motion: baseball pitching

There are many preceding researches to analyze human body motion with use of motion capture systems including Kinect. Also, there are some researches to give automatic feedback messages to learners to refine their motion. The authors arranged these researches as shown in Table 1 in general.

Table 1: Preceding Researches

	Analysis	Feedback
Commercial/ Original 3D Motion CaptureSystem	Bideau et al. (2004) Brodie et al. (2008) Corazza et al. (2006) Hachimura et al. (2004)	Ishii et al. (2011) Kwon and Gross (2005) Soga and Myojin (2008)
Microsoft Kinect	Fujimoto et al. (2012) Hsu (2011) Kato et al. (2012) Marquardt et al. (2012) Mitchell and Clarke (2011) Ogawa and Kambayashi (2012)	Chye and Nakajima (2012)

Papers at upper left side in Table 1 utilize commercial or original 3D motion capture systems to analyze 3-D motion. Bideau et al. (2004) utilized Vicon 370 system to analyze relationship of movement between throwers and a goalkeeper of handball. Brodie et al. (2008) synthesized a body model of a ski racer from GPS information and video motion graphics. Corazza et al. (2006) synthesized a body model with use of 8 motion cameras and replays it in a virtual environment. Hachimura et al. (2004) developed a dance training support system with use of magnetic sensor system Fastrak and HMD.

At upper right side, there are researches to give feedback messages to learners, based on 3-D captured data. Ishii et al. (2011) utilized a motion capture system IGS-190 for baseball batting movement. It also provided a comparing function between “goal motion” and learner’s one. Based on the comparison, the system showed messages to refine learner’s motion. Kwon and Gross (2005) developed an original motion capture system for Taekwondo training. It also displayed a visual feedback to adjust one’s movement. Soga and Myojin (2008) proposed a training support system for rhythmic gymnastics. It adopted an optical motion capture system, compared the captured data and ideal motion data, and displayed feedback messages in the screen.

At lower left side in Table 1, there are researches to analyze human motion with use of Kinect. Fujimoto et al. (2012) developed a dance training support system. It showed learner’s image and instructor’s ideal motion image in overlaying manner. Hsu (2011) discussed many possibilities of Kinect utilization in various sports learning activities. Kato et al. (2012) developed a system to compare a professional player and a novice learner of soccer. Marquardt et al. (2012) diagnosed a pose of ballet dancer with use of Kinect. It is called “Super Mirror”, because common ballet studios use a mirror to check and adjust one’s pose. Mitchell and Clarke (2011) developed a Kinect based system to diagnose hand movement for playground game. Ogawa and Kambayashi (2012) developed a distance learning system. An instructor and a learner share a common virtual space, and compare their body motions.

Finally, at the lower right side, there is one preceding research similar to the proposing method. Chye and Nakajima (2012) utilized Kinect to diagnose Karate pose. He compared 4 joint points of an instructor and a learner, calculates their Euclid distances, and gives feedback messages to the learner.

As mentioned above, this research focuses on the motion of baseball pitching. According to this topic, there are some related researches. Lin et al. (2003) used Qualisys motion analysis system by Qualisys AB. The goal is to analyze the movement and velocity of “center of mass” in the pitching cycle. They found different moving speeds in different pitching phases (lowest position and late cocking). Theobalt et al. (2004) used a proprietary, multiple video camera-based motion capturing system. Aguinaldo et al. (2007) used Real-Time Motion Capture System by Motion Analysis Corp (120 fps). They set 18 marker points for upper body, in order to analyze peak shoulder internal rotation torque. 38 pitchers are categorized into 4 groups (Pro, College, High School, Youth). The

Professional group showed a significant result of rotation timing. Lapinski et al. (2009) developed a proprietary wearable sensing device with use of Analog Devices iMEMS ADSXR300 and ADXL210E, that also provide a wireless communication function. For reference, they utilized a motion analysis system from XOS Technologies with 10 motion analysis cameras. They analyzed acceleration and G-Force of pitchers' body points (hand, waist, forearm, chest and waist) and compared the results between their original system and optical motion analysis system. Ukita et al. (2014) developed a Kinect-based automatic rehabilitation system, and applied its motion as a baseball pitching in left tow and right hand velocity.

In above papers, most of them are utilizing commercial or dedicated motion capture systems. Only Ukita et al. (2014) utilized Kinect, but its focus is rehabilitation. In this sense, there is no preceding research to focus on baseball pitching training with use of Kinect device.

2. Proposed System

This paper proposes a system that will process data in three steps:
(a) acquisition of 2-D video images and 3-D position data for body points with use of Kinect,
(b) assessment of whether the baseball pitching movement is adequate or not, based on the position data acquired in (a), and
(c) display of feedback messages with 2-D motion image from (a) based on the results of the assessment in (b).

Details of these processes are given below.

2.1 Kinect and its Data Acquisition

Kinect is a device with a function to analyze the motion of human subjects in real time and 3-D manners. It was initially developed as a peripheral device to be connected to Microsoft's Xbox gaming system. Kinect includes a CMOS camera, infrared projector, image depth sensor, microphone, and one USB port for connection to a Windows PC. Kinect projects patterned infrared rays that are analyzed by CMOS camera to recognize the distance between the target player and the device. Also, with use of a machine learning function called "human pose estimation" developed by Microsoft Research Cambridge, Kinect is able to recognize 20 positions of target human joints and body parts with reasonable accuracy. 3-D coordinates of these positions are transferred to a connected Windows PC in 30 fps (frames per second) through the libraries included in the device's SDK.

2.2 Judgment of Pitching Form

This paper focuses on the assessment of baseball pitching movement. However, target skill levels of learners are totally diverse in beginner and professional learners. This study focused on absolute beginners and made assessments by comparing whether their pitching motions match a basic and standard one. The authors set two hypothetical criteria of the pitching motion: (a) maximum angle of elbow, and (b) hand twist. One of the authors (Shima) analyzed these aspects and tendencies during his coaching in college classes. They are shown in Figure 1.

The aspect (a) judges whether the right elbow (for a right-hand pitcher) is raised in enough height of not. An amateur pitcher tends not to raise his right elbow enough. In order to judge this point, the authors calculated the angle of hip center (#0) – shoulder center (#2) – right elbow (position #9) as θ_1. These numbers of 0, 2 and 9 are specified in Kinect SDK to identify 20 human joints and body parts. Kinect is able to measure the absolute height of elbow position, but stature varies on individual. Then in this paper, this angle of $\theta_1$was decided as a criterion. This θ_1 is judged as below. Points shown in parenthesis are used to evaluate performance at the time of experiment shown in Chapter 3.

- $125 <= \theta_1$: too much (35 points)
- $115 <= \theta_1 < 125$: a little much (40 points)
- $105 <= \theta_1 < 115$: good (50 points)
- $90 <= \theta_1 < 105$: a little few (40 points)

- $\theta_1 < 90$: too few (10 points)

The aspect (b) is twist of the right hand (for a right-hand pitcher). In order to throw a fast ball, right elbow should go forward rather than his body. In other words, his elbow should be twisted, not be strait. In order to measure it, the authors focus on the angle of pitching target – right elbow (#9) – right hand (#10) as θ_2. In this paper, this angle of θ_2 is judged as;

- $110 <= \theta_2$: good (50 points)
- $100 <= \theta_2 < 110$: moderate (40 points)
- $\theta_2 < 100$: not enough (30 points)

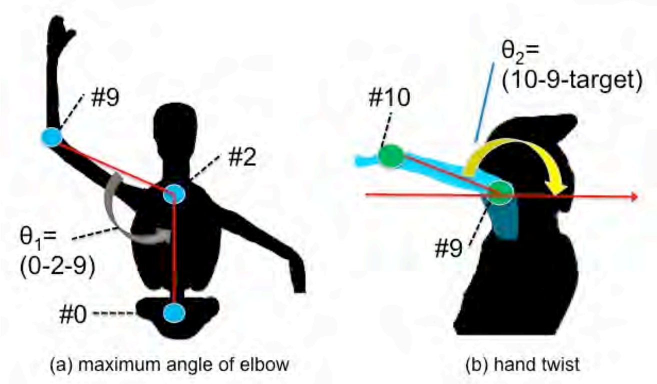


Figure 1. Judgment of Pitching Form

In the previous studies (Tamura, Yamaoka et al. (2013) and Tamura, Uehara et al. (2013)), the target motion was flying disc throw. In this case, the main focus was height of right hand (#10). If this height changes widely, he is difficult to throw the disc to the intended target. However, in the baseball pitching, the main issues are the maximum angle of elbow and hand twist, as shown in Figure 1. This difference can be implemented after the judgment criteria is identified and quantified, but the criteria itself should be determined in discussion with a professional sports trainer.

As introduced in Chapter 1, there are some preceding researches to analyze baseball pitching forms. These researches focus on mainly speed, torque, acceleration and G-force of body points. However, they have little insights why and how the speed etc. come from. In other words, They do not show a "good" pitching form. In this paper, the authors set a hypothesis of a "good" form, and show judgment criteria as stated above.

2.3 Visual Feedback

In order to provide visual feedback to pitching learners, the authors also developed a visual feedback interface. It contains a real time 2-D video of pitching action and both quantitative numbers and qualitative feedback sentences of form judgment result described in section 2.2. Figure 2 shows a snapshot of the developed visual feedback interface. The proposing system is reset when a target person raises his left hand and then starts the analysis.

3. Experiment

In order to verify the effectiveness of the proposing system described above, the authors performed an experiment. The authors collected 40 testees in Sophia University. Some of them belong to athletic clubs in the college, but none of them have been trained as a baseball pitcher.

First, as a pre-test, all of the 40 testees were examined the precision of pitching forms 5 times. After that, the testees were divided into two groups of a target group (TG) of 20 and a control group (CG) of 20, whereas pitching performances of the two groups are to be statistically non significant. Next, as a test, the TG members were given feedback in 15 times of pitching movements with use of the proposing system. On the other hand, the CG members had no feedback from the proposing system in same 15 times of pitching. Last, as a post test, all members including TG and CG were assessed their pitching forms with use of the proposing system, without any feedback.

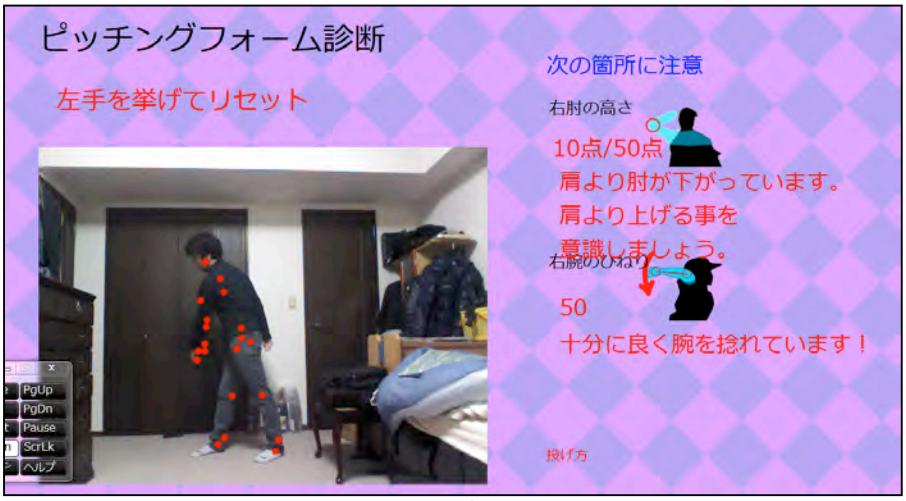


Figure 2. Screen Shot of the Proposing System

The authors compared the result of the post test, focusing on improvement of their performance. The result of t-test among all TG and CG testees, p value was p=0.227, which showed there was no significance between TG and CG in performance improvement. Then, in order to perform more detailed analysis, total testees were divided into six groups according to TG and CG, and also their performance groups; upper, intermediate, and lower. When TG and CG are compared in the same groups, p values were: Upper group: p=0.515, Intermediate group: p=0.706, and Lower group: p=0.061.

This result shows that the p value of lower groups shows a tendency of significance, while upper and intermediate groups are not significant totally. This phenomenon is similar to the experiment performed previously in Tamura, Yamaoka et al. (2013) and Tamura, Uehara et al. (2013).

4. Discussion & Conclusion

The result of experiment in Chapter 3 shows a tendency of significance only for relatively lower-graded learners. A possible reason of this point is that judgment points and their evaluations hypothesized in Section 2.2 fit exercises for novice learners of baseball pitching. In other words, there should be some other judgment points and evaluations for upper than intermediate learners. This point is thought as one of the future issues.

As a conclusion, this paper has presented a system with use of Kinect device for analysis of and feedback on the motion of baseball pitching. A result of experiment shows that this method is thought to be useful for relatively novice learners to improve their movement. Future research issues contain a work to refine the current system vis-à-vis the points noted in Chapter 3, and retry to validate the efficiency of the proposing system with improved methods and sequence.

Acknowledgements

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grand-in-Aid for Scientific Research (C) 25350779, 2013.

References

Aguinaldo, A. L., Buttermore, J. & Chambers, H. (2007). Effects of upper trunk rotation on shoulder joint torque among baseball pitchers of various levels, Journal of Applied Biomechanics, 23(1), 42.

Barris, S. & Button, C. (2008). A Review of Vision-based Motion Analysis in Sport, Sports Medicine, 38(12), 1025-1043.

Bideau, B., Multon, F., Kulpa, R., Fradet, L., Arnaldi, B. & Delamarche, P. (2004). Using virtual reality to analyze links between handball thrower kinematics and goalkeeper's reactions, Neuroscience Letters, 372(1-2), 119-122.

Brodie, M., Walmsley, A. & Page, W. (2008). Fusion motion capture: a prototype system using inertial measurement units and GPS for the biomechanical analysis of ski racing, Sports Technology, 1(1), 17-28.

Chye, C. & Nakajima, T. (2012). Game based approach to learn Martial Arts for Beginners, IEEE 18th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), 19-22.

Corazza, S., Muendermann, L., Chaudhari, A. M., Demattio, T., Cobelli, C. & Andriacchi, T. P. (2006). A Markerless Motion Capture System to Study Musculoskeletal Biomechanics: Visual Hull and Simulated Annealing Approach, Annals of Biomedical Engineering, 34(6), 1019-1029.

Fujimoto, M., Terada.T. & Tsukamoto. M. (2012). A Dance Training System that Maps Self-Images onto an Instruction Video. ACHI 2012 (The Fifth International Conference on Advances in Computer-Human Interactions), IARIA, 309-314.

Hachimura, K., Kato, H. & Tamura, H (2004). A Prototype Dance Training Support System with Motion Capture and Mixed Reality Technologies, Proc. IEEE International Workshop on Robot and Human Interactive Communication, 20-22.

Hsu, H. J. (2011). The Potential of Kinect as Interactive Educational Technology, 2nd International Conference on Education and Management Technology, 13, 334–338.

Ishii, K., Soga, M. & Taki, H. (2011). Proposal and development of reversal motion skill leaning support environment, International Conference on Information Technology Based Higher Education and Training (ITHET), 1-6.

Kato, K., Namiki, Y., Thepvilojanapong, N., Tetsutani, N. & Ohta, Y. (2012). Exploring Analysis of Football Actions Cosidering Correlation of Joints in the Body, Ninth International Conference on Networked Sensing Systems (INSS), 11-14.

Kwon, D.Y. & Gross, M. (2005). Combining Body Sensors and Visual Sensors for Motion Training, ACM SIGCHI International Conference on Advances in computer entertainment technology, 94-101.

Lapinski, M., Berkson, E., Gill, T., Reinold, M. & Paradiso, J. A. (2009). A distributed wearable, wireless sensor system for evaluating professional baseball pitchers and batters, Proc. ISWC'09, 131-138.

Lin, H. T., Su, F. C., Nakamura, M. & Chao, E. Y. (2003). Complex chain of momentum transfer of body segments in the baseball pitching motion, Journal of the Chinese Institute of Engineers, 26(6), 861-868.

Marquardt, Z., Beira, J., Em, N., Paiva, I. & Kox, S. (2012). Super Mirror: a kinect interface for ballet dancers, Proc. 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts (CHI EA '12), 1619-1624.

Miles, H.C., Pop, S. R., Watt, S. J., Lawrence, G. P. & John, N. W. (2012). A review of Virtual Environments for Training in Ball Sports, Computers & Graphics, 36, 714-726.

Mitchell, G. & Clarke, A. (2011). Capturing and Visualising Playground Games and Performance: A Wii and Kinect based Motion Capture System, Proc. Int'l Conf. Electronic Visualisation and the Arts (EVA'11), 218-225.

Moeslund, T.B., Hilton, A. & Krüger, V. (2006). A Survey of Advances in Vision-based Human Motion Capture and Analysis, Computer Vision and Image Understanding, 104, 90-126.

Ogawa, T. & Kambayashi, Y. (2012). Physical Instructional Support System Using Virtual Avatars. ACHI 2012, The Fifth International Conference on Advances in Computer-Human Interactions. IARIA, 262-265.

Soga, A. & Myojin, Y. (2008). Educational Applications for Learning Rhythmic Gymnastics Rules Using Motion Data, Journal of Media Information and Television Engineering, 62(2), 222-226. (in Japanese)

Tamura, Y., Yamaoka, K., Uehara, M. & Shima, T. (2013). Capture and Feedback in Flying Disc Throw with use of Kinect, Int'l Journal of Sports Medicine and Sports Sciences, 74, 313-317.

Tamura, Y., Uehara, M., Maruyama, T. & Shima, T. (2013). Feedback of Flying Disc Throw with Kinect: Improved Experiment, Proc. ICCE 2013, 196-201.

Theobalt, C., Albrecht, I., Haber, J., Magnor, M. & Seidel, H.P. (2004). Pitching a baseball: tracking high-speed motion with multi-exposure images, ACM Transactions on Graphics (TOG), 23(3), 540-547.

Ukita, N., Eimon, K. & Röcker, C. (2014). Mining Crucial Features for Automatic Rehabilitation Coaching Systems, Int'l Workshop on User-Centered Design of Pervasive Healthcare Applications.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Integration of Augmented Reality Mobile Learning and Self-Regulated Learning by using Concept Mapping - A Case Study of the Plants in Campus

Po-Han WU[a*], Gwo-Haur HWANG[b], Yu-Syuan WANG[a] & Yen-Ru SHI[c]
[a] *Department of Mathematics and Information Education, National Taipei University of Education, Taiwan*
[b] *Department of Information Networking and System Administration, Ling Tung University, Taiwan*
[c] *Department of Information and Learning Technology, National University of Tainan, Taiwan*
*cincowu@gmail.com

Abstract: With the rapid development of mobile devices, ubiquitous learning becomes the new trend of digital learning. In recent years, many scholars found that learning through augmented reality methods can improve students' interest and motivation. Therefore, in this study, an integration system of augmented reality mobile learning and self-regulated learning by using concept mapping is proposed for conducting learning activities. Students can observe the plants in campus and draw the concept maps by augmented reality technology and self-regulated learning strategy. According to the drawn concept maps and the learning contents from real situations, the system can provide students the guidance and feedback to help their thinking. Students can set their learning goals and adjust their learning progress according to the nature of the learning tasks via self-regulated strategies. The experimental results were expected that the proposed approach is able to improve the students' the cognition of the plants for students in campus and raise students' learning motivation and effectiveness.

Keywords: Augmented reality, mobile navigation, self-regulated learning, concept mapping, campus plant

1. Introduction

In traditional teaching environment, it is difficult to express the abstract concept of knowledge. Thus, with the improvement of information technologies, many scholars tried to import some multimedia technologies into courses in the past. And some scholars found out that the multimedia learning systems can improve students' learning motivation and effectiveness (Govaere Jan, de Kruif, & Valcke, 2012; Uluyol & Agca, 2012). But students think that multimedia learning systems are not in real environments. Because of the difference between virtual environments and real environments, students cannot associate learning contents with everyday life. (Barfield, Sheridan, Zeltzer, & Slater, 1995; Eysenck & Keane, 2005). However, augmented reality (AR) is a technology combined with virtual and reality images (Azuma, 1997). Thus, students can directly scan learning objects in real environments and get the relational teaching contents immediately. In the past, many scholars applied AR technologies in classes. For example, Martín-Gutiérrez et al. (2010) used AR technologies in learning the spatial integration ability in geography, history, arts and humanities. The results of their study showed that AR technologies can raise the attraction of students and have the easy manipulation characteristics, and it can improve their space ability. Wernhuar and Ou (2012) developed a smartphone butterfly ecological system in virtual environments by AR technologies. The results of their study showed that AR technologies for learning can improve students' learning effectiveness. Hwang et al. (2013) developed a multi-language learning system by AR technologies. This system can appeared correspondence learning contents when students scan the objects in real environments. In addition, it can also provide multi-language spelling and voice teaching by the objects. The purpose of this system is to let students study languages. At the same time, it can also let students practice pronunciation.

In traditional teaching environments, teachers faced too many students and cannot consider the individual learning situation for students. Thus, the learning effectiveness is not very good (Hwang, Tsai, & Yang, 2008; Wu, Hwang, Milrad, Ke, & Huang, 2012). In the past, some scholars indicated that students can use mobile electronic devices in real situations. For students, those methods cannot only solve the problems in one-to-many teaching environment (Wu, Hwang, Su, & Huang, 2012), but also can let students study according to their own learning pace. For teachers, they can act the leader roles and understand the students' learning situations from the systems.

Bandura (1986) indicated that self-regulated learning is defined as "personalization means perception, motivated willingness, and active behaviors in the learning process". And self-regulated learning is built learning goal by students. The process of setting goal contains self-monitoring, standard setting, evaluative judgment, self-appraisal, and affective self-reaction (Schunk, 2001). In real environments, learning activities are very complex and difficult. If teachers use unsuitable teaching tools, they need to take attention to individual students and provide guidance and assistance, and it is a difficult thing. In the past, many researches showed that concept mapping was thought an effective learning strategy, and it is a visual method. Concept mapping expresses the cognitive structure between "Concept" and "Relationship". The purpose is to let students understand learning contents (Hwang, Panjaburee, Triampo, & Shih, 2013; Panjaburee, Triampo, Hwang, Chuedoung, & Triampo, 2013).

Therefore, in this study, we developed an integration of augmented reality mobile learning and self-regulated learning by using concept mapping. This system is different from traditional paper textbooks and traditional media materials, which appeared text or images only. Thus, this system use augmented reality methods to improve students' learning willing. At the same time, it uses self-regulated methods to complete students' goals. And it uses the concept mapping method, so that students may conduct reflection to finish their learning goals.

2. Development of System

2.1 System Architecture

The system architecture (includes three learning modules) which is composed of self-regulated learning, AR learning navigation, and concept mapping (see Figure 1). The self-regulated learning module function includes setting the observed goals and getting the learning feedback. When students use the AR learning navigation, they can collect data such as photos, videos and audio, so that more supplementary materials can display on the screen. The concept mapping module provides a way to improve the learning effectiveness of students. It allows students to create their own concept maps and gives them a rank with appropriate materials. The system is built by HTML5 and converted into the mobile platform through Phone Gap. All of the learning data is stored in the machine in JSON format, and upload to the server when the internet is connected.

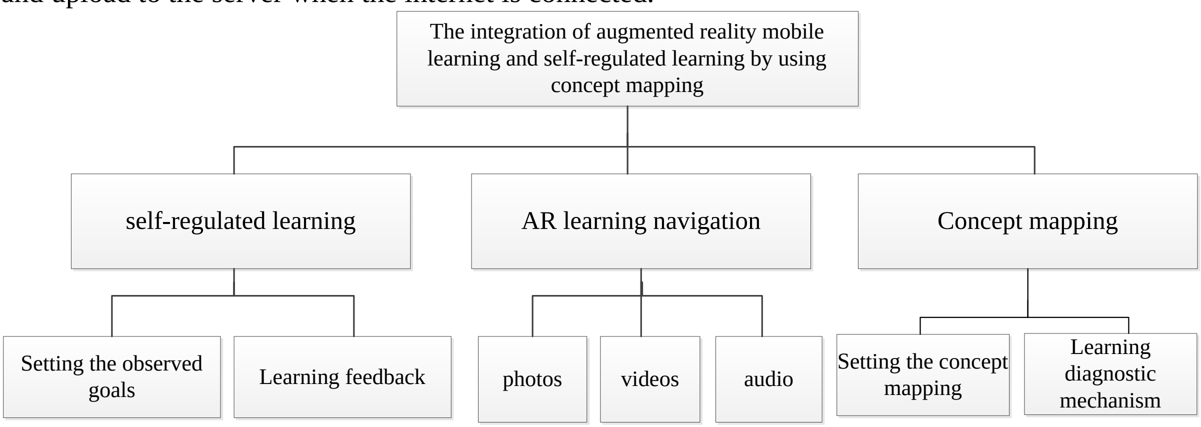


Figure 1. The system architecture.

2.2 Interfaces for Users

At first, students may login the homepage of the system by usernames and passwords. In self-regulated mode, this system provided learning goal contents for students. Students can use the dropdown menu to select their learning objects, and the system will provide the learning contents according to their learning goals. Then students can select their learning goals for the learning object. Students also can choose the confirm button to learning, as shown in Figure 2. After setting goal, the system will appear the main screen. The main screen contains four bottoms: data collection, concept mapping, operating instructions and log out, as shown in Figure 3. The four buttons are described in detail as follows.

In the AR learning navigation module, students click data collection button that they enter data collection interface according to system tasks guidance. In the AR learning navigation module, students can click data collection button, so that they can enter the data collection interface and record the data by photo, video, or audio format. According to their collection, they also can record the names, the found locations and the number of the found plants, as shown in Figure 4. In the concept mapping module, students can draw relationships by using image methods according to the learned content. In the concept mapping module, there exists a learning diagnostic mechanism. In the learning diagnostic mechanism, when students click the edit-finished button, the system will evaluate the concept maps draw by students and produce summary reports, as shown in Figure 5. The completion rates of learning goal and self-learning goal will appear in the summary report. The concept maps drawn by students will appear at the bottom of the summary report, as shown in Figure 6. When an error happened in the concept map drawn by a student, an error-list button will appear at the bottom of the summary report. In the error list, the system will provide some information for the concept map drawn by a student and correct his/her misconception. According to the unfinished learning tasks, students can adjust the goal to learning, as shown in Figure 7. In addition, when a student clicks the operating-instructions button, the system can appear operating instructions interface, as shown in Figure 8. The system can provide guide and learning-goal materials for students, and they can learn the materials via the AR technologies, as shown in Figure 9.

Figure 2. Setting goal.

Figure 3. Main screen.

Figure 4. Data collection.

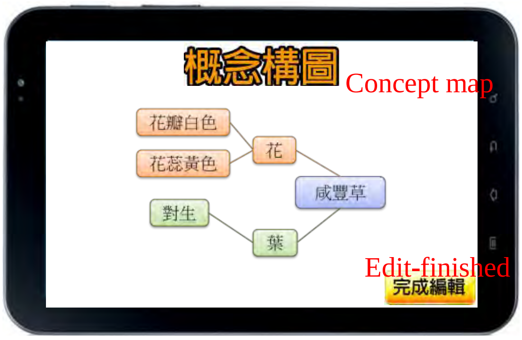


Figure 5. Concept mapping.

Figure 6. Summary reports.

Figure 7. Error messages.

Figure 8. Operating instructions.

Figure 9. AR learning materials.

3. Conclusion and Future Works

3.1 Conclusion

In this study, an integration system of augmented reality mobile learning and self-regulated learning by using concept mapping is developed. Students can use AR technologies to learn in a real environment. At the same time, students can also learn according to their own learning pace. And students can understand the learning contents and the cognitive structure between the concept and the relationship through concept mapping strategies. In the other hand, the system provides the guide and the learning diagnostic mechanisms, which can correct the misconceptions of students and provide the feedback to students. The experimental results were expected that the proposed approach is able to improve students' learning motivation and effectiveness

3.2 Future Works

In future works, we will continue the experiment of this study. We will conduct teaching experiment at a certain school in the Northern region of Taiwan. The participants are about 60 students in two classes. One class is the experimental group, which is using our system. The other class is the control group, which is using AR mobile navigation learning strategy only. All students will conduct pretest, posttest and questionnaire, which contain learning effectiveness, learning motivation, learning attitude, learning satisfaction and cognitive load. After the experiment, we will analyze whether the learning motivation and the effectiveness between the experimental group and the control group exist significant difference. In addition, we will analyze the log data of students to find what behavior may influence students' learning effectiveness, as shown in Figure 10.

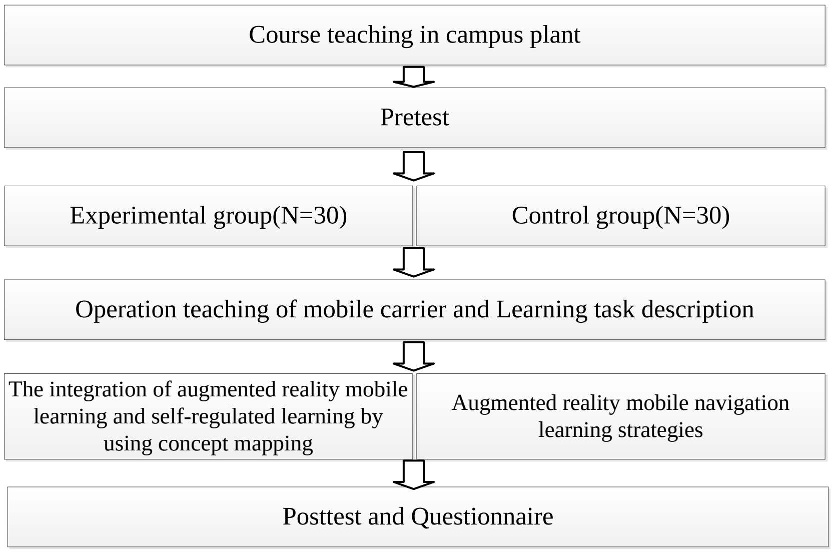


Figure 10. The flowchart of the experiment.

Acknowledgements

This study is supported in part by the Science Council of the Taiwan under Contract No. NSC 103-2511-S-152-001-MY2 and MOST 103-2511-S-275-002-MY2.

References

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6,* 355-385.

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barfield, W., Sheridan, T., Zeltzer, D., & Slater, M. (1995) *Presence and performance within virtual environments.* Virtual Environments and Advanced Interface Design. W. Barfield and T. Furness (eds). New York, NY: Oxford University Press.

Eysenck, M. W., & Keane, M. (2005). *Cognitive psychology a student's handbook.* East sussex, UK: Psycholoty Press.

Govaere Jan, L. J., de Kruif, A., & Valcke, M. (2012). Differential impact of unguided versus guided use of a multimedia introduction to equine obstetrics in veterinary education. *Computers & Education, 58*(4), 1076-1084.

Hwang, G. H., Lee, C. Y., Hwang, H. L., Huang,G. L., LIN, J. Y. & Cai, J. J. (2013). *Using augmented reality to assist an interactive multi-language learning system in an elementary school.* Workshop Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education.

Hwang, G. J., Panjaburee, P., Triampo, W., & Shih, B. Y. (2013). A group decision approach to developing concept effect models for diagnosing student learning problems in mathematics. *British Journal of Educational Technology, 44*(3), 453-468.

Hwang, G. J., Tsai, C. C., & Yang, S. J. H. (2008). Criteria, strategies and research issues of context-aware ubiquitious learning. *Educational Technology & Society, 11*(2), 81-91.

Martín-Gutiérrez, J., Luís Saorín, J., Contero, M., Alcañiz, M., Pérez-López, D. C., & Ortega, M. (2010). Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics, 34*(1), 77-91.

Panjaburee, P., Triampo, W., Hwang, G. J., Chuedoung, M., & Triampo, D. (2013). Development of a diagnostic and remedial learning system based on an enhanced concept effect model. *Innovations in Education and Teching International, 50*(1), 72-84.

Schunk, D. H. (2001). *Social cognitive theory and self-regulated learning.* In B. J. Zimmerman, & D. H. Schunk (2nd eds), Self-regulated Learning and Academic Achievement: Theoretical Perspectives (pp. 125-151). Mahwah, NJ: Lawrence Erlvaum Associates Publishers.

Uluyol, C., & Agca, R. K (2012). Integrating mobile multimedia into textbooks: 2D barcodes. *Computers & Education, 59*(4), 1192-1198.

Wernhuar, T., & Ou, K. L. (2012, March).*A study of campus butterfly ecology learning system based on augmented reality and mobile learning*. Paper presented at the 7th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technolgies on Education (WMUTE 2012), Kagawa, Japan.

Wu, P. H., Hwang, G. J., Milrad, M., Ke, H. R., & Huang, Y. M. (2012). An innovative concept map approach for improving students' learning performance with an instant feedback mechanism. British Journal of *Educational Technology, 43*(2), 217-232.

Wu, P. H., Hwang, G. J., Su, L. H. & Huang, Y. M. (2012). A context-aware mobile learning system for supporting cognitive apprenticeships in nursing skills training. *Educational Technology & Soiety*, *15*(1), 223-236.

Huang, S. H. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Application of Teams-Games-Tournament Strategy to Investigate Learning Effectiveness in Primary Schools

Shu-Hsien Huang[a], Ting-Ting Wu[b], Yueh-Min Huang[a*]
[a]*Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan*
[b]*Graduate School of Technological and Vocational Education, National Yunlin University of Science and Technology, Yunlin, Taiwan*
*huang@mail.ncku.edu.tw

Abstract: Mathematics is the foundation course in many fields, but many students are afraid to of it. Determining a method by which to promote the student interest and enthusiasm is still a problem many researchers are trying to solve. However, a lot of studies have found that a combination of mathematics and everyday life things can improve student interest, and the math learning process, combined with appropriate teaching strategies, can enhance learning achievement. Therefore, in this study, addition and subtraction, which are the basis for mathematics and a beginning course, are chosen as the learning content, and Teams-Games-Tournament (TGT) learning strategy is used in the entire learning process. It is expected that students' motivation and interest would be aroused by using game-based learning and that learning performance would be improved as a result of competition in a peer tournament. All of the participants in this study were second grade primary students who used an interactive game-based learning system constructed as a virtual grocery store and combined with TGT learning strategy for the purpose of learning addition and subtraction. The results of the experiment indicated that the students who used the proposed learning system with the TGT learning strategy had higher performance during the learning process and that interactive game-based learning can effectively promote motivation and attitude toward a math course.

Keywords: Addition and Subtraction, cooperative, Teams-Games-Tournament, game-based learning, learning performance

1. Introduction

Previous studies have indicated that motivation will enhance the enjoyment of learning (Johnson & Johnson, 1990; Jacob, 1999; Huang, Huang, & Wu, 2014) and that this in turn will improve learning achievement. The characteristics of game-based learning include interactivity, enjoyment and liveliness (Mayer, 2003; Lisi & Wolford, 2002), among others. If these characteristics along with instruction are combined and implemented into a pedagogy, this will result in a learning environment with little pressure that is enjoyable as well.

In the area of multiculturalism, cooperative learning strategy became a popular learning method starting in 1970 (Johnson & Johnson, 1990). Cooperative learning was connected with the peer cooperative relationship and sharing in order to achieve shared personal and group learning goals (Mentz, Walt, & Goosen, 2008). In cooperative learning strategies, Teams-Games-Tournaments (TGTs, as presented by David Devries and Keith Edwards in 1970, have been used to promote learning performance among peers. Slavin (1995) also indicated that TGT strategy has a structured pedagogy that is independent of the learning context, e.g. mathematical computation and applications, sentence patterns, geography and graphical skill, and science concepts.

The TGT strategy has been implemented in kingdom and at the first grade level. The results indicated that TGT positively affects mathematics learning (Jacob, 1999).

Two effects of game-based learning (GBL) are the provision of interaction and the development of intrinsic motivation. Learners either work alone or in group activities in the interactive learning

environment, and learning performance is enhanced through their operative experience and interactive communication with each other (Mayer, Mautone, & Prothero, 2002). Thus, learning achievement of game-based learning strategy comes from experience in games and is a result of immediate feedback. Learning through games initiates tournaments and cooperative as well as fun, entitlement, and overwhelming (Prensky, 2001). Moreover, interdependence and personal responsibility for individual performance exist between members in the group learning and competition process (Johnson & Johnson, 1994; Slavin, 1981). Group learning can consist of both active participation in learning activities and interaction with situations. Cognitive structures and fundamental models are constructed based on situations. In addition, learning in situations also stimulates reflection and problem-solving skills (Brown, Collins & Duguid,1989; Huang & Wu, 2010; Huang, Huang, Huang, & Lin, 2011).
When situations are integrated into in-class learning environments, interactive game-based learning not only enhances learning motivation but also helps learners immerse into the interactive learning environment. Interactive learning includes good instruction and a multimedia environment, and some design concepts, like "learning by doing," "interaction," and "cooperative." In such a learning environment, learners will find it easier to enhance both their learning performance and motivation. Instructional activities based on a game-based learning environment involve structuring the learning content and integrating the knowledge structure instead of attempting to obtain a pure recreational effect (Druckman, 1995; Eskelinen, 2001).
This study combined digital games, situational learning, and mathematics instruction as well as the design of a life-oriented virtual shop. In the game-based learning environment, learners simulate trading activities in order to learn addition and subtraction skills. Moreover, the learning environment integrates the TGT strategy into the game. Learners will use the TGT learning environment to promote constructive competition with their peers and to promote learning aspirations. The TGT strategy has funny and interactive characteristics that will attract active learning, promote learning motivation, and enhance learning performance. In this study, the experiment participates are the second grade students. This study investigates the learning effect between general game-based learning and life-oriented virtual shop game-based learning.

2. The TGT Strategy Method for Addition and Subtraction

2.1 Teams-Games-Tournament, TGT

The main characteristic of TGT strategy is a focus on team-game tournaments after the learning instruction. Before the games, learners are divided according to ability, gender, and so on. Each group includes three to six persons. All of the groups initially do interclass learning by themselves. The team game tournaments are implemented after the interclass learning. They have to invent a game to get 20 points in the tournament process. Five games lead to a total score of 100 points. Each game scores student performance in lieu of a paper-based test. In the final step of the TGT process, the total score for each group is recorded. Moreover, the teacher praises the winning groups in order to motivate the students. This process consists of both personal learning and group performance as well as increasing the interaction among peer groups. The TGT process is shown in Figure 1.

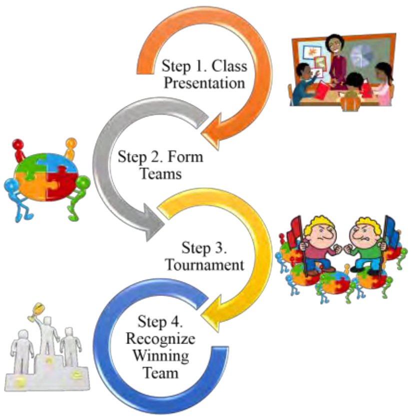


Figure 1. The TGT strategy learning process

2.2 TGT Platform- Addition and Subtraction

The aim of this study is designing a game-based TGT learning system for addition and subtraction. The learning system can assist primary students with learning addition and subtraction. Group ranking of is one function in this learning system. This function will help the tournaments held among the students promote group motivation. In this system, a virtual shop function is designed for real grocery store situational learning. Learners can use the system, which resembles a real-life experience, to learn addition and subtraction skills. Figure 2. shows the platform for the game-based TGT learning system for addition and subtraction.

a.) The virtual store interface

b.) Ranking interface

Figure 2. Illustration of the TGT learning platform

3. Experimental Environment

3.1 Participants

The participants in this study are second grade primary students. They are randomly divided into two groups. One group is the control group (n=27), and the other one is the experimental group (n=28). The control group is learning with general game-based learning. The experimental group is learning with game-based TGT learning. Both groups are learning addition and subtraction with the same teacher.

3.2 *Experimental Procedure*

This study experiment was conducted for 240 minutes in order to investigate the experimental activities with the experimental group and the control group and to evaluate their learning performance. A teacher with five years teaching experience was invited to conduct the teaching of addition and subtraction. Before the experiment, in order to determine the changes in the comprehension level in the two classes regarding addition and subtraction through the experimental design, prior to the teaching of addition and subtraction, a mathematics addition and subtraction pre-test was conducted. Before the experimental process, the two groups were instructed in mathematics addition and subtraction and in system operations. There were two classes 40 minutes long each week. In the experimental process, the experimental group used a game-based learning system with the TGT strategy. The control group used the traditional game-based learning system. At the end of the experiment activities in the last week, a post-test and attitude questionnaire related to addition and subtraction was given to the students to determine their learning performance and their attitude after implementing the experimental activities (as shown in Figure 3).

Attitude toward mathematics affects learning performance. A positive attitude toward mathematics causes students to spend more time and effort on learning mathematics as well as helping them become effective learners, and vice versa in the case of students who have negative attitudes toward math (Ma & Kishor, 1997). Aiken (1970) indicated that attitude and achievement mutually influence the learning of mathematics. Corbo (1992) also indicated that attitude toward math is significantly different in the case of learners at different levels. Some researchers have pointed out that mathematics performance is related to achievement. For the purposes of this study, an existing mathematic attitude questionnaire was modified. The questionnaire had 24 items and was scored with a five point Likert scale. This questionnaire is used to investigate the relationship between attitude toward mathematics and achievement.

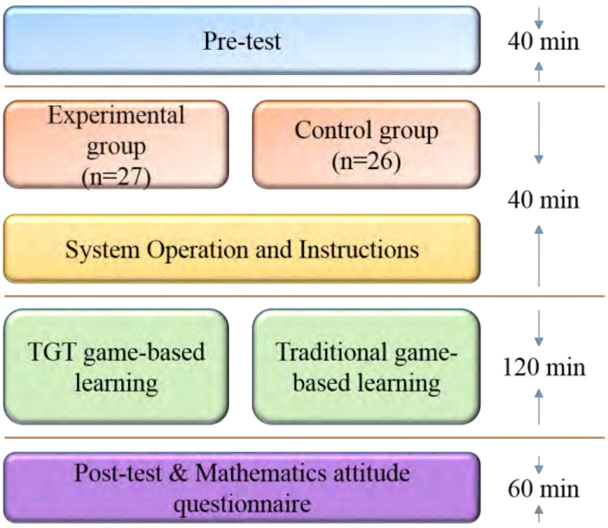


Figure 3. The experimental procedure

4. Results and Analysis

4.1 *Learning Achievement*

In order to investigate the differences between the two groups, a learning achievement pre-test was essential. The pre-test was used to ensure the ability to measure learner achievement differences. A t-test was used to evaluate the results of the pre-test. The experimental group's mean and standard deviation were 75.11 and 9.07, respectively. The control group's mean and standard deviation were 73.69 and 9.73, respectively. The variance in the t-test was .548 (p>.05). The results of the pre-test indicated that there were no significant differences between the two groups. Their learning achievement was at the same level before the learning activity. Therefore, the learning activity could be implemented.

After the learning activity, a post-test was scored and an analysis of covariance (ANCOVA) was used for analysis in order to investigate the effects of the activity on learning achievement. In the analysis process, the pre-test score was the covariance variable and the learning method was the independent variable. The analysis results are shown in Table 1. In the post test ANCOVA analysis, the mean, standard deviation and adjustment average for the experimental group were 85.33, 7.63 and 85.17,. The mean, standard deviation and adjustment average for the control group were 76.65, 8.72 and 76.81, respectively. These results indicated that there were significant differences between the experimental group and the control group (F=7.284, p<.05, as shown in Table 1). The results indicated that the experimental group and control group were significantly different in the post test. The mean for the experimental group was higher than that of the control group. This result also can be interpreted to mean that game-based TGT learning was superior to general game-based learning with regard to learning achievement. On the other hand, the standard deviation for the experimental group was lower than that of the control group. This means that the learning difference in the experimental group was smaller than that of the control group. Therefore, game-based TGT learning promoted learning achievement and lower learning differences among peers.

Table 1. The post-test ANCOVA analysis

Group	N	Mean	SD	Adj.	SE	F
Experimental	27	85.33	7.63	85.17	3.09	7.284*
Control	26	76.65	8.72	76.81	3.09	

*p<.05

4.2 Mathematics Attitude Questionnaire

From the mathematics attitude questionnaire shown in Table 2, a t-test was used to ensure differences between the two groups. The results indicated that the two groups were not significantly different. The possibly reason for this result was that both two groups used a game-based learning system and the participants were primary students. Aikden (1970) reported that mathematics attitude and achievement have a relationship at the primary school level, but the relationship is not significant. This result also echoed Wenger’s (1992) report. The results of Wenger’s study indicated that learners with high levels of positive attitude toward mathematics have better ability than those with low levels of positive attitude. The results also indicated that mathematics attitude was not positively related to mathematics achievement. Therefore, both groups’ attitudes toward mathematics were not significantly different.

Table 2. Descriptive statistics for mathematics attitude questionnaire

Group	N	Mean	SD	*t*
Experimental	27	4.18	3.21	1.72
Control	26	3.97	4.17	

5. Conclusions

This study investigated game-based TGT learning of addition and subtraction. The learning environment was a combination of game-based learning, team tournaments and virtual shop trading. The teams invented games in the learning environment. Mutual survival and personal responsibility were learned by inter-group relations. The results indicated that TGT game-based learning achievement was higher than general game-based learning. Thus, constructive competition promotes peer learning and learning achievement. Hence, the two groups were not significant different in terms of attitude toward mathematics. Related work is in progress, and qualitative research will be used to collect interview data. Participant information (e.g. learning motivation, math anxiety, etc.) will be collected in the future in order to investigate the mathematics learning status of students.

Acknowledgements

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 102-2511-S-041-003 and NSC 101-3113-P-006 023, as well as Ministry of Science and Technology (MOST), Taiwan, ROC, under Grant MOST 103-2511-S-224-003, and MOST 103-2511-S-006-002-MY3.

References

Aiken, L. R., Jr. (1970). Attitudes toward mathematics. *Review of Educational Research, 40*(4), 551-596.

ASCD（1990）.Cooperative Learning series: Facilitator's Manual. Association for Supervision and Curriculum Development.

Brown, J. S., Collins, A., ＆Duguid, P. （1989）. Situated cognition and the culture of learning. *Educational Researcher, 18(1)*, 32-42.

Corbo, N. J. (1992). *Mathematics attitude and achievement in grades five through seven in a south-central Pennsylvania district*. Mich: UMI.

Druckman, D.（1995）.*Situational levers of position change: further explorations*. The Annals of the American Academy of Political and Social Science, 542, 61-80.

Eggen, P. D. & Kauchak, D. P.（2001）. Strategies for Teachers. Allyn & Bacon

Eskelinen, M.（2001）.The game situation. *The International Journal of Computer Game Research, 1*(1),152-171.

Gunter, M. A.、Estes T. H. & Schwab J. H. （1990）：Instruction: a model approach. Allyn & Bacon

Huang, Y.M., Huang, S. H., & Wu, T. T. （2014）. Embedding diagnostic mechanisms in a digital game for learning mathematics. *Educational Technology Research and Development*, 62(2), 187-207.

Huang, Y.M., Kuo, Y.H., Lin, Y.T., & Cheng, S.C. （2008）. Toward interactive mobile synchronous learning environment with context-awareness service. *Computers & Education, 51*, 1205–1226.

Huang, Y. M., Huang, Y. M., Huang, S. H., & Lin, Y. T. （2011）. A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education, 58*, 273–282.

Lisi, R. and Wolford, J. L.（2002）. Improving Children's Mental Rotation Accuracy with Computer Game Playing. *The Journal of Genetic Psychology, 163*, 272-282.

Jacob, E. （1999）. Cooperative Learning in Context: An Education Innovation in Everyday Classroom. *SUNY Series' the Social Context of Education*, ED 431773.

Johnson, D. W., & Johnson, R. T.（1990）. Social skills for successful group work. *Educational Leadership, 47*(4), 29-33.

Johnson, D., & Johnson, R.（1994）. *Learning together and alone: Cooperative, com-Petitive, and individualistic learning* (4th ed.). Boston: Allyn & Bacon.

Ma, X, & Kishor, N. (1997). Attitude toward self, social factors, and achievement in mathematics: A meta-analytic review. *Educational Psychology Review, 9*(2), 89-120.

Mayer, R. E., Mautone, P., & Prothero, W. （2002）. Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Education Psychology, 94*, 171-185.

Mayer, R. E.（2003）. The Promise of Multimedia Learning: Using the Same Instructional. *Campus-Wide information System, 13*, 125-139.

Mentz, E., Van Der Walt, J. L., & Goosen, L. （2008）. The effect of incorporating cooperative learning principles in pair programming for student teachers. *Computer science education, 18*(4), 247-260.

Prensky, M. （2001）. Digital Game-Based Learning. McGraw-Hill.

Slavin, R. E.（1995）. *Cooperative Learning: Theory, Research, and Practice* (2nd ed.) Massachusetts: Allyn & Bacon.

Slavin, R. E. （1981）. Synthesis of research on cooperative learning. *Educatioal leadership, 38*, 655-658.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Education 3.0 and beyond: A learner-led experience of Education

Pranav KOTHARI[a*], Anurima CHATTERJEE[b]
[a]*Educational Initiatives Pvt. Ltd., India*
[b]*Educational Initiatives Pvt. Ltd., India*
[*]pranav.kothari@ei-india.com

Abstract: The last decade has seen technology entering the education space in multiple ways – through digitised learning content, adaptive online tests, pre-recorded or real-time videos of teachers, social networking, MOOCs, etc all that have been classified as Education 1.0, 2.0 and 3.0. The predominant reference of these terms has been for higher education (college level and beyond). It assumes that the student is a mature and motivated adult who is able to take in discrete pieces of information and assimiliate it coherently to learn. This paper explores how Education 3.0 concepts are implemented in the realm of elementary education with a focus towards increasing students' learning outcomes. This paper combines our research in heutagogy by drawing on data from an intelligent tutoring system called Mindspark that has been used by over 80,000 students in 150 private schools over the past five years as well as in 5 after-school remedial centres in urban slums of Delhi. There is a shift in educational processes - characterized by the transformation from a teacher-led delivery model to a student-led learning experience. In this Education 3.0 model, the learner is no longer a passive recipient but rather becomes an active player in defining their learning environment, the teacher-learner interaction, affective factors of learning, evaluation, interventions and finally achievement – all towards achieving higher outcomes on students' learning.

Keywords: Education 2.0, Education 3.0, intelligent tutoring system, elementary schooling, technology in education

1. Education 3.0 – Newer trends in elementary education

Education technology has seen varied changes resonating users' social experience and interaction with the Web. In line with the various versions of the Web there is an evolution of educational processes as Education 1.0, Education 2.0, Education 3.0 and finally Education 4.0. This evolution draws in from the ideas like progressivism and social constructivism. The ideas of progressivism in education (as purported by thinkers like John Dewey) talk about how "learning by doing" allows for a more active environment for learning where the learner is able to develop problem-solving and decision-making skills. Social constructivism recognises that culture plays a large role in cognitive processes and learning (Vygotsky) and student dialogue and discussion and collaborative learning processes are emphasised. Technology has been a catalyst for these ideas in education, wherein content can be made more meaningful and interactive for the learner and executed even in operational constraints like low availability (in quantum or quality) of teachers and teaching resources.

Education 3.0 is a shift in user-driven motivation, teacher facilitation and in pedagogic styles. According to Jackie Gerstein, use patterns of students should drive the type of mobile learning activities so that the transfer outside the learning setting can occur. The role of the educator in this new environment transforms from being gatekeepers of knowledge to becoming model learners who can demonstrate self-directed learning (Gerstein). While Education 3.0 and 4.0 suggests ubiquitous learning is possible with freely available resources and user-generated content and processes, one needs to interrogate how effective this process is for elementary school students. Can such large amounts of learning content be moderated intelligently by students? Can a 5th grader be expected to read large volumes of data on historic wars and discern reliable sources from non-reliable sources? Can a 2nd grader have the requisite knowledge to decide what she/he needs to learn after finishing addition of fractions? To make it effective, Education 3.0 for elementary education will need to adapt to some of the things that are taken for granted in adult education in Education 3.0.

In this light, we will look at how the learning process in elementary education is aligned to the principles of Education 3.0 in three aspects. The first aspect of analysis will be the learning environment which will describe the online interface with its various features for students to engage with their teaching learning content, understanding their own learning achievement and progress, communicate their emotional states on specific content, receive rewards based on achievement for motivation to continue further and many others. The second aspect studies the role of the teacher in this new learning environment. The third aspect is the dynamic nature of the learning content that is served in real time driven by user response. This section will describe the logic of learning content sequencing through data driven misconception identification and remediation, recognition of prior learning, variable learning path and students' learning while being in flow. Table 1 below is a brief understanding of the shifts in the various generations of education as explained by John Moravec and Arthur M. Harkins (Moravec, 2008 and Harkins 2008) and the next section will describe how the above concepts of Education 3.0 have been implemented utilizing Mindspark.

Table 1: The various generations of educational technology

	"Download" Education 1.0	**"Open Access" Education 2.0**	**"Knowledge Producing" Education 3.0**	**"Innovation Producing" Education 4.0**
Meaning is..	Dictated	Socially constructed, with aid of Internet access	Socially constructed and contextually reinvented knowledge	Build through selective individual and team-driven embodiments in practice
Technology is ...	Confiscated at the classroom door	Cautiously adopted open access	Everywhere	Always changing with learners as a source of innovation production
Teaching is done...	Teacher to student	Teacher to student and student to student	Teacher to student, student to student, student to teacher, people-technology-people	Amplified by positive innovation feedback loops; ubiquitously and creatively 24/7 in all phases of living, learning and working
Schools are located...	In a building (brick)	In a building or online but increasingly on the Web	Everywhere in the "creative society"	In the globally networked human body, a continuously evolving instrument innovatively supplementing and replacing classrooms
Parents view schools as...	Daycare	Daycare with a laboratory	Places for students to create knowledge	Schools are viewed as one of many innovation venues
Teachers are...	Licensed professionals	Licensed professionals who team with students and parents	Everybody, everywhere backed by wireless devices for knowledge production	Everybody, everywhere is an innovation production source backed up intuitive software "partners" and human collaborators.

2. Mindspark – Intelligent Tutoring System

Educational Initiatives, India developed an intelligent tutoring system (ITS) called Mindspark in 2008 which is used for teaching Mathematics and Language to elementary school students. In 2014, the Mathematics program is used by over 80,000 children in private and government schools and the

Language program is used by 5,000 children in government schools and after school remedial centers. An Intelligent Tutoring System (ITS) is a computer based tutoring program that provides personalized learning content to students based on factors like student performance and prior knowledge (Corbett, Koedinger and Anderson, 1997). The sequencing of learning content is done in a manner to avoid cognitive mismatch such as cognitive overload for low performers and boredom for high performers (Brusilovsky and Milln, 2007) to ensure that child is engaged and is performing at the optimum level (Csikszentmihalyi, 1998) . In Mindspark, the questions are 'finely-graded', meaning that there are a very large number of questions of gradually increasing levels of difficulty. In a way, Mindspark adapts to the needs of every individual student. Questions are specially designed to test understanding and to help students clear their misconceptions. When a student answers a question or combination of questions incorrectly, the intelligent system diagnoses the child's misconceptions / weak areas. The child may be further provided with a simple or detailed explanation, or be redirected to questions that strengthen the basic understanding. These decisions are taken by an adaptive logic which is expected to get better and better with increased student usage (Rajendran and Muralidharan, 2013). There is very little emphasis on instruction due to the belief that students learn when they have to think – either by answering a question, or by doing an activity on the computer (Prince, 2004).

An independent third party evaluation of Mindspark shows an effect size of 0.40 against a comparison group with usage as low as 1 hour per week (50 hrs per year). This study finds a favorable outcome despite the teacher effects and school effects that may occur. This study was done across more than 4,400 students from 18 different schools that used Mindspark comparing with more than 15,000 students across 150 schools that did not (IDInsight, 2014).

3. Technology in Education 3.0 is…

3.1 Providing student intelligent and limited choice

The Mindspark portal hosts the interface for every student, where the content is available to students depending on their learning levels. This becomes a dynamic classroom for every child where they have a certain choice of topics to choose from based on extent of completion, performance on the pre-requisite topics, teacher intervention and the adaptive algorithm. Intentionally the entire content is not left open at the discretion of the student. These are shown on the left side in Figure 1. This session is bookmarked to the place where the student left the previous session and allows the student to continue from the same place.

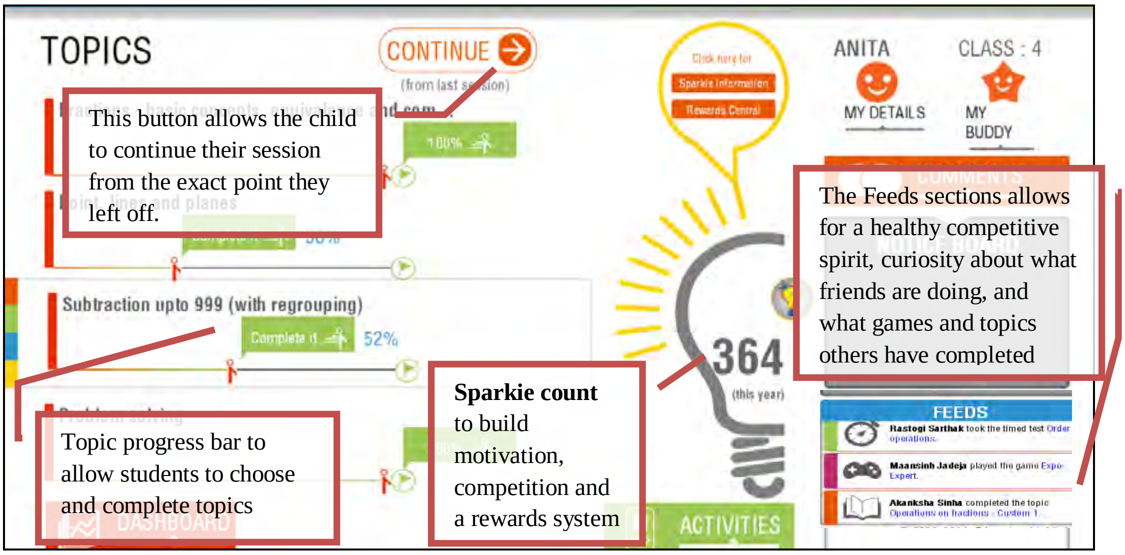


Figure 1. Dashboard for Mindspark

3.2. Providing the encouragement and incentives customized to each child's needs

While technology has allowed for systematic archiving of content and assessing student performance, it is often felt that affective factors like boredom, motivation, and sensing emotional states of the learner are outside the purview of technology. Although intrinsic motivation to learn often helps students approach technology based learning tools, at times, they need a little push to persist at it, with minimal human interaction. Skinner's work on rewards and reinforcements (Skinner, 1938) has been applied extensively by teachers and educationists in working with student behaviour and motivation. These are replicated by Mindspark through creating extrinsic rewards called "Sparkies" to promote positive learning habits like grit and persistence required to help students learn and progress from the mistakes they might make; and to cope with the increasing difficulty levels of questions. Sparkies are shown by the number 364 in Figure 1 above.

3.3. Capturing students' emotions

In addition to this, the tutoring system tries to gauge the student's emotional state through an Emote Toolbar shown on the right side of Figure 2 (to share feedback on moods and emotions related to content). This also provides data to analyze questions that have been repeatedly marked by students as "Boring", "Exciting" or "Confusing". While this data is currently used to review and modify the learning content and its sequencing for greater student engagement, research on affective states of learners indicates that there in future there is a possibility to predict frustration and also address it in real-time through the program's learning environment (Rajendran, Iyer, Murthy, Wilson and Sheard, 2013).

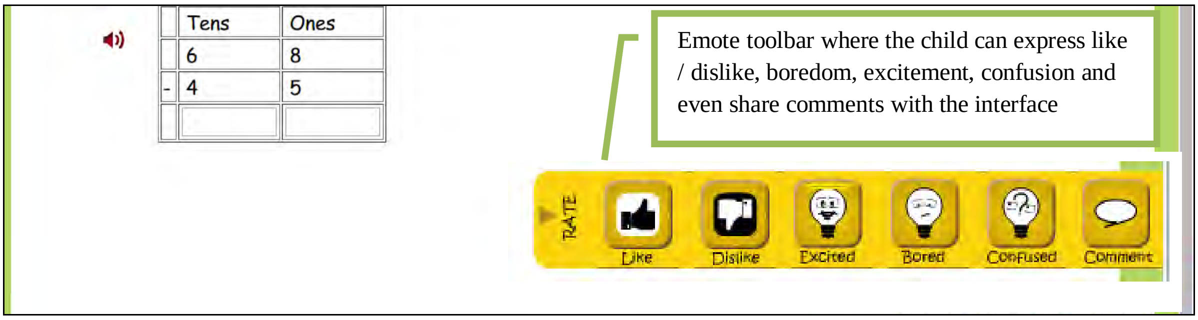


Figure 2. Emote toolbar to capture levels of interest, boredom, frustration etc.

3.4. Providing students transparency and visibility into their learning

Mindspark helps children keep track of their goals through a tabulated Cluster chart shown in Figure 3 to track their levels of completion as well as success. Mindspark believes that keeping students well-informed about their learning process will help them be more aware of themselves as learners and therefore, understanding the value of what is being presented, do better. This is why students are also given prompts telling them when they have finished a learning unit successfully and when they are being made to repeat a learning unit or attempt more basic learning units, as is the case when they get a number of questions incorrect.

Add Sub up to 999 - Pre Math module	Arranging and adding any whole numbers	Knowing facts for division by single digit number	Dividing a 2- digit number by 1 digit number	Division - Pre math module	Problems based on multiplication and division
Dividing a 2-digit number by 1 digit number standard	Dividing a number by multiples of 10	Multiplying two 2-digit numbers vertically	Multiplying a 3-digit number by 1, 2 digit number	Problems involving more than one operation	Estimating products
Introduction to basic fractions (half and quarter)	Introduction to fractions	Basic understanding of half and quarter	Naming fractions and understanding fraction notation	Understanding fractions as part of a collection	Identifying right angles
Save TOTO game	Measurement of length -interactive	Using standard units of length (m and cm)	Using standard units of length (km and m)	Concept and computation of perimeter	Informal understanding of area

Figure 3. Cluster charts that show extent of completion and accuracy

3.5. Providing encouragement to those who have done well, and support to those who need it

Encouragement is provided to students when they achieve certain milestones (Figure 4: "Great going, Anita") so that their energy and enthusiasm is renewed for the next topic. In addition to this, the reward system (Figure 5) applauds diverse qualities like being consistent and perseverant, academic achievement etc. There is a monthly competition among students to become champions in their classroom based on the speed, accuracy, complexity of questions as well as reading explanations to encourage higher usage.

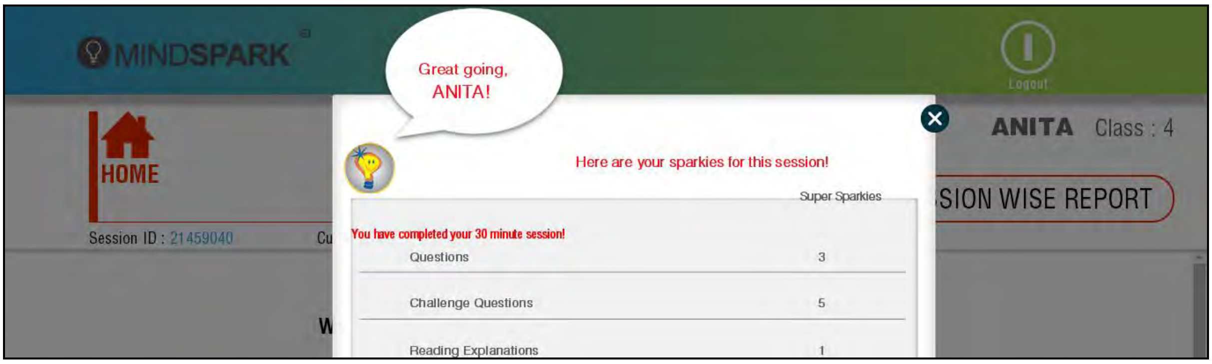


Figure 4. The student interface showing various interactive elements like Sparkies (rewards), encouragement statements, and emotional sensors

Figure 5. Reward system at the classroom level

3.6. Personalization based on the student, time of year, geographic location, festivals, etc

Mindspark allows for personalization of celebrations like an animated "Happy Birthday" greeting card on the student's birthday that opens up on the first login of that day. It also allows for localization of content through interface themes related to upcoming festivals as shown below on the left hand side of the login page during an Indian festival (Figure 6). This space is also used to profile accomplished mathematicians thereby providing the inspiration to become one.

Figure 6. Mindspark login page showing public announcements for Sparkie Champs and greetings

3.7. Providing specific step by step intelligent responses

Technology in Education 3.0 is used more for ensuring that children are learning by providing intelligent responses instead of simply marking students right or wrong (which can be discouraging). Here feedback provided is customized to the type of error that a student is making as shown in an example leading to solving a linear equation with popular mistakes that students make.

Solve for x

$3(x+7)=15$

Hide	$3x+7$	=	15	You seem to have made an error in expanding the term: $+3(x+7)$
Hide	$3x$	=	15	Please check the term on left side of the equal to sign.
Step 1	$3x+21$	=	15	✓
Hide	$3x$	=	$15+21$	You made a mistake. To move $+21$ from the left side to the right side you need to change the sign of the term.
Hide	$3x$	=	36	You made a mistake. You seem to have made a mistake in moving terms to isolate terms with x on the left side. You made a mistake while adding together the terms $+15$ and -21 on the right side. Please try again.
Step 2	$3x$	=	$15-21$	✓
Step 3	$3x$	=	-6	✓
Hide	x	=	-6	To isolate x you need to divide by 3 on both sides.
Step 4	x	=	-2	✓

Figure 7. Intelligent responses to errors made by students while solving linear equations

4. In Education 3.0 Teaching is done…

4.1 By reading accurate and instantaneously generated reports of classrooms

Teachers are able to gauge performance, students' needing attention, level of comprehension through system generated reports to aid in teaching.

Students	Progress	Total Q's ↓	% Correct	Total Attempts	Learning units not cleared	Trail
Aswin Benedict	35.9%	273	48.7	1	Representing numbers up to 9999 using base ten blocks and abacus	Trail
Pradeep K	↓ 33.6%	234	42.3	1		Trail
Susmitha V	60.3%	188	46.8	1		Trail
Santhosh M	↓R 100%	183	80.3	2		Trail

Figure 8. Class level reports showing topic progress across various students

4.2 Using real time data for effective monitoring

In this model, the teacher is provided by system driven prompts to help facilitate their teaching in a Mindspark class. Through the dashboard (shown below in Figure 9), a teacher is able to ascertain the child's comfort with the learning content and provide support to students struggling with concepts. This is done by a combination of the pace at which questions are done (system flags if some student is really slow), the accuracy with which it is done and shows relevant parameters (such as class grade level vs. actual level, etc)

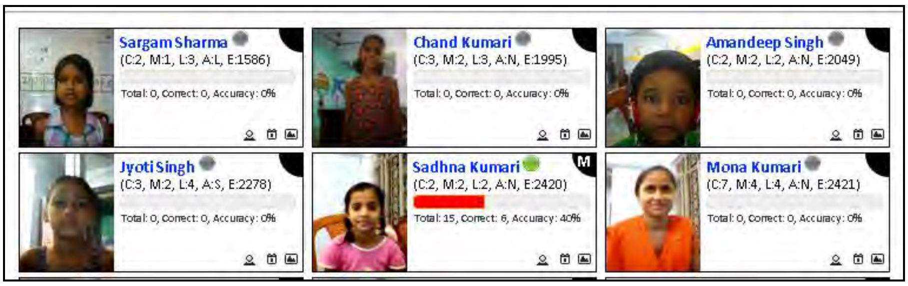


Figure 9. Teacher dashboard for monitoring students' performance

4.3 Better communication with parents

Through system generated reports, the teacher is able to accurately gauge the learning levels of each child and communicate the learning gains and the challenges with parents and to set goals for their students for subsequent months. Also through the teacher interface, the teacher is directly able to email students' parents who can also monitor the progress via the parent portal.

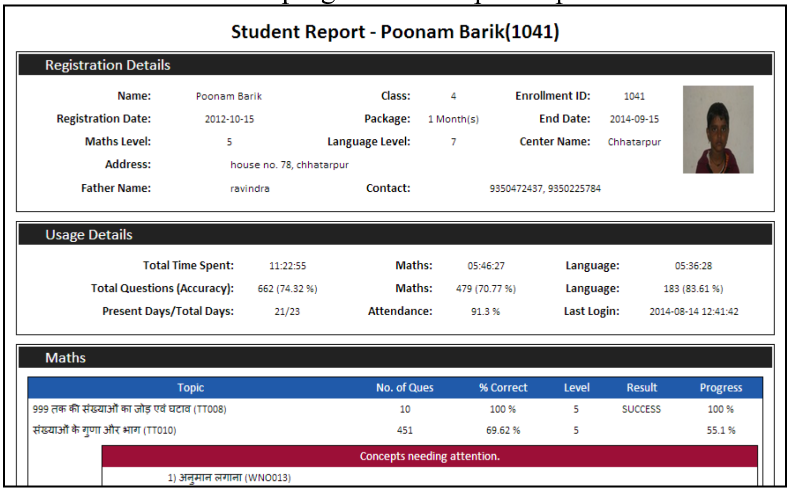


Figure 10. Monthly report on student usage and accuracy shared with parents

4.4 Data driven instruction

Through data on student performance across levels, topics and questions, teachers are able to identify learning misconceptions and plan their offline instruction to help kids through these. For eg. in image below the question shows the "Most common wrong answer" that children have picked as a result of a popular misconception as well as kids who never got the concept right even after being given multiple times. This allows the human teacher to intervene and teach the particular concept. Also the Common Wrong Answer report can be downloaded as a worksheet – printed copies of which can be given to children for practice.

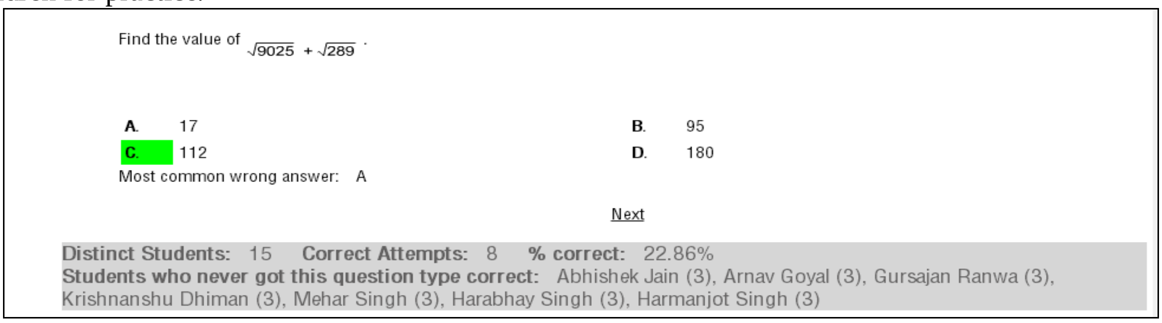


Figure 11. Common Wrong Answer reports

4.5 Teacher-led customization of learning content

While Mindspark provides the learning path for every child, flexibility and autonomy is given to the teachers to allow students to focus on a particular concept determined important by teacher or school. This can be done at the individual student level or at the entire class level.

Figure 12. Topic customization by teacher – for individual level or class level

5. In education 3.0 the learning content is…

5.1 Integration of online content with school curriculum

Teachers using this program integrate Mindspark into their curriculum plan whereby they 'activate' topics for students based on content taught in class and use the performance results of students to identify degree of comprehension, common misconceptions, and low performing students. It thus aims to use not just the interactivity of the computer, but its intelligence; and to mimic the diagnostic capabilities of a good teacher. In addition it also serves as a powerful teaching tool in these cases allowing teachers exposure to good learning materials.

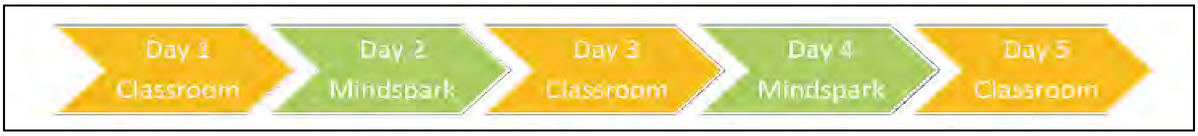


Figure 13. Integration of Mindspark in the school curricula

5.2 Informed through data on student performance and misconceptions

Similar to the discussion on Common Wrong Answer reports, the program captures data on student performance across topics to identify misconceptions. The learning content recognises these misconceptions and attempt to address these through its content flow. Examples of student misconceptions diagnosed and addressed by Mindspark are

- 20 = 5x4 is considered wrong by many students (most students think that 5x4=20 is the only correct notation)
- Confusion between terms like k + 3, 3k, k^3, etc

5.3 Appropriate to the learner's abilities and learning levels

The learner's journey begins with a baseline diagnostic test that assesses the actual understanding level (i.e. "current level") of each child for their Language and Mathematical competencies. In cases where the current level is lower than the grade level of school, the program adapts to the child's learning level and through performance at each learning unit, the child progresses to the next unit or moves to a lower level learning unit or a remediation plan. In Figure 11, this child is successfully completing learning units (downwards) but needs to move to previous units at various junctures.

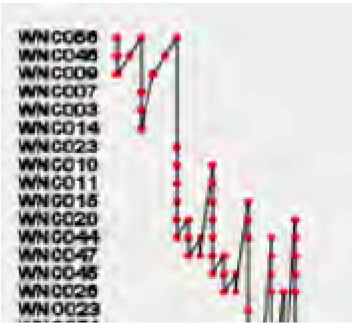


Figure 11. Learning path of learner

6. Concluding remarks

The best use of technology in elementary eduction is not for putting up fancy smartboards and projectors in classroom which typically deal with the delivery of material to students; but instead employing it towards a personalized and adaptive learning program with the right support system. Student attendance tracking, greater tracking of attendance and accountability of teachers by monitoring their login behaviour, involving parents in student behaviour through system-driven phone calls or texts, measuring metrics like retention, usage, and performance in real-time are some ways that would allow for greater facilitation and efficiency of learning for children under the Education 3.0 generation.

Applying the principles of Education 3.0 in the context of elementary education is a complex one due to the nature of the target group. Catering to learners of different learning abilities, varying levels of reading, and requiring diverse skill sets to achieve the goals of the curriculum requires us to think beyond the framework of Education 3.0. While knowledge must be freely accessed and produced by the learner, the learner can't be left to moderate that knowledge on their own. Hence, the learning platform needs to be responsive to the access and creation of that knowledge to provide appropriate learning support and avoid a 'knowledge deluge'. In addition to this, one can't ignore the importance of a human intervention required with young learners. While advocating for technology-led learning, the changing role of the teacher and parent must be understood to allow for effective learning for every child. Through adaptive learning platforms like Mindspark, it is possible to provide young learners with the new learning experience synonymous with Education 3.0 that creates the incentives for a child to learn and for teachers to monitor and ensure that each child is performing at his/her best potential.

Acknowledgements

The authors would like to thank Central Square Foundation and Tech Mahindra Foundation for their financial support and guidance. We are very grateful to Sridhar Rajagopalan, Ritika Datta and Ritika Subhash for reviewing this paper and providing suggestions.

References

Albert T Corbett, Kenneth R Koedinger, and John R Anderson. (1997) Intelligent Tutoring Systems. Handbook of Human computer interaction, pages 849–874

Brusilovsky and E. Milln. (2007) User models for adaptive hypermedia and adaptive educational systems. In The adaptive web: methods and strategies of web personalization. Springer-Verlag. pages 3–53.

Csikszentmihalyi, Mihaly (1998). Finding Flow: The Psychology of Engagement With Everyday Life. Basic Books.

Gerstein, Jackie. User Generated Knowledge. https://usergeneratededucation.wordpress.com/

Govinda, R. & Bandyopadhyay, M. (2011) Overcoming exclusion through quality schooling. CREATE-PTA. Research Monograph No. 65. University of Sussex, UK

IDinsight. (2014) Impact Evaluation of Mindspark. http://www.ei-india.com/wp-content/uploads/2011/11/IDinsight-EI-Mindspark-evaluation.pdf

Harkins, A. M. (2008). Leapfrog Principles and Practices: Core Components of Education 3.0 and 4.0. Futures Research Quarterly, 24(1), 19-31.

Michael Prince (2004) Does Active Learning Work? A Review of the Research, Journal of Engineering Education Volume 93, Issue 3 pages 223–231

Moravec, J. W. (2008). Moving beyond Education 2.0. Retrieved March 18, 2008, from Education Futures Web site: http://www.educationfutures.com/2008/02/15/moving-beyond-education-20/

Rajendran, R. Muralidharan, A. (2013) Impact of Mindspark's Adaptive Logic on Student Learning. T4E

Rajendran, R.; Iyer, S.; Murthy, S.; Wilson, C.; Sheard, J. (2013) "A Theory-Driven Approach to Predict Frustration in an ITS," Learning Technologies, IEEE Transactions on , vol.6, no.4, pp.378,388, Oct.-Dec

Skinner, B.F. (1938) The Behavior of Organisms: An Experimental Analysis

The Probe Team (1999) Public Report on Basic Education in India. Oxford University Press. New Delhi

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

The Impact of Affective Tutoring System and Information Literacy on Elementary School Students' Cognitive Load and Learning Outcomes

Ching-Ju CHAO[ab*], Shang-Chin TSAI[ac], Chia-Hsun LEE[a], Tao-Hua WANG[d], Hao-Chiang Koong LIN[a]

[a]*Dept. of Information and Learning Technology, National University of Tainan, Taiwan*
[b]*Dept. of Applied Foreign Languages, Tung Fang Design Institute, Taiwan*
[c]*Dept. of Motion Picture, Tung Fang Design Institute, Taiwan*
[d]*Degree Program of Fashion and Performance, Kun Shan University, Taiwan*
*chingju@mail.tf.edu.tw

Abstract: In this study, an affective tutoring system (ATS) is developed by integrating dengue fever prevention courses with an intelligent tutoring system (ITS) under a semantic identification module for a teaching experiment. A one-group pretest-posttest research design is adopted for this teaching experiment, which involves 66 research participants and lasts for two weeks. Further, students are divided into a high score group and a low score group based on students' information literacy and cognitive load, and nine students are selected to participate in a focus group interview. Our research team uses a "primary school information literacy questionnaire" and a "primary school student's cognitive load scale", qualitative research methods, and auxiliary quantitative research methods to cross-verify the effect of the dengue fever prevention affective tutoring system on the cognitive load and learning outcomes of primary school students with different level of information literacy.

According to results of this study, both qualitative and quantitative analysis endorse the hypothesis that the dengue fever prevention affective tutoring system enhances students' learning outcomes. Students with varied information literacy show significantly different cognitive load, yet there is no significant difference in students' overall cognitive load, content of cognitive load, and learning outcomes. Quantitative analysis shows partially significantly negative correlation between students' information literacy and cognitive load, yet the correlation between students' information literacy, cognitive load, and learning outcomes is not endorsed by quantitative analysis.

Keywords: Affection Tutoring System, Intelligence Tutoring Systems, dengue fever, information literacy, cognitive load

1. Introduction

Dengue fever, which is rampant in southern Taiwan every summer, is an acute infection transmitted by the bite of a mosquito infected with dengue viruses. Due to the general public's inadequate knowledge of mosquito species that transmits dengue fever, the time when these mosquitos are active, and how dengue transmission happens, people tend to pay little attention to environmental hygiene and end up aggravating the severity of dengue fever outbreaks in metropolitan areas. To reinforce the promulgation of dengue fever prevention knowledge, the government is devoted to advocating the importance of dengue fever prevention in schools at all levels. In this study, our research team observed that in the course of receiving traditional indoctrination, students may be subject to various influencing factors, develop negative emotions in the process of learning, and end up having learning outcomes that are below expectation.

Given that a growing number of recent studies have pointed out the important impact of emotions on learning, we suggest integrating an affective tutoring system (ATS) with courses of dengue fever prevention, which involves discerning students' emotional state in the process of interacting with students, giving timely feedback, and adjusting the pace of teaching. Such teaching mode can only be performed through students' human-computer interaction. At present, students are accustomed to being surrounded by abundant information. Nevertheless, whether students are equipped with better information literacy to minimize cognitive load caused by an overall curriculum in the process of learning, sources of such cognitive load, along with the interplay between students' cognitive load and learning outcomes are still what researchers are curious about. As such, with the aid of the innovative teaching mode, we propose to examine the interplay of cognitive load and learning outcomes of students with dissimilar information literacy in students' learning process. Through this study, we also aspire to suggest more diverse promotion approaches of dengue fever prevention to teachers in the field of education.

2. Literature Review

2.1 ITS & ATS

An intelligent tutoring system (ITS) is a system which provides students with personalized guidance or direct feedback through computer analysis (Sarrafzadeh, 2002). Knowledge of an intelligent tutoring system is constructed by three modules: a student module made up of students' knowledge, a tutoring module made up of teaching-related knowledge, and an expert module made up of knowledge in specific fields. Further, knowledge of an intelligent tutoring system and the user interface module are integrated to make a comprehensive intelligent tutoring system structure (Koedinger& Corbett, 2006). Also, an intelligent tutoring system, which determines the content and methods of teaching based on individual students' characteristics, is like a real tutor. As such, an intelligent tutoring system enjoys more advantages as it provides useful, uncritical, and tailor-made feedback (Anderson, Corbett, Koedinger, & Pelletier, 1995; Johnson, Rickel & Lester, 2000).

Affective tutoring systems, which are developed from intelligent tutoring systems, are expected to improve intelligent tutoring system and further make intelligent tutoring systems like a real tutor, adapt to learners' emotional state, and help learners learn effectively (Sarrafzadeh et. al., 2003; Alexander, Sarrafzadeh & Hill, 2006). A tutoring research group at the University of Memphis attempted to add emotional components into the AutoTutor of an intelligent tutoring system and used this system to conduct a test on approximately 1,000 learners with physics or computer background knowledge. The result was that learners had achieved remarkable learning outcomes in both surface-level knowledge and deeper learning (Craig, Graesser, Sullins & Gholson, 2004). Moreover, in the course of real teaching and teaching through an intelligent tutoring system, learners' common emotions such as joy, frustration, surprise, curiosity, wonder, etc. had stronger correlation with and made great impact on learners' learning experiences (Burleson & Picard, 2004; Craig et al., 2004; D'Mello et al., 2008; Graesser et al., 2006). On the other hand, Picard's research put forward a conceptual module which could affect learners' emotions while learning, identify learners' emotional state, give timely feedback, and improve learners' learning (Picard, Kort & Reilly, 2001).

2.2 Identifying semantics and emotions in Chinese language

Through natural language processing and semantic analysis, we could effectively understand the semantic content, acquire correct information from the semantic content, and assist to identify emotions (Yan, Bracewell, Ren& Kuroiwa, 2008; Quan & Ren,2009). This is because subsequent actions such as emotion recognition and feedback could only be performed after the acquisition of correct information. However, in practice, semantic and emotion recognition methods would vary according to a sentence or an article's hierarchical structure (Kao et.al,2009; Quan &Ren,2010; Xu, Meng &Wang,2010).

Calix(2012) extracted descriptive texts from fairy tales and incorporated these texts into the Support Vector Machines (SVM) training corpus, determined learners' emotions based on learners' textual inputs of learners after training, and let a specially-designed 3D character model show corresponding emotional responses based on learners' emotions. In 2009, Goh & Huang mentioned about using text mining algorithms to search for wordings which convey negative emotions on blogs as a way to prevent teenage depression and suicide incidents.

Sun, Chen, Liu, Liu & Soo(2010) used the Chinese word segmentation system developed by Academia Sinica Institute of Information Science to segment words and phrases. Also, the popular social networking site “Plurk” was used as the source where words and phrases were gathered from. Further, the collected short Chinese phrases were categorized based on these phrases’ emotions and the characteristic lexicons were converted to vectors and joined with a probability model, which was from the semantic dictionary developed by the Natural Language Processing Laboratory at National Taiwan University, to form a hybrid model to analyze emotions of short Chinese phrases. As indicated by studies in recent years, results of research on Chinese semantic identification technology have been good. Marvelous progress in Chinese semantic identification technology has strengthened its potential of enhancing the effectiveness of emotion recognition research.

3. Research Method

3.1 Research Architecture

With the one-group pretest-posttest design as the research design and the research framework in Figure 1, this study proposes the following three hypotheses:

H1: Teaching by means of an affective tutoring system can enhance students’ learning outcomes.

H2: Students with varied information literacy would experience significantly different cognitive load while receiving teaching through an affective tutoring system.

H3: Students with varied information literacy would have significantly different learning outcomes after receiving teaching through an affective tutoring system.

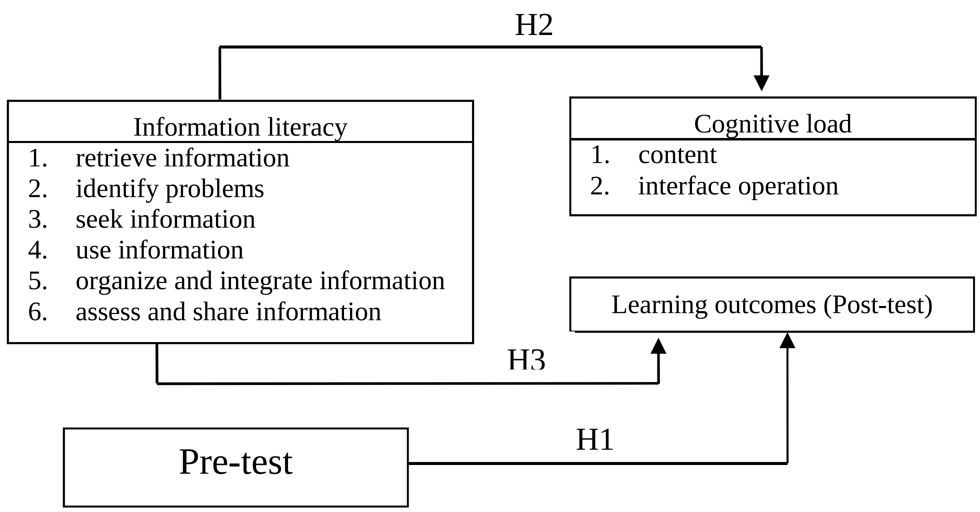


Figure 1 Research Architecture

3.2 Participants and Instrument

77 students in grades 5-6 from a public primary school in Kaohsiung were chosen as research participants in this study. 77 questionnaire copies were distributed and 70 questionnaire copies were collected, a response rate of 91%. After four invalid questionnaire copies with incomplete data or answers were removed, 66 valid questionnaire copies were left, an effective response rate of 94%.

Research instruments of this study include: an affective tutoring system about dengue fever prevention, a primary school students’ information literacy questionnaire, a primary school students’ cognitive load scale, and a quiz of students’ dengue fever prevention knowledge.

3.2.1 Affective Tutoring System about dengue fever prevention

Other than using dengue prevention courses designed and distributed by the environmental protection administration as teaching materials, our research team also added the module of semantic and emotion recognition and relevant modules to this teaching mode. This system, which can be divided into affective computing and course teaching, consists of four modules: a semantic identification module, a pedagogical assistant agent module, a dengue fever prevention teaching materials module, and a teaching strategies module.

(1) semantic identification module

The procedures of building a semantic identification module under the category of affective computing include compiling a dictionary of emotions, processing semantic structures, and using Semantic Clues Emotion Voting Algorithm (SeCeVa) to identify emotions. Methods of how the semantic identification module was constructed and the operating procedures are shown in Figure 2:

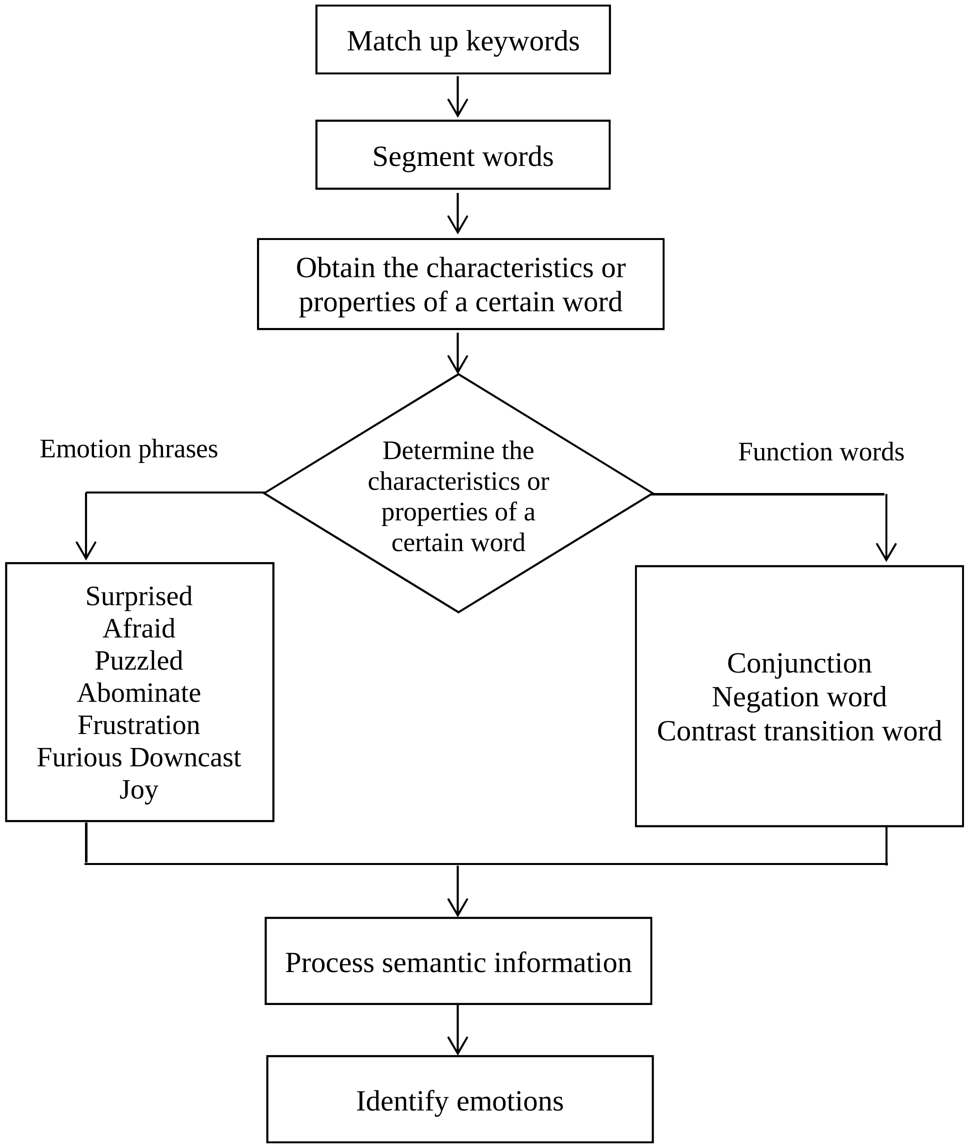


Figure 2 methods of how the semantic identification module was constructed and the operating procedures

Considering the similarity of categories of emotions and the system's promptness in perceiving learners' emotions and giving feedback, the six basic emotions proposed by Ekman (1972) and eight academic emotions proposed by Pekrun et al.(2009) were consolidated and classified to become this system's eight emotion categories derived from semantic identification. After giving definitions to each emotion, the collected emotion keywords were classified and the characteristic or property of a certain word was marked next to the word. How the emotion keywords were consolidated is shown in Figure 3.

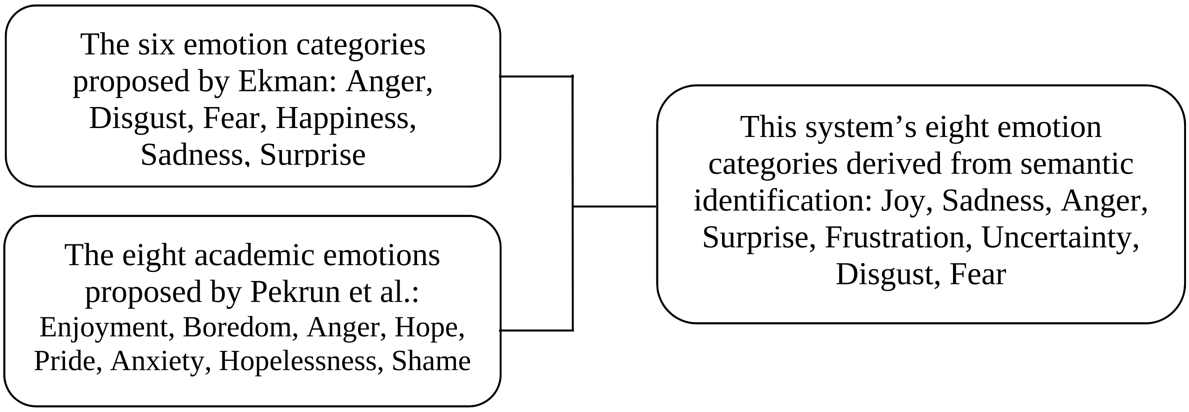


Figure 3 a flowchart of the integration of emotion keywords

(2) A pedagogical assistant agent module

A pedagogical agent, who acts as an intermediary between learners and messages from the system, could provide learners with opportune feedback through a two-way interactive mechanism which involves informing learners the system's current actions and understanding learners' needs. A screen shot showing the interaction between a learner and a pedagogical agent is shown in Figure 4.

Figure4 the interaction between a learner and a pedagogical agent

3.3 Experimental procedures

The one-group pretest-posttest research design is adopted for this study's experiment. Prior to using the affective tutoring system, students are instructed to fill out an information literacy questionnaire, take the dengue fever prevention knowledge pre-test, and receive a score from the pre-test. The same procedures are performed in each class for a total of four times. The next week, students are instructed to fill out a cognitive load assessment questionnaire, take the dengue fever prevention knowledge post-test in a computer classroom immediately after using the affective tutoring system, and receive a score from the post-test. The same procedures are also performed in each class for a total of four times, and the experiment lasts for two weeks in total. Following the affective tutoring system-based course and questionnaires, students are divided into groups based on their information literacy. The top 27 % of students are allocated to the higher score group, the bottom 27% are allocated to the lower score group, and the remaining 46% are allocated to a new group. Further, three students are selected from each group and a total of nine students are selected to participate in a focus group interview. Also, the whole interview is recorded with a digital video camera instead of a sound recording device for better accuracy of recording. The experimental procedures are shown in Figure 5.

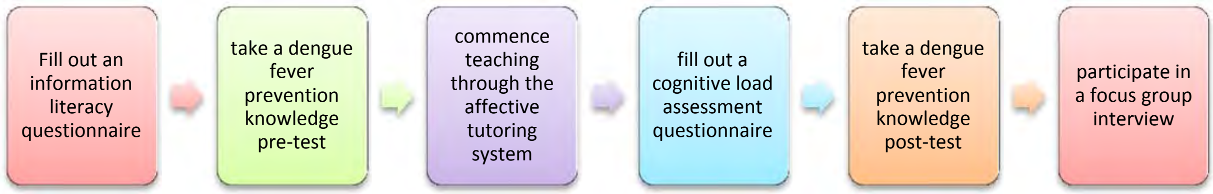


Figure5 Experimental procedures

4. Experimental results

This study aims to explore: 1. whether an affective tutoring system enhances students' learning outcomes; 2. difference in students' cognitive load in the process of using an affective tutoring system and students' learning outcomes afterwards due to students' varied information literacy; 3. the correlation between students' information literacy, cognitive load, and learning outcomes. T-test is employed to analyze students' scores at the pre-test and the post-test to compare students' learning outcomes before and after using the affective tutoring system and to test H1. Mean scores, standard deviation, and t-test results, which show students' learning outcomes before and after using the affective tutoring system, are shown in Table 1.

Table 1: Summarized statistics of students' learning outcomes

	Test	Number of students	Mean	Standard deviation	T value
Learning outcomes	Pre-test	66	76.73	14.50	-4.357***
	Post-test	66	84.73	11.38	

***p<.001

As shown in Table 1, "H1: teaching by means of an affective tutoring system can enhance students' learning outcomes" is supported as students' mean score on the post-test is higher than the mean score on the pre-test. Also, the quantitative data and results of qualitative data analysis validate each other as all students in the interview unanimously speak highly of the effectiveness of teaching by means of an affective tutoring system.

Independent sample t-test is employed to analyze the completed primary school students' information literacy questionnaire copies to find out difference in students' cognitive load and learning outcomes as well as to test "H1: teaching by means of an affective tutoring system can enhance students' learning outcomes" and "H2: students with varied information literacy would experience significantly different cognitive load while receiving teaching through an affective tutoring system." Mean scores, standard deviation, and t test results in relation to the two aspects of cognitive load and the overall cognitive load of students with varied information literacy are shown in Table 2.

As shown in Table 2, primary school students in the higher information literacy group and the lower information literacy group demonstrate significant difference of .05 on the interface operation aspect of cognitive load:. The t-value is -2.20, indicating that in comparison with students in the lower information literacy group, students in the higher information literacy group demonstrate significantly lower cognitive load on the interface operation aspect.

Nevertheless, the difference on content and overall cognitive load does not reach a level of significance, indicating primary school students' varied information literacy makes no significant difference on overall cognitive load and cognitive load on the content aspect. Further, mean scores, standard deviation, and t test results of learning outcomes of primary school students with varied information literacy are presented in Table 3. According to Table 13, difference on learning outcomes of students in the higher information literacy group and students in the lower information literacy group does not reach a level of significance, meaning that primary school students' varied information literacy does not make a significant difference on students' learning outcomes.

Table 2 Summarized statistics of the interplay between students' varied information literacy and students' cognitive load

Aspects in Relation to Cognitive Load	Information Literacy Group	Number of Students	Mean	Standard deviation	T value
Content	higher information literacy group	18	12.94	3.50	1.42
	lower information literacy group	19	11.11	4.29	
Interface Operation	higher information literacy group	18	15.06	6.88	*
	lower information literacy group	19	20.16	7.21	
Overall Cognitive Load	higher information literacy group	18	28.00	8.04	-1.25
	lower information literacy group	19	31.26	7.91	

* $p<.05$

Table 3 Summarized statistics of the interplay between students' varied information literacy and students' learning outcomes

	Information Literacy Group	Number of Students	Mean	Standard deviation	T value
Learning Outcomes	higher information literacy group	18	90.22	9.33	1.78
	lower information literacy group	19	82.74	15.38	

5. Conclusions and implications

This study aims to explore the correlation between cognitive load and learning outcomes of primary school students with varied information literacy when an affective teaching system is used. According to results of this study, the current status of students' overall information literacy is relatively good and is at an upper - intermediate level. As for each aspect of students' information literacy, students receive the highest score in the "retrieving information" aspect and the lowest score in the "assessing and sharing information" aspect, indicating students' relative strength in retrieving information yet relative weakness in assessing and sharing information. While students are using the affective teaching system, the "content" aspect of students' perceived cognitive load receives a higher score while the "interface operation" aspect receives a lower score, indicating that content causes more cognitive load for students than interface operation does. With respect to whether teaching with an affective tutoring system could enhance students' learning outcomes, statistical analysis by means of t-tests is performed on students' scores before and after receiving the affective tutoring system-assisted teaching. Results of the analysis show that students' mean scores at the pre-test and the post-test have reached a level of significance, indicating that teaching by means of an affective tutoring system could effectively enhance students' learning outcomes. However, students with varied information literacy are only different in the interface operation aspect of cognitive load yet show no significant difference in overall cognitive load and the content aspect of cognitive load. Nevertheless, a multitude of factors, which include students' intelligence, motivation to learn, learning styles, leaning attitudes, learning anxiety, parents' attitudes, and many more, may affect outcomes of teaching by means of an affective tutoring system.

Therefore, these factors may be counted as research variables in future research which explores the influence or correlation between other variables and different aspects of learning outcomes. In this way, researchers may better understand the importance of different variables on different aspects of students' learning outcomes as well as provide a reference for teachers who use an affective tutoring system to teach.

References

Alexander, S., Sarrafzadeh, A., & Hill, S. (2006, June). Easy with eve: A functional affective tutoring system. In Workshop on Motivational and Affective Issues in ITS. 8th International Conference on ITS (pp. 5-12).

Ammar M. B., Neji M., Alimi A. M., & Gouardères G. (2010). The Affective Tutoring System. Expert Systems with Applications 37, 3013–3023.

Anderson, J. R., Corbett, A. T., Koedinger, K. R. & Pelletier, R. (1995). Cognitive tutors: lessons learned. Journal of the Learning Sciences 4, 167–207.

Beale.R, & Creed.C (2009), Affective Interaction: How emotional agents affect users. International Journal of Human-Computer Studies,67(9), 755-776.

Bhattacherjee, A.(1998). Managerial influences on intraorganizational information technology use: A principal-agent model.Decision Sciences, 29(1),39-162.

Burleson, W., & Picard, R. (2004). Affective agents: Sustaining motivation to learn through failure and a state of stuck.

Calix, R. A. (2012, May). Emotion Expression 3-D Synthesis From Predicted Emotion Magnitudes. In FLAIRS Conference.

Chakraborty S., Roy D. &Basu A. (2010). Development of Knowledge Based Intelligent Tutoring System Advanced Knowledge Based Systems.

Chandler, P., & Sweller, J. "Cognitive load theory and the format of instruction.", Cognition and Instruction, 8, pp. 293-332, 1991.

Chou, Y.C. (2004). "A study of primary school students' information literacy: primary schools in Beigang Township, Yunlin County, as examples". Unpublished, master's thesis, Graduate Institute of Information Management, Nanhua University, Chiayi.

Craig, S., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. Journal of Educational Media, 29(3), 241-250..

D'Mello, S. K., Craig, S. D., Witherspoon, A., Mcdaniel, B., & Graesser, A. (2008). Automatic detection of learner's affect from conversational cues. User modeling and user-adapted interaction, 18(1-2), 45-80.

Eisenberg, Mike & Lee, Hur-Li (1999). New Perspective on Learning and Teaching: The Big Six Approach to Information Literacy Instruction. The International Conference on Information Literacy and a Lifelong Learning Society, pp.398-409. Taipei City: National Taiwan Normal University

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. Academy of Management Review, 14(1), 57-74.

Fuji Ren. (2009). Affective Information Processing and Recognizing Human Emotion. lectronic Notes in Theoretical Computer Science 225 (2009), 39–50.

Goh, T. T., & Huang, Y. P. (2009). Monitoring youth depression risk in Web 2.0. VINE, 39(3), 192-202.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. Behavior Research Methods, Instruments, & Computers, 36(2), 180-192.

Graesser, A. C., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. (2006). Detection of emotions during learning with AutoTutor. In Proceedings of the 28th Annual Meetings of the Cognitive Science Society (pp. 285-290).

Halstead SB.,Pathogenesis of Dengue (1988).Challonges to molecular biology. Science, 239, 476--481.

Jensen, M. C. and Meckling, W. H. (1976). The theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(1), 305-360.

Jex, H. R.(1988). Measuring mental working: Problems, progress and promises. In P. A. Hancock & N. M. Meshkati, (Eds). Human mental workload. Amsterdam North-Holland: Elsevier, 5-40.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. International Journal of Artificial intelligence in education, 11(1), 47-78.

Kao, E. C., Liu, C. C., Yang, T. H., Hsieh, C. T., & Soo, V. W. (2009, April). Towards Text-based Emotion Detection A Survey and Possible Improvements. In Information Management and Engineering, 2009. ICIME'09. International Conference on (pp. 70-74). IEEE.

Koedinger, K. R., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. na.

Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A. and Bhogal, R.S. (1997). The persona effect: Affective impact of animated pedagogical agents. Proceedings of CHI, 359-366.

Mao X.,& Li Z. (2010). Agent based affective tutoring systems: A pilot study. Computers & Education 55, 202-208.

Mayer, R. E.& Moreno, R.(2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. EducationPsychology, 38(1), 43-52.

McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011). A politeness effect in learning with web-based intelligent tutors. International journal of human-computer studies, 69(1), 70-79.

Miller, G.A. (1956). "The magic number seven plus or minus two: some limits on our capacity to process information". Psychological Review 63 (2): 81–97. doi:10.1037/h0043158.

Monath, T. P. (1994). Dengue: the risk to developed and developing countries. Proceedings of the National Academy of Sciences, 91(7), 2395-2400.

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87(2), 319-334.

Paas, F., Renkel, A., & Sweller, J. (2004). "Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture". Instructional Science 32: 1–8. doi:10.1023/B:TRUC.0000021806.17516.d0

Paas, F.G.W.C., and Van Merriënboer, J.J.G. (1993). "The efficiency of instructional conditions: An approach to combine mental-effort and performance measures". Human Factors 35 (4): 737–743.

Pass, F., Renkle, A., & Sweller, J.(2003). Cognitive load theory and instructional design: Recent development. Educational Psychologist, 38(1), 1-4.

Picard, R. & Kort, B., Reilly, R., (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy building a learning companion. In Proceedings IEEE international conference on advanced learning technology: Issues, achievements and challenges, 43–48.

Ptaszynski, M., Rzepka, R., Araki, K., & Momouchi, Y. (2012, March). A robust ontology of emotion objects. In Proceedings of The Eighteenth Annual Meeting of The Association for Natural Language Processing (NLP-2012) (pp. 719-722).

Quan, C., & Ren, F. (2009, August). Construction of a blog emotion corpus for Chinese emotional expression analysis. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3-Volume 3 (pp. 1446-1454). Association for Computational Linguistics.

Quan, C., Ren, F. 2010. "A blog emotion corpus for emotional expression analysis in Chinese", Computer Speech & Language, Vol. 24, Issue 4, pp. 726-749.

Sarrafzadeh, A. (2002). Representing domain knowledge structure in Intelligent Tutoring Systems. Proceedings of the International Conference on Information and Communication Technologies in Education. Spain, November 02, 665-9.

Sarrafzadeh, A., Hosseini, H. G., Fan, C., & Overmyer, S. P. (2003, July). Facial expression analysis for estimating learner's emotional state in intelligent tutoring systems. In Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference on (pp. 336-337). IEEE.

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., & Bigdeli, A. (2008). "How do you know that I don't understand?" A look at the future of intelligent tutoring systems. Computers in Human Behavior, 24(4), 1342-1363.

Sun, Y. T., Chen, C. L., Liu, C. C., Liu, C. L., & Soo, V. W. (2010). Sentiment classification of short Chinese sentences. In Proceedings of 22nd Conference on Computational Linguistics and Speech Processing (pp. 184-198).

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive science, 12(2), 257-285.

Sweller, J., Van Merrienboer, J. J. G. & Pass, F. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10, 251-296.

Van Merrienboer, J. J. G., Kirschner, P. A., & Kester, L.(2003). Taking the load off a learner's mind: Instructional design for complex learning. Educational Psychologist, 38(1), 5-13.

Wang Y., Guan L. (2005). Recognizing human emotion from audiovisual information. Proceedings of IEEE International Conference on Acoustics, Speeh, and Signal Processing(ICASSP), 2, 1125-1128, Philadelphia, PA, USA, March 18-23.

Wang, C.-Y., Chang, C.-W.& Chen, G.-D. (2009). Design an empathic virtual human to encourage and persuade learners in e-learning systems. ACM Multimedia MDTL.

Wei, W.J. (2007). "A study of primary school students' cognitive load in a digital learning environment: scientific articles as examples". Unpublished, master's thesis, Graduate Institute of Applied physics and Chemistry, University of Taipei, Taipei City.

Xu, G., Meng, X., & Wang, H. (2010, August). Build Chinese emotion lexicons using a graph-based algorithm and multiple resources. In Proceedings of the 23rd International Conference on Computational Linguistics (pp. 1209-1217). Association for Computational Linguistics.

Yan, J., Bracewell, D. B., Ren, F., & Kuroiwa, S. (2008). The Creation of a Chinese Emotion Ontology Based on HowNet. Engineering Letters, 16(1), 166-171.

Zhang, C., Zuo, W., Peng, T., & He, F. (2008, November). Sentiment classification for chinese reviews using machine learning methods based on string kernel. In Proceedings of the 2008 Third International Conference on Convergence and Hybrid Information Technology (Vol. 2, pp. 909-914).

Development of Digital Game-based Biology Learning Experience on Cell Cycle through DSLM Instructional Approach

Porntip KANYAPASIT[a] **& Niwat SRISAWASDI**[b*]
[a]*Science and Technology Education Program, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
[b]*Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Khon Kaen, Thailand*
*niwsri@kku.ac.th

Abstract: Digital games were several used as instructional tool in science education in several countries. Using games could increase students' motivation and interest to learn in regular class. The purposes of this study were to design and develop digital game-based learning by inquiry regarding biology of cell cycle based on dual-situated learning model (DSLM) approach, and evaluate effectiveness of the digital game-based learning on improvement of students' conceptual understanding of cell cycle. In this paper, the researchers report results of an investigation of students' existing ideas on biological concepts about cell cycle. 36 twelfth-grade students were recruited into the investigation by undertaking a series of open-ended conceptual questions covering five main concepts i.e. cell cycle, interphase, mitosis, meiosis and cytokinesis. The results indicated that they hold different patterns of alternative- and misconceptions about cell cycle, and some had no conception on the concepts. Then, these findings on specific mental set students lack were used to particularly design dual-situated learning events for the cell cycle concepts. Purposively, a digital game was created by the researchers in order to use for eliminating or decreasing students' alternative- and misconceptions about cell cycle. Besides, the digital game was designing to use as inquiry-based tool to support constructing scientific understanding on the concepts and it is illustrated in this paper. An implication of this paper is to provide pedagogical guidance for developing digital game-based biology learning environment with regarding effective strategy for science learning, open-inquiry learning process.

Keywords: digital game, open inquiry, dual-situated learning model (DSLM), biology education, cell cycle

1. Introduction

In recent years, there are several studies with purpose to promote students' scientific conceptual understanding and conceptual change in biology education. Several researchers developed instructional models and teaching techniques for enhancing the leaning of biological concepts such as cellular respiration (Songer and Mintzes, 2006), genetics (Termtachatipongsa, 2014; Opfer and Siegler, 2004). The results from previous studies reported that students often hold alternative conceptions about cell cycle, especially interphase, mitotic and meiotic cell division, and cytokinesis, due to its abstraction and invisible by nature (Termtachatipongsa, 2014). Also, researcher mentioned that teaching method used in biology class might be a major factor which causes students alternative- and misconceptions (Obaidat and Malkawi, 2009). Moreover, lack of prior knowledge and appropriate ideas could affect ineffective conceptual learning in science (Obaidat and Malkawi, 2009; She, 2004). As such, to promote students' scientific understanding and conceptual development in science, several researchers have attempted to develop instructional materials and models for teaching and learning of science concepts. The dual-situated learning model (DSLM) was proposed by She (2004) and it is one of the instructional models which considers students' alternative conceptions to be a very important consideration in process of learning, and many researchers reported successful on the use of DSLM for enhancing students' conceptual understanding in science (Lee and She 2010; Liao and She 2009; She and Liao 2010; Srisawasdi and Kroothkeaw, 2014; Srisawasdi and Sornkhatha, 2014). However, no study has utilized digital game into this learning model.

In recent years, several researchers have paid attention to a new research trend that focused on using digital game for teaching biological concepts and other science concepts. With technological features, development of interactive computer-based learning materials for science teaching and learning provide opportunities to help student understand the concepts better by visualizing abstract science concepts into concrete experience to change students' alternative conceptions to scientific conceptual understanding (Srisawasdi, Kerdcharoen, and Suits, 2008; Suits and Srisawasdi, 2013). Moreover, teaching-learning process by using computer technology such as digital game could be a novel pedagogy to promote meaningful learning and students' motivation better than traditional teaching-learning process. Therefore, with the abovementioned reasons, importance of using digital game cooperated with DSLM for conceptual change in science learning, the researchers interested to develop digital game-based inquiry learning in biological concepts about cell cycle through DSLM approach for facilitating students' scientific conceptual understanding and cognitive process of conceptual change.

2. Literature review

2.1 Dual-situated Learning Model (DSLM)

Dual-situated learning model (DSLM) (Lee and She 2010; Liao and She 2009; She and Liao 2010). is one of instructional model that considered prior knowledge, alternative conceptions of students and use these data for generating instructional tools for helping students to have correct science concepts through process of conceptual change. The DSLM comprised of six major stages: (1) examining the attributes of the science concept to provide information in which essential mental sets are needed to construct a scientific view of the concepts; (2) probing students' misconception on the concept; (3) analyzing for mental sets in which the students lack to pinpoint which and how many particular mental sets students lack for restructuring the science concepts based upon the first pair of DSLM theory; (4) designing dual-situated learning events including the ideas of second and third duals of DSLM; (5) instructing with dual-situated learning events to provide students an opportunity to make predictions and provide explanations before and after the event, and to further explain why they changed their conceptions or retained their original conceptions and (6) instructing with challenging situated learning event to provide an opportunity for the students to apply the mental sets they have acquired to a new situation, ensuring that successful conceptual change to occur (She and Liao 2010).

According to several previous studies, DSLM had been used in learning of physics and chemistry, and the results showed that students had meaningful learning in science concept through process of conceptual change (Srisawasdi and Kroothkeaw, 2014; Srisawasdi and Sornkhatha, 2014) However, there was no study on application of DSLM in biology class before.

2.2 Games-based Learning in Science

A game is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (Salen and Zimmerman, 2004). General characteristics of game are fun and challenge (e.g. leader board). Effective educational computer game should have five characteristics: (i) built on learning principle; (ii) provide personalized learning opportunities; (iii) provide more engagement for learner; (iv) teach 21th century skills; and (v) provide an environment for authentic and relevant assessment (McClarty et al., 2012).

Currently, digital game or computer game plays common role to provide fun and relax because the game can challenges the player to a digital control of the game on their own. As a result, educators are interesting into how to use the game to facilitate and enhance teaching and learning in subject matters. Kumar (2000) suggested utilization of computer games as instructional tools to stimulate intrinsic motivational factors that encourages curiosity and creates the impression to the students by self-controlled learning. In additions, there were several researches showed that digital games can promote students' learning and their positive attitudes (Sung and Hwang, 2013), support development of critical thinking (Squire, 2006), and creativity (Annetta, Cheng, and Holmes, 2010). Compared to traditional class, digital game can increase motivation that make students interested to learning better than the traditional (Bergin and Reilly, 2005; U.S. Department of Education, 2010). Moreover, learning with digital games promote collaborative process for learning (Gee, 2005; Sung and Hwang, 2013; Wu, Chiou, Kao, Hu, and Huang, 2012).

3. Methods

3.1 Participants

The participants in this study were 36 twelfth-grade students, age ranging from 17 to 18 years old in a public school in Northeast of Thailand. They are attending a biology course for basic education level and they study in special science program for gifted students. Based on background knowledge of twelfth grade students were learned contents about cell cycle before, in tenth grade.

3.2 Instrument

The instrument in this study is ten open-ended conceptual question items regarding five major concepts about biology of cell cycle, i.e. cell cycle, interphase, mitosis, meiosis and cytokinesis. Before employ the instrument to elicit students' understanding of these concepts, it was processed construct and communication validity by four independent experts, including two biologists and two biology teachers, and two of ten items were revised following the experts suggestions.

3.3 Procedures and Data Analysis

All students took 60 minutes to complete the series of open-ended conceptual question items. Following this, content analysis was the primary method for analysis of students' written responses to the open-ended question items, represented their conceptual understanding about biology of cell cycle. The researchers began with repeatedly read the students' written responses and then development of a general conceptual understanding category. The researchers have analyzed, interpreted, and classified their responses into five categories i.e. scientific conception (SC), incomplete scientific conception (IC), alternative conception (AC), misconception (MC), and no conception (NC). Then the researchers have designed a series of dual-situated learning events for facilitating mechanism of change and revise of their alternative and misconceptions of cell cycle into scientific conception. The dual-situated learning events were emphasizing into the design of a digital game of cell cycle.

4. Results and Discussion

Based on the five categories (SC, IC, AC, MC and NC) interpreted students' unscientific conceptions, the percentages of quantity of a combination between students' alternative- and misconceptions and their no conceptions on the interphase concept was displayed in Figure 1.

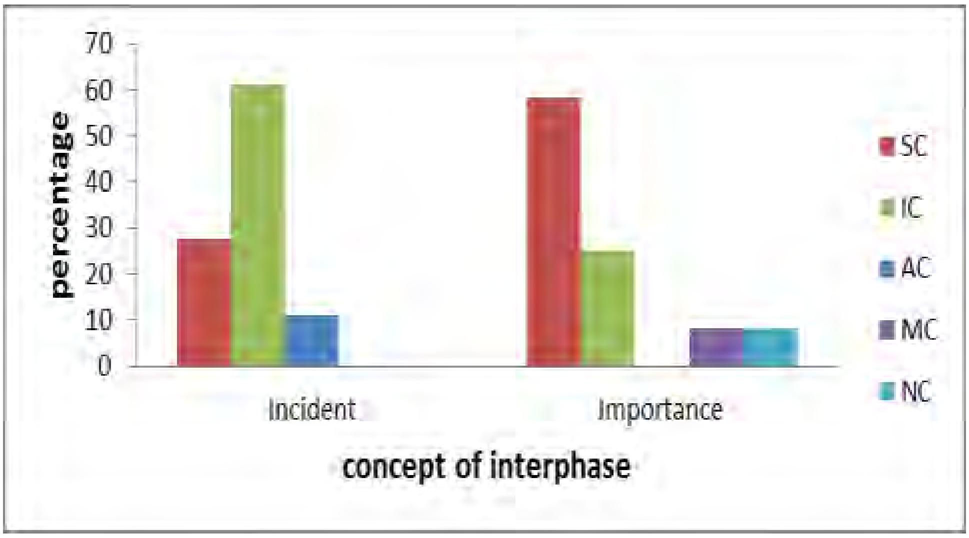


Figure 1. Distribution of students' alternative conceptions on interphase concept

According to Figure 1, the percentages for combination of alternative- and misconceptions for incident of interphase and importance of interphase were 11.1% and 8.3%, respectively. The percentages of no conception for incident of interphase and importance of interphase were 0% and 8.3% respectively. The result of students' alternative-, mis-, no conceptions on meiosis concept was illustrated in Figure 2.

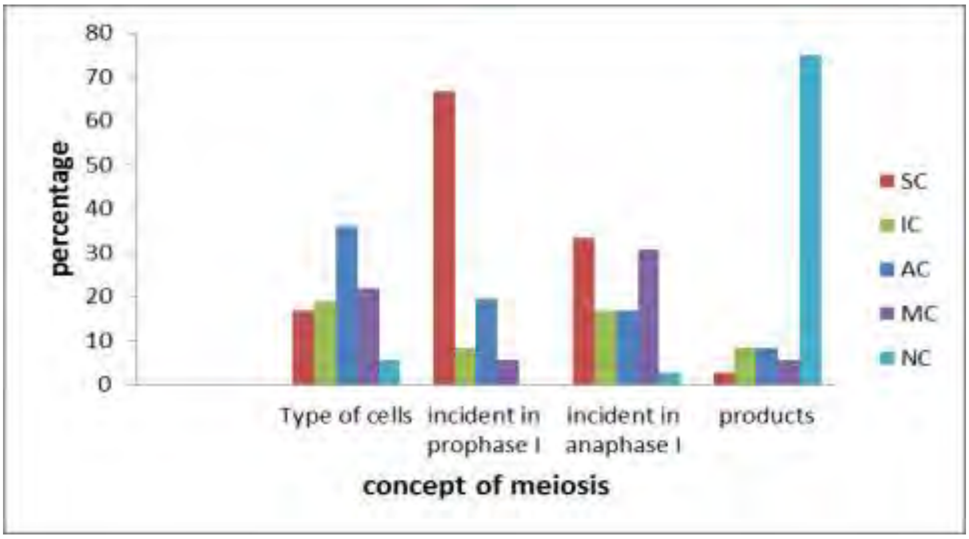


Figure 2. Distribution of conceptual understanding on meiosis concept.

As seen in Figure 2, the highest percentages for combination of alternative- and misconceptions on the concept of meiosis were (i) type of cells and properties of organism, (ii) important incident in anaphase I, (iii) important incident in prophase I, and (iv) products, and there were 58%, 47.3%, 25%, and 13.9% respectively. The percentage of no conception was highest on the products concept (75%), and there was none of the students who had no conception on the important incident in prophase I. This means the researchers should focus to design learning event for the process of conceptual change for this sub-concept, but it need to be ground the students' conceptual understanding in the learning events for another three meiosis sub-concepts. In the next, Figure 3, the results of students' conceptions of process and products within mitosis concept were presented.

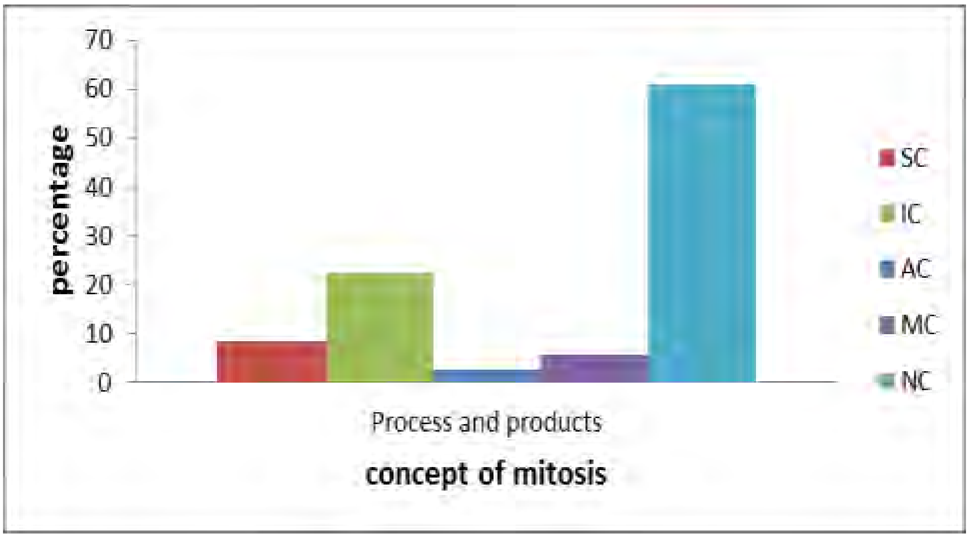


Figure 3. Distribution of conceptual understanding on concept of mitosis.

As show in Figure 3, there was a small number of the students who hold alternative- and misconceptions on process and products sub-concept (8.4%). However, more than a half of them showed no conception on the concept (61%). This means there need help for facilitating construction of conceptual understanding regarding the concept of process and products., The result of students' alternative-, mis-, no conceptions on cytokinesis concept was depicted in Figure 4.

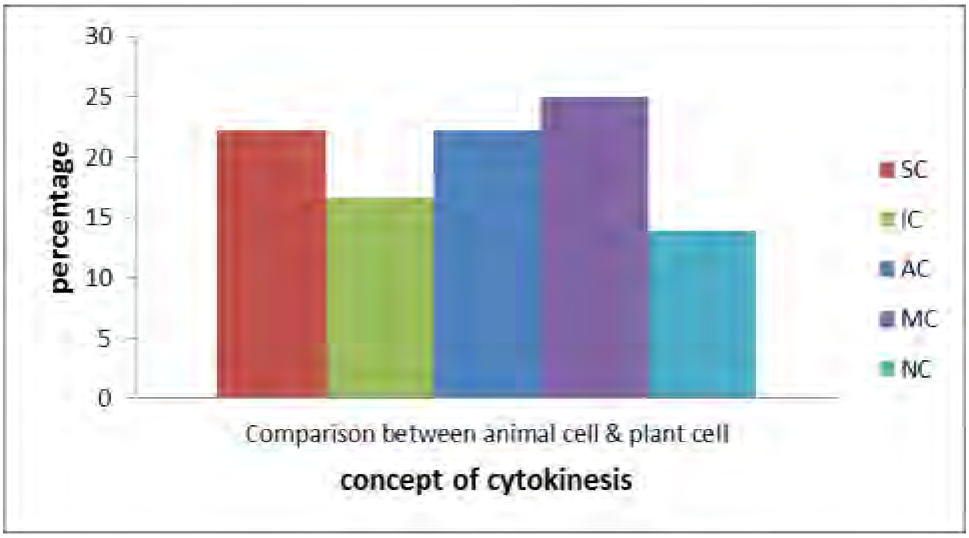


Figure 4. Distribution of conceptual understanding on cytokinesis concept.

Figure 4 shows the percentage for combination of alternative- and misconceptions for comparison between animal and plant cell was about the half of them (47.2%), and for no conception was 13.9%. This revealed that this concept need a couple for student learning, both grounding scientific concept and changing unscientific concept. For the main concept of cell cycle, Figure 5 displayed students' conceptions on the sub-concept of important of cell cycle and types of cell.

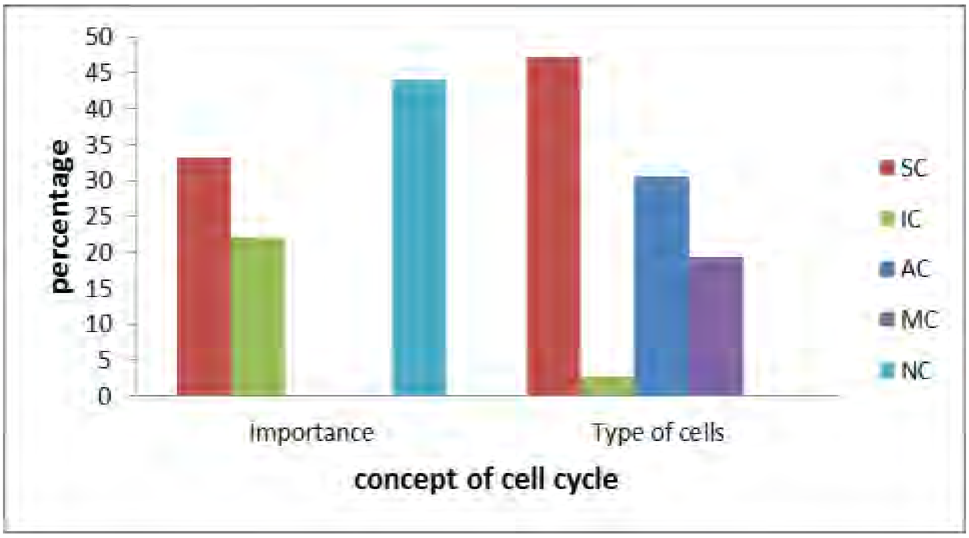


Figure 5. Distribution of conceptual understanding on the cell cycle concept.

In Figure 5, none of students' unscientific conceptions on importance of cell cycle were found and only 22.2% of them showed incomplete conceptions in this sub-concept. This means most of them were able to comprehend the sub-concept from the regular class. Nevertheless, a half of them (50%) hold unscientific conceptions on type of cells sub-concept, and there was 44% who had no conception about the concept.

The findings of this study consistent with the previous finding that many students hold alternative conceptions in mitotic and meiotic division (Ozcan, Yildirim, and Ozgur, 2012). Moreover, the results also confirm to a research carried out by Lewis and Robinson (2000) that high school students hold unscientific understanding, and encounter learning difficulty and they cannot understood on biological concepts of DNA, gene, chromosome, mitotic and meiotic division.

According to the DSLM approach the researcher used, Table 1 shown examples of students' alternative conceptions in cell cycle concepts and the design of learning events to address and facilitate conceptual change. Moreover, the next section illustrates the design of digital game-based inquiry learning in biology of cell cycle.

Table 1: Designing dual-situated learning events regarded students' alternative conceptions on cell cycle.

Concept	Sub-concept (item)	Example of students' alternative conceptions	Design of learning event to address the alternative conceptions
Interphase	Incident of Interphase (1)	DNA synthesis occurs in G_1 phase.	Comparison of necessary compounds for biological process occurred in G_1, S, and G_2 stage (Learning Event#1)
	Importance of Interphase (2)	-	-
Meiosis	Type of cells & properties of organism (3)	Meiosis occurs in reproductive organs, for example, testes and ovaries.	Comparison between the origin of somatic cell and reproductive cell (spermatogenesis and oogenesis) (Learning Event#2)
	Important incident in prophase I (5)	Crossing over occurs in prophase II of meiosis process.	Identification of key characteristics of "crossing over" (Learning Event#3)
	Important incident in anaphase I (6)	Segregation of homologous chromosomes occurs in telophase I.	Comparisons among meiosis I, meiosis II, and mitosis (Learning Event#4)
	products (7)	Segregation of homologous chromosomes occurs in anaphase II.	

Mitosis	Process and products (4)	When mitosis completed, there is followed by meiosis.	
Cytokinesis	Comparison between animal cell and plant cell (8)	Similarly, both plant and animal cell have process of mitotic karyokinesis.	Comparison between cytokinesis of plant cell and animal cell (Learning Event#5)
Cell cycle	Importance of Cell cycle (9)	-	
	Type of cells (10)	There may be fix over.	Comparison of different types of cell (Learning Event#6)

4. The Design of Digital Games about Cell Cycle

According to the stage 4 of DSLM, learning events associated students' unscientific conceptions on cell cycle has been design and presented in the previous section (See Table 1). To develop the digital games which facilitate process of conceptual change, the researchers employed the designed learning events as a basis to create a prototype of "The Cell Cycle Game". This game consists of two parts; game playing and visualization. Figure 6 illustrates interaction part of game playing in the cell cycle game, and Figure 7 displays visual representation about cell division, both mitosis and meiosis.

Figure 6. Example illustration of "The Cell Cycle Game": (A) a screen for selecting a car represented a type of cell; (B) a screen for controlling the car (a type of cell) and collecting nutritive essence for cell development.

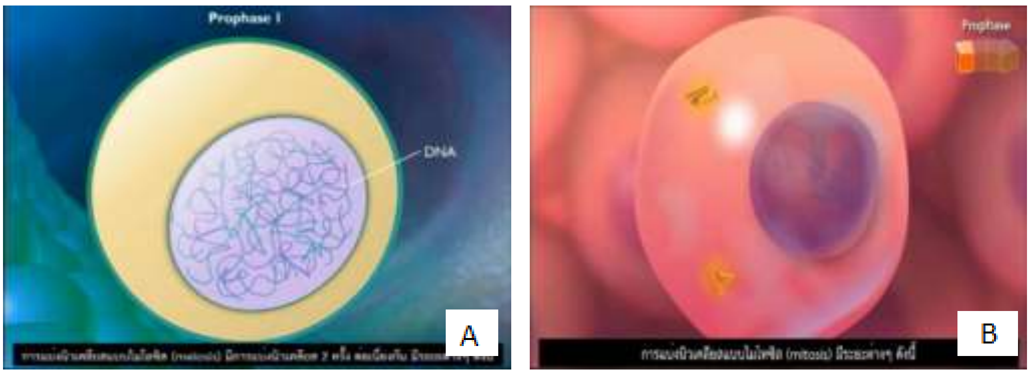


Figure 7. Example illustration of visualization, embedding in the game: (A) a screen representing mitosis process of cell division; (B) a screen representing meiosis process of cell division.

5. A Proposed Learning Process of Digital Game-based Open Inquiry with Dual- situated Learning Model on Cell Cycle

This section presents a proposed inquiry learning process using the cell cycle game. The simulation-based open inquiry with dual-situated learning model (Srisawasdi and Kroothkeaw, 2014;

Srisawasdi and Sornkhatha, 2014) was applied for the proposed learning process. Table 2 displayed an example of the learning process on cell cycle concept.

Table 2: Example of the design of learning process by game-based open inquiry with dual-situated learning model

The main stage	Components of the proposed learning process	Example of learning process
Pre-gaming	Open-ended inquiry question	Teacher provides an open-ended inquiry question: "How many DNA during the cell division?"
	Scientific background/ information	Teacher encourages collaborative discussion on basic information related the question by presenting a concept map of cell cycle.
Collaborative gaming	Procedure/design	Students are classified into small group to play game. Each group designs their own task, independent selecting a type of cell, to play game. 
	Data and result analysis	After the interacting with the game, students make a decision to analyze obtained playing game data e.g. scores and interpret it into a graphical representation. 
Post-gaming	Result communication	Students have to select the way to present, communicate, and discuss the meaning of data for whole class.
	Conclusion	Students have to collaboratively make a relationship between each group results and then show conclusion as the best answer to the provided question.

In this learning process, student will collaborative work together in small groups of three to five members. This pedagogy begins with an open-ended driving question targeted to alternative conceptions about cell cycle commonly found in students. To assist the process of hypothesis generation addressed the question, essential scientific backgrounds are provided to students. Then, students are required to perform generating testable hypotheses, designing an investigation with the cell cycle game. During playing the game, each group was assigned to access Google Drive spreadsheet, preparing by instructor, for recording scores and what they found into a predetermined table. In an addition, each group was assigned to analyze the recorded data by comparing individual score and also use Google Chat for discussing in the group. When they finished the game, all groups have to communicate findings among groups by creating a PowerPoint presentation via Google Drive presentation. Finally, instructor induces students into a forum for drawing a conclusion based on evidence and collaborative explaining the result of hypotheses testing.

Acknowledgements

This work was financially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) and Graduate School, Khon Kaen University. The author would like to express gratefully acknowledge to Faculty of Education, Khon Kean University, for supporting this contribution.

References

Annetta, L. A., Cheng, M.-T., & Holmes, S. (2010). Assessing twenty-first century skills through a teacher created video game for high school biology teachers. *Research in Science and Technological Education*, 28(2), 101-114.

Bergin, S., & Reilly, R. (2005). The influence of motivation and comfort-level on learning to program. In *Proceedings of the 17th workshop on psychology of programming PPIG'05*, 293-304.

Duncan R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: students' understandings of molecular genetics. *Journal of Research in Science Teaching, 44*(7), 938-959.

Gee, J. P. (2005). *Why video games are good for your soul*: Pleasure and learning. Champaign, IL: Common Ground Publishing.

Kumar, D. (2000). Pedagogical dimensions of game playing. *ACM Intelligence Magazine*, 10(1), 9-10.

Lee, C. Q., & She, H. C. (2010). Facilitating students' conceptual change and scientific reasoning involving the unit of combustion. *Research in Science Education*, 40(4), 479–504.

Lewis, J., Leach, J. and Wood-Robinson, C. (2000). Chromosomes: the missing link-young people's understanding of mitosis, meiosis, and fertilization. *Journal of Biological Education*, 34(4), 189-191.

Liao, Y. W., & She, H. C. (2009). Enhancing eight grade students' scientific conceptual change and scientific reasoning through a web-based learning program. *Educational Technology & Society*, 12(4), 228–240.

McClarty, K.L., Orr, A., Frey, P.M., Dolan, R.P., Vassileva, V., & McVay, A. (2012). A Literature Review of Gaming in Education. Retrieved fromhttp://researchnetwork.pearson.com/wp-content/uploads /Lit_Review_of_Gaming_in_Education.pdf

Obaidat, I., & Malkawi, E. (2009) 'The grasp of physics concepts of motion: identifying particular patterns in students' thinking', *International Journal for the Scholarship of Teaching & Learning*, 3(1), 1–16.

Opfer, J. E., & Siegler S. R. (2004). Revisiting preschoolers' living things concept: A microgenetic analysis of conceptual change in basic biology. *Cognitive Psychology*, 49 (2004), 301–332.

Ozcan, T., Yildirim, O., & Ozgur, S. (2012). Determining of the university freshmen students' misconceptions and alternative conceptions about mitosis and meiosis. *Procedia - Social and Behavioral Sciences*, 46, 3677 – 3680.

Salen, K., & Zimmerman, E. (2004). Rules of play [Kindle]. Cambridge,MA:MITPress.

Sadler T. D., Romine W. L., Stuart P. E. and Merle-Johnson D. (2013). Game-Based Curricula in Biology Classes: Differential EffectsAmongVarying Academic Levels. *Journal of Research in Science Teaching*, 50(4), 479-499.

She, H.C. (2004) 'Fostering "radical" conceptual change through dual situated learning model', *Journal of Research in Science Teaching*, 41(2), 142–164.

She, H. C., & Liao, Y. W. (2010). Bridging scientific reasoning and conceptual change through adaptiveweb-based learning. *Journal of Research in Science Teaching*, 47(1), 91–119.

Songer C.J., & Mintzes J. J. (2006). Understanding cellular respiration: An analysis of conceptual change in college biology. *Journal of Research in Science Teaching*, DOI: 10.1002/tea.3660310605.

Squire, K. (2006). From content to context: Videogames as designed experience. *Educational Researcher*, 35(8), 19–29.

Srisawasdi, N. (2014). The effect of simulation-based inquiry on students' conceptual learning and its potential applications in mobile learning. *International Journal Mobile Learning and Organisation*, 8(1), 28-49.

Srisawasdi, N., Kerdcharoen, T., & Suits, J.P. (2008). 'Turning scientific laboratory research into innovative instructional material for science education: case studies from practical experience', *International Journal of Learning*, 5(15), 201–210.

Srisawasdi, N., & Kroothkeaw (2014). Supporting students' conceptual development of light refraction by simulation-based open inquiry with dual-situated learning model. *Journal Computer Education*, DOI 10.1007/s40692-014-0005-y.

Suits, J.P. and Srisawasdi, N. (2013). 'Use of an interactive computer-simulated experiment to enhance students' mental models of hydrogen bonding phenomena', in Suits, J.P. and Sanger, M.J. (Eds): *Pedagogic Roles of Animations and Simulations in Chemistry Courses*, ACS Symposium Series 1142, American Chemical Society, Washington, DC.

Sung, H., & Hwang, G. (2012). A collaborative game-based learning approach to improving students learning performa nce in science courses. *Computers & Education*, 63 (2013), 43–51.

Termtachatipongsa, P. (2014). Instructional Context, Motivational beliefs, Attitude and Strategies in Teaching Genetics for Conceptual Change. *Procedia - Social and Behavioral Sciences***,** 116(2014), 2023 – 2029.

U.S. Department of Education. (2010, November). *Transforming American education: Learning powered by technology*. Retrieved from http://www.ed.gov/sites/default/files/netp2010-execsumm.pdf

Wu, W., Chiou, W., Kao, H., Hu,C. A., & Huang, S. (2012). Re-exploring game-assisted learning research: The perspective of learning theoretical base. *Computers & Education*, 59(2012), 1153– 1161.

Mobile Augmented Reality in Supporting Performance Assessment: An Implementation in a Cooking Course

Kuo-Hung CHAO [a], Chung-Hsien LAN[b*], Yao-Tang LEE[c], Kinshuk[d], Kuo-En CHANG[e] , & Yao-Ting SUNG[f]

[a]*Assistant Professor, Hsing Wu University, Taiwan*

[b]*Assistant Professor, Taoyuan Innovation Institute of Technology ,Taiwan*

[c]*Assistant Professor, Nan Kai University of Technology, Taiwan*

[d] *Professor, Athabasca University, Canada*

[e,f] *Professor, National Taiwan Normal University, Taiwan*

f

*chlan@tiit.edu.tw

Abstract: In this study, a conceptual system framework was developed by integrating augmented reality (AR) technology to reduce the limitations in observation and assessment during performance assessment. Thus, the efficiency and reliability of mobile performance assessment can be enhanced. The processes of student performance can be presented from diverse approaches by using the characteristics of mobile devices and AR and considering the categories and situations of student performance. In this study, a novel mobile performance assessment system that incorporates AR to the processes of observation and assessment in performance assessment was developed. The mechanism emphasises the approaches adopted to present student works and provides opportunities for enhancing student communication and interaction. In addition, the system enables students to explain their works and incorporate the feedback they receive into future work. More importantly, mobile AR can be applied to offer personalised features and appropriate information in particular areas. Hence, students can interact with real or virtual information based on their needs. During this process, students can observe their own works from varying perspectives, acquire vital knowledge, develop the skills of critical thinking, and transform the process into a substantial self-established learning process.

Keywords: Augmented reality, performance assessment, mobile learning

1. Introduction

Conventional paper-based tests are adopted extensively in school education because the scoring is objective and the tests are easy to implement and have relatively high reliability and validity. However, this type of test has been criticised for being fragmented, having situations removed from the learning process, overemphasising cognition, and confining learning. Based on the popular educational philosophy of allowing students to develop diverse capabilities and achieve active knowledge building, performance assessment should be considered as a vital link in teaching. In addition to assigning a final score to students, the purpose of assessment is to develop a highly in-depth understanding regarding the process that students undergo during learning and to provide feedback to assist in student growth. Performance assessment, which is an assessment method that became widely used again after the 1980s, emphasises assessing actual performance and behaviour. This approach can be employed to compensate for the inadequacy of the paper-based test approach and assess the diverse disposition- and skill-based achievements of students. O'Neil (1992) indicated that assessment plays a vital role in teaching and that the process of assessment consists of goal setting, data collection, organisation, and result analysis. The results can be used to enhance teaching and report the actual progress of students. By conducting assessments, teachers can evaluate students' learning outcomes and identify relevant factors, which can be used to devise improved methods for enhancing the effectiveness and quality of teaching. Considering the value of assessment, establishing a sound assessment mechanism is imperative for achieving multiple purposes and functions. Most teachers have had the experience of being questioned by parents if their children have failed tests because the teachers are biased against their children. In the assessment of design works, particular students have the misconception that they should receive a passing grade as long as they submit their assignments. This phenomenon has generated the discussion whether a well-defined set of assessment criteria should be applied when teachers are assessing the design works of students, who are highly self-aware. These students directly question teachers regarding the assessment criteria once they realise that their grades are unsatisfactory. In other words, students are no longer passive recipients of assessment results, but rather are attentive regarding their learning outcomes.

Nevertheless, a major problem encountered by the education community is determining the appropriateness of educational evaluation. Beywl and Speer (2004) found that previous studies had emphasised that the various dimensions of assessment in the narrow and broad senses must be consistent for an evaluation to be substantial, indicating that the assessment methods employed should be diverse. Problem-solving abilities are often regarded as abilities that require high-level reasoning or thinking and intelligence. Performance assessment has been recognised as one of the most effective methods for assessing this type of high-level thinking because this approach emphasises the application and demonstration of abilities in problem-solving situations and the complexity of problem-solving processes (Wiggins, 1993).

Although less objective and convenient compared with conventional paper-based tests, performance assessment can result in encouraging education outcomes when implemented under certain conditions. Previous studies (Dunbar, Koretz, & Hoover, 1991; Jiang, Smith, & Nichols, 1997; Baker, 1996; Messick, 1992) have indicated that the primary limitations and disadvantages of the performance assessment approach include the lack of comparison, limited reliability, unsatisfactory economic performance, and low validity. However, the majority of these factors can be attributed to the subjective consciousness of the assessors and errors in the measured situations. By contrast, augmented reality (AR) technology can be employed to display, in real situations, real-time information that is necessary for assessing or learning. From the perspective of cognitive psychology, this approach can be applied to reduce the errors resulting from the process of performance assessment and to minimise the time and economic costs that teachers must bear when observing student behaviour. Therefore, we examined the meaning, relevant studies, and limitations of performance assessment before investigating the effects that incorporating AR technology exert on improving performance assessment systems. Subsequently, we applied an AR-based performance assessment system to a cooking course to explore the effects of the application. The results yielded by conducting the performance assessment and paper-based tests were compared before a conclusion and recommendations were provided. The results of this study can serve as a reference for implementing performance assessment in teaching.

2. Performance Assessment

Performance assessment requires that students apply the knowledge and skills they have learned to perform hands-on practice rather than simply revalidating and recollecting the experience of learning (Roeber, 1990). This assessment method satisfies the needs of the current trend of constructivist learning and teaching (Chang, 2002b). Performance assessment motivates students to integrate the knowledge, skills, and dispositions required in the subject, and the results of the assessment can reflect students' problem-solving abilities in real life and the interest and needs of the students. Performance assessments, which can be conducted to evaluate high-level cognitive abilities and the dispositions and skills of students, are more comprehensive in the contents tested compared with conventional paper-based tests. Performance assessment can be integrated into teaching activities rather than being separated from teaching. Thus, teaching can be performed without interruptions. Furthermore, by adopting specific performance and assessment criteria, teachers can provide students with specific feedback, which motivates students to take initiatives in learning and assume the responsibility to critique their own works and strategies. Therefore, performance assessment is an effective approach for facilitating teaching and learning.

Performance assessment was initially called performance-based assessment when the concept was first applied to education. Specifically, performance refers to the process of completing a task. The concept emphasises the authenticity and representativeness of the assessment regarding particular abilities as well as the significance of learning and evaluating in meaningful and real situations. By contrast, the concept of performance emphasises the necessity in challenging the intelligence, knowledge, and skills of the learners and the necessity in mastering the contents, processes, and outcomes of learning (Wiggins, 1993). Airasian (1996), Fitzpatrick and Morrison (1971), and Wiggins (1992) have indicated that in performance assessment, students are required to participate in an activity or create a piece of work to demonstrate the knowledge and skills they possess. Thus, students are required to demonstrate what they know and are capable of in actual situations. Aschbacher (1991) argued that the performance assessment in teaching-learning situations refers to teachers using their professional judgment to assess students' learning performance, which includes students' responses to tasks, the works delivered, and the process of learning. The characteristics of the performance assessment in this context are listed as follows: (a) students are expected to perform tasks or create objects that require high-level thinking or problem-solving skills; (b) the tasks based on which the students are assessed are meaningful, challenging, and integrated with teaching activities; (c) the tasks based on which the students are assessed are connected with real life; (d) processes and products are often the focuses of assessment; and (e) the assessment criteria and standards, which are vital dimensions and standards of the assessment, must be defined in advance (Herman et al., 1990).

The purpose of performance assessment is to motivate students to engage in useful, beneficial, and meaningful activities. Regarding form, this type of assessment is a component of learning activities, which require high-level thinking skills, an understanding of relevant concepts, and the ability to link various forms of knowledge. In addition, this type of assessment involves a specific explanation regarding the bases upon which student works are assessed. Therefore, the assessment is essentially a process of standard building rather than standard testing. Thus, the ultimate objective of performance assessment is to motivate students to comprehend the teaching contents and reach achievement standards by participating in teaching activities (Dorn, 1999). When performance assessment is conducted, students are required to apply particular learning results to daily situations. During the process of problem solving, students can reference necessary knowledge and independently build subject-specific knowledge and evaluate the possibility of the results. By so doing, students are motivated to independently determine what they want to learn and thus acquire knowledge actively. During performance assessment, the processes of thinking and the results yielded from the processes are both assessed. When implementing performance assessment, teachers can evaluate students' understanding of the problems, involvement, problem-solving skills, and self-expression abilities. Thus, the learning outcomes and learning processes of students can be fully reflected.

The purpose of performance assessment is using assessment to promote student development. The tasks designed for performance assessment, such as design-related experiments and system operation tasks, are often difficult for students to complete independently. Instead, students must work in teams and cooperate with their team members. Hence, during assessment, allowing other students and teachers to provide feedback on the assessment standards, record the processes, and evaluate the progress made by the student under assessment offers more opportunities for teacher–student and peer interactions. During these interactions, students can communicate and explain their learning experience and contemplate the learning process. In addition, the process of teamwork enables students to develop the abilities to communicate and cooperate with their peers and to develop favourable work attitudes. The difference between teachers (experts) and novices is that experts understand how to effectively use the knowledge they have acquired. Performance assessment enables teachers to comprehend the thinking process students undergo by observing the process of students' operations. Thus, teachers can understand whether the students' operations comply with the prescribed procedure.

To determine the effectiveness of a test or an evaluation method, we must explore the effectiveness of the method and the results based on the intended purposes. Ou (2002) asserted that one of the purposes of performance assessment is to offer a real and specific situation for students to demonstrate their knowledge and skills and thus make correct inferences regarding students' learning achievements. In this context, the accuracy with which the assessment results reflect the teaching goals can be enhanced. One of the characteristics of performance assessment is that students are allowed to flexibly use days of a week or several weeks to engage in high-level thinking before completing their tasks. In terms of presentation, no correct answers or uniform patterns apply in performance assessment. Instead, the ambiguity of situations enables students to adopt various approaches to demonstrate individual creativity.

Gronlund (1993) and Linn (1991) have indicated that the purpose of performance assessment is to establish a model that enables students to focus on real learning activities. A course-oriented assessment approach is established when performance assessment is closely integrated with a course. This form of assessment system can inspire students and teachers to strive to enhance learning outcomes. The methods used in performance assessment are often one of the following: (a) checklists, which are used to evaluate whether the assessees exhibit a particular behaviour; (b) rating scales, which are employed to evaluate whether the assessees exhibit a particular behaviour and to rate the behaviour based on the extent; and (c) anecdotal records, where texts are used to describe and interpret assessee behaviour.

3. Limitation of Performance Assessment in Learning

Performance assessment, where the rating is often performed by professionals based on their observation and judgment, is subjective, demanding, and low in reliability compared with paper-based tests. In addition, the fairness of performance assessment is often questioned because the results cannot provide immediate feedback on student performance. Therefore, a critical problem that necessitates solution in implementing performance assessment is devising fair and objective rating criteria that are easy to apply and can be used to provide feedback to students (Lu et al., 2005). According to Lu et al. (2005), the criteria must also be able to provide specific answers to parent questioning, enable students to understand the dimensions of their learning capabilities, provide students with information that can be used to examine and evaluate their performance, offer feedback on student performance, and enable teachers to determine whether the assessment results truly reflect the response processes of students. Performance assessment often simultaneously involves multiple rating standards, some of which are objective and quantitative (e.g., completion time, quantity of completed work, materials consumed, and error) and some are subjective and qualitative (e.g., the originality and comprehensiveness of the completed work, the proficiency of action, and safety; Ou, 2002).

The application of performance assessment ranges from classroom teaching to large-scale surveys such as those conducted for appraisal purposes. A major concern in these forms of assessments is general quality control (Dunbar, Koretz, & Hoover, 1991). A previous study regarding performance assessment found that errors in the generalisation of performance assessment is primarily affected by the following four factors: (a) the items or activities used in the assessment, (b) the assessors, (c) the situations in which the assessments are conducted, and (d) the unintentional influence of assessees or other people (Jiang, Smith, & Nichols, 1997).

To determine whether a student has mastered a skill, the evaluator must collect performance data on multiple occasions. The number of observations necessary for making decisions can be determined based on the importance of the decision, the amount of time that an observation consumes, and whether the teacher has collected sufficient samples for evaluating student performance and behaviour. Hence, carefully and comprehensively observing all details in a single observation is equally vital for the assessors and assessees. Thus, the assessors can obtain all details by conducting only a minimal number of observations, thereby reducing the cost of assessment. Simultaneously, the assessees can benefit from fair assessments performed based on records that contain all details regarding their performance.

Another question involves assessor selection: Should teachers or professionals act as the assessors or should students perform self-assessment or peer assessment? Regardless, the assessors should have received training on rating.

Performance assessment has long been extensively applied to various fields primarily because this assessment method has advantages that cannot be achieved by conducting paper-based tests. Nevertheless, this method has limitations that must be overcome. Previous studies have shown that the limitations of performance assessment include the following concerns: rating is subjective; the criteria adopted by various assessors are inconsistent; the assessors do not truly understand the connotation of the assessment; the assessors do not follow the standardised assessment processes; and the rating standards are undefined and time- and energy-consuming. To enhance the fairness and objectivity of performance assessment, scholars have developed procedures for implementing performance assessment. Among the steps defined by multiple scholars, the following four steps are indispensable: (a) defining the purpose of assessment, (b) confirming the assessment standards, (c) designing tasks or activities, and (d) selecting a rating or assessment method (Chen & Martin, 2000).

In summation, when performance assessment is applied to assess the operation and production of actual works, the fairness, objectivity, convenience, and timeliness of assessment must be considered to overcome the limitations of this method. These factors were used as the references for developing the questionnaire employed in this study. The following paragraphs list the primary limitations of performance assessment and how AR technology was employed to solve the problems:

Lack of comparability: In conventional standardised tests, the results can be compared against established norms; therefore, result interpretation is specific and clear. By contrast, the results of performance assessment are often affected by the subjective judgments of teachers; additionally, the criteria employed are occasionally confusing, thereby increasing the difficulty involved in comparing and interpreting the assessment results. The process of student performance can be recorded and students and teachers can employ AR technology during the rating process. Thus, the assessors and assessees can appear in real-time situations and serve as the direct references for assessment processes, thereby enhancing the accuracy of assessment.

Limited reliability: The majority of manual assessment methods are subject to the subjective influence of the assessors. Unlike standardised tests, for which computer scoring can be adopted, performance assessment relies on assessor observation and judgment. Consequently, the reliability of the assessors should not be overlooked. The errors in assessor reliability result from the assessors, and a satisfactory rating system can reduce assessor errors (Baker, 1996). We can employ AR to present the processes of work production or the implicit details hidden in the works. Thus, assessment accuracies can be increased substantially and the risks of rating errors resulting from assessor negligence or excessively short observation time can be reduced.

Unsatisfactory economic performance: The amounts of time and money spent on performance assessment are considerably greater than those spent on paper-based tests. Using AR to present the production processes of works can reduce the travel costs that teachers would otherwise spend for conducting on-site observations. In addition, the assessors can watch videos repeatedly to reduce rating errors and the amount of effort that teachers must spend on assessments can be reduced. Furthermore, occasionally teachers must simultaneously observe multiple students, thereby rendering them unable to observe all details within a particular period. Adopting AR can prevent this problem.

Low validity: In performance assessment, ambiguous problem situations are designed to test the high-level thinking abilities of assessees. Nevertheless, the validity of ambiguous problems is difficult to control; consequently, the assessment can be irrelevant to the teaching contents. An AR-integrated system can show in real-time the rating standards and the feedback from teachers or peers; thus, the associated cognition of feedback materials and student works can be enhanced. Therefore, students can more effectively immerse themselves into the teaching situations, thereby improving the validity of assessments.

4. Methods of the Performance Assessment Conducted in Hands-on Performance Courses

The assessment conducted in implementation activities are considered assessments conducted during activities. Performance assessment is often based on observation; thus, it can also be referred to as the work evaluation method (Tsai, 1996).

Establishing a set of criterion to be used in performance assessment enables designers to perform self-assessment during the processes of creation and development and the completion stage and enables assessors to eliminate uncertainty and overcome the complexity involved in the assessment. According to Wolansky (1985), a U.S. scholar specialising in vocational-education studies, teachers must focus on the following concerns when evaluating student works or productions: (a) the performance of the works should accomplish the required purposes; (b) records regarding the production processes of the final products should be made available for assessment purposes; and (c) teachers should provide well-defined explanations or standards regarding the quality of the final products. Wolansky emphasised that a standard table or a criteria table is a convenient and informative tool that enables students to understand the standards for excellent work and the contents of assessment. Yunghans (1981) indicated that the standards for evaluating art works should be (a) the purity and openness of expressions, (b) the problem-solving methods exhibited, (c) the duration of focused attention on production, and (d) the attention to detail in images.

Khattri et al. (1998) argued that the results of performance assessment differ when the methods or systems employed vary. Specifically, the following five characteristics of performance assessments should be considered: (a) the purpose of a performance assessment, (b) the format of assessment, (c) the subject areas being assessed, (d) the levels of students, and (e) the implementation of performance assessments.

Gronlund (1993) categorised performance assessments into the following types based on the extent to which the situations used in the tests are true to reality: (a) paper-and-pencil performance, (b) identification tests, (c) structured performance tests, (d) simulated performance, and (e) work samples. The experiments in this study were conducted using a structured performance test and simulated performance.

In summary, the specific steps of rating criterion design are (a) teachers must first determine the target of assessment, be it the process, the result, or both; (b) subsequently, teachers must identify the contents or scopes of observation, list the focuses of the observation and assessment, and explain the significance of the criteria adopted; (c) teachers can discuss the rating criteria with students and ensure that the students truly understand the connotations of the criteria; the rating criteria can also be established by students or jointly by students and teachers; and (d) before conducting a performance assessment, teachers must carefully examine the items regarding detailed behaviour and apply necessary revisions.

5. Performance Assessment with Mobile Augmented Reality

AR enables users to visualise real environments in a real world with the digital information overlaid on real environments (objects or locations), thereby improving user experiences (Berryman, 2012). The combination of additional information and real situations can enhance the senses of reality and presence for people. The theoretical basis for the mobile AR system that integrates human–computer-context interactions is situated cognition. The fundamental argument of the theory is that knowledge acquisition and learning occurs after people interact with situations, which include social environments such as people and social culture, and physical environments such as the contexts formed by scenes and artefacts (Brown, Collins, & Dugid, 1989; Greeno, Collins, & Resnick, 1996).

Applications of AR and high-tech products to situated teaching activities are lacking because the attention that users direct toward additional information and real scenes is difficult to balance. In addition, human–computer-context interactions are difficult to achieve. Participants may focus excessively on human–computer interactions and overlook human–context (objects in scenes and information contexts) interactions, which are more crucial than human–computer interactions in real situations. Therefore, the link between additional information and real environments should be emphasised in the virtual contents presented in AR (Klopfer & Squire, 2008; Chang et al., 2014; Zhang et al., 2014).

In informal learning, the application of mobile devices has recently attracted an increasing amount of attention (Semper & Spasojevic, 2002; Kwak, 2004; Cabrera et al., 2005; Chang et al., 2006; Sung et al., 2010a). However, studies regarding the application of AR navigation are scant (Barber et al., 2001; Sparacino, 2002; Damala et al., 2007; Damala et al., 2008; Portalés et al., 2009).

The mixed reality spectrum (Fig. 2-2) developed by Milgram and Kishino (1994) offers a valuable basis for exploring the integration of reality and virtual reality. AR is situated on the spectrum between virtual and real environments. Based on the definition of mixed reality, Milgram and Kishino developed the linear spectrum, showing the transition from real environments to virtual environments. With real environments on the left end of the spectrum and virtual environment on the right, AR is located toward the left end of the spectrum, indicating that the main subject in AR is real objects and that virtual objects are additional and supplementary. When AR is applied to spaces, implicit spatial information is transformed to explicit spatial information by employing technologies that incorporate virtual objects with the real world. Hence, additional values and meaning are added to spaces.

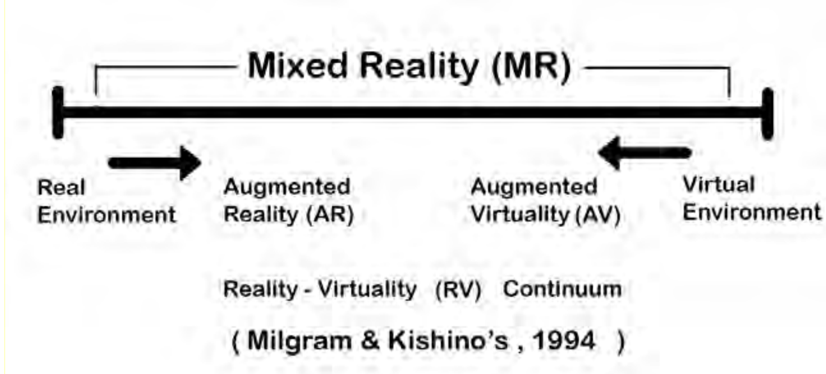


Figure 0. Reality-Virtual Continuum

AR is a technology that incorporates virtual and real objects in real environments (Azuma et al., 2001). AR enables users to visualise real environments in a real world with digital information overlaid on actual environments (objects or locations), thereby improving user experience (Berryman, 2012). Initially, most studies used head-mounted displays to present the results of virtual–real environment integration. Azuma (1997) defined three criteria for AR: (a) the combination of virtual and real environments, (b) real-time interaction, and (c) 3D referencing. Scholars generally agree that AR can be used to enhance the experience that users have when interacting with real environments. In addition, virtual information enables users to obtain information that otherwise cannot be directly acquired from the real world. Because of this feature, AR is considered an effective tool that users can employ to achieve objectives in the real world (Azuma, 1997).

Technically, AR presentation can be divided into the following two types: marker-based and marker-less identification. Specifically, marker-based identification operates based on the principle of quick response codes, which are 2D bar codes or dot matrices in square icons. The markers are locked and read using the cameras on mobile devices and identification software. Subsequently, the interaction is activated using 3D objects or videos. By contrast, marker-less identification is based on the global positioning system. Users can use mobile devices to locate objects that interest them and floating markers or chat boxes are shown to display information through the cameras installed in devices.

Since the 1990s, AR has been applied in various fields, including geography (Vlahakis et al., 2002; Portalés, Lerma, & Pérez, 2009; Priestnall, 2009), linguistics (Liu, 2009), social sciences (Hedley et al., 2002; Mathews, 2010; McCall et al., 2011), mathematical sciences (Wang, 2007; Yim & Seong, 2010), natural sciences (Klopfer & Squire, 2008; Liu, Tan, & Chu, 2009; Zhang et al., 2014), biomedicine (Vilkoniene, 2009; Strickland et al., 2011), arts and humanities (Shen, Ong, & Nee, 2010; Chang et al., 2014), leisure and recreation (Portalés et al., 2010; Wang & Chen, 2009), and advertising and marketing (Moltenbrey, 2011).

Barber et al. (2001) indicated that using smartphones to display additional information and placing the screens next to student works is essentially integrating virtual and real environments into the same view. Thus, the number of times users switch between the exhibited works and description plaques can be reduced and the number of searches users must make can also be minimised. In a study conducted by Dunleavy et al. (2008), the students engaged in role-playing tasks by using AR. Specifically, the students walked around campus while the cameras on the mobile devices they were holding displayed digital objects and virtual characters that were overlaid onto real spaces. Video, audio, and text files were used to provide clues and challenges for narration, navigation, and cooperation. Thus, the learning objectives for subjects such as math, language arts, and scientific literacy at junior and senior high schools can be accomplished.

Damala et al. (2008) agreed that the integration of virtual and real environments revolutionised the interaction between people and objects in an unanticipated manner. The tiny screens on the devices can represent complete environmental spaces and facilitate establishing a close relationship between the

appreciated works and additional information. However, AR technology requires improvement when applied to learning activities (Billinghurst et al., 2003; Dunleavy et al., 2008; Wang & Chen, 2009; McCall et al., 2011). For example, considering the coexistence of virtual and real environments in AR, the additional information presented in AR may be designed to attract participant attention so that the participants can see the information. Consequently, the participants may excessively focus on the contents shown in the AR system, particularly on the additional information, and ignore the actual environment and surroundings (Billinghurst et al., 2003; Dunleavy et al., 2008; Wang & Chen, 2009).
In summation, to design a performance assessment learning system that achieves human–computer-context interactions, we employed AR technology to develop a performance assessment system that enables peer assessment. In this system, the criteria based on which learners produce their works or assessors evaluate the works are predefined. Thus, students can evaluate their own works or peers' works based on sufficient information, thereby developing strong learning motivations and achieving great efficacy. In addition, teachers or assessors can spend comparatively less time and simultaneously evaluate assessees' works accurately and fairly.

6. System Realization and Illustrative Example

6.1 System Architecture

In the field of education, numerous situations cannot be experienced or represented in the classroom setting. AR is the most appropriate technology for incorporating or adjusting students' learning experience based on specific needs. AR is defined as a real-world environment whose elements are built upon computer-generated sensory input such as sound, video, graphics or GPS data. In this study, AR allows students to see virtual objects about peers' works or contents in a real world environment with the aid of camera during the assessment process. The overall framework of the use of mobile AR technique in performance assessment is described in figure 1.

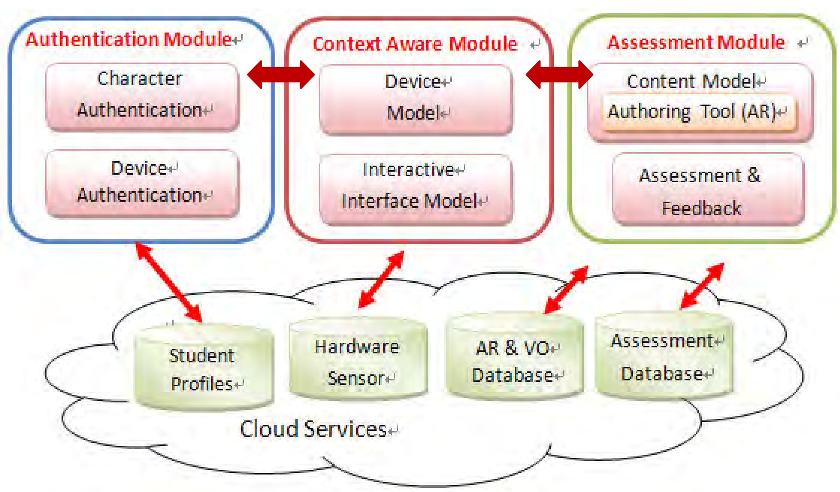


Figure 1. The architecture of the mobile AR technique

The entire processes of learning and assessment can be divided into the following three modules: the authentication module, the context-aware module, and the interactive assessment module. In addition, the process is supported by four databases on cloud servers, which are student profiles, a hardware sensor, the AR and virtual object database, and the assessment database. The authentication module enables authorised people to obtain appropriate information for completing corresponding tasks. The context-aware module enables assessors to employ appropriate device functions for accessing suitable information for performing rating. In the context-aware module, mobile devices list appropriate learning contents after detecting student locations and collecting onsite information. Thus, learners can select appropriate learning materials from the content model. The content model retrieves appropriate materials from the virtual object database before providing them to assessors. Subsequently, learners can use the authoring tool provided by the system to establish an AR marker, work descriptions, and an AR context object. Once all steps are completed and the information is uploaded to the system, appropriate virtual information is used through AR technology to overlay images onto corresponding objects in the real world. Thus, assessors can rate the works conveniently and accurately. Through the AR work presentation technology employed in the system, the assessment module enables assessors to conveniently and directly observe the works of peers. Hence, assessors can provide feedback for the peers they evaluate. In addition, the system can be employed to develop a work-specific exhibition situation for peer references, thereby enabling peer assessors to provide feedback. Additionally, teachers can use the AR performance assessment system to understand the peer assessment performed by students before providing feedback for the assessors and assessees. More importantly, teachers can integrate previous cases to develop new teaching situations that are highly appropriate and inspiring.

6.2 Walk-through Illustrative Example

The methods for conducting performance assessment are diverse, including observation, document records, and real-time performance. The methods adopted in this study were real-time performance and peer assessment. Peers who possessed similar knowledge levels observed and learned from each other before offering recommendations. Specifically, a class of 50 sophomore students at the culinary department of a technical institute were recruited as the participants of this study. A performance assessment experiment was conducted during a training course for cooking licenses in Western cuisine. The students were divided into groups of five, obtaining a total of 10 groups. The group members divided the labour between themselves. The students were randomly assigned to the groups without considering sex or cooking skills. During class, the teacher designated an item from the licensing examination as a task. The teacher demonstrated the cooking procedure once, after which the completed set was recorded and used as an item marker in the AR performance assessment system. Before the students began the performance, they used mobile devices (tablet personal computers) to photograph the sample. Subsequently, the system displayed real-time information (learning mode) that corresponds

Figure 2. The process of the mobile augmented reality p assessment system

to the dish onto the dish image, such as the ingredients that should be prepared and the steps of cooking. Thus, students can follow the instructions during the performance and record the process of cooking by using mobile devices. Subsequently, the videos were uploaded to cloud servers and arranged by the names of the dishes. After all the dishes were completed, the system integrated all the data and was prepared for peer assessment. The system enables assessors to review the records of assessed dishes for reference. After the assessment mode began, the system listed the content and assessment criteria of the set for the assessees rather than listing information regarding the dishes after the samples were scanned. The assessees prepared the ingredients for cooking the dishes and had every step recorded before the videos were uploaded to cloud servers. Subsequently, the system integrated the information for the teachers to provide ratings and feedback. The procedure of the experiment is shown in Fig. 2.

During assessment, the AR performance assessment system identifies each dish and lists the contents when the assessors use the cameras on their mobile devices to photograph the sets completed by the assessees. When an assessor selects the name of a particular dish, the video showing the cooking process is immediately shown on the screen. In addition, the assessment criteria are displayed simultaneously, enabling the assessors to perform the assessment intuitively and clearly. Because the entire cooking procedures were videotaped, the assessors were able to observe all the details that interested them. Thus, the assessors did not miss crucial details as they otherwise would when simultaneously observing several groups of students. Additionally, the assessment criteria adopted are consistent because they are shown in real-time. Hence, the errors in performance assessment can be minimised.

Furthermore, the assessors can provide real-time feedback and recommendations during assessment. The feedback can be uploaded to cloud servers immediately following assessment. Thus, the students can immediately review and share the feedback and recommendations regarding their works and further discuss among themselves by using the system. Real-time sharing and the real-time display of assessment criteria enable students to immediately understand the advantages and disadvantages of their works and to use the feedback to improve their works. Thus, the learning objectives were achieved.

7. Conclusion

This paper presents the novel framework of an enhanced performance assessment system complemented by the use of smart and mobile devices. Integrating the AR technology overcame some of the limitations of conventional performance assessment systems, such as the implementation method, excessively high costs, and substantial errors. In this framework, the AR technology enables students to observe how their peers completed their works by displaying videos of the cooking process over the completed dishes. During the assessment, students can determine whether their peers followed the instructions correctly by comparing the performance against the assessment criteria. By doing so, students can discover their own inadequacies or learn from other people’s methods. In addition, the system provides each student stable and convenient information and digital content based on environmental parameters or the identification of particular objects. Thus, students can learn while engaging in activities based on which their performance is assessed. Students can obtain appropriate learning information by using mobile devices to photograph and identify target objects at appropriate moments and particular locations. The novel framework developed in this study, in which the AR technology was integrated, enables students to use various methods to observe the cooking processes and completed works of their peers. Simultaneously, the students can receive real-time feedback and recommendations regarding their own works. Hence, the barrier resulting from conflicting opinions between students can be eliminated, and students’ understanding of each other’s opinions can be enhanced. Thus, the accuracy of the results of performance assessment can be improved. From the perspective of cognitive psychology, showing assessment criteria and feedback in real-time situations during assessment enables students to develop strong impressions of the feedback. Therefore, the students are highly capable of incorporating the feedback into their future work to achieve improvement and growth. The novel AR-integrated framework used in this experiment is almost complete. However, additional work is necessary. For example, studies can be conducted using an experiment sample that is larger than that used in this study, a large number of performance tests in a classroom setting, and an enhanced system.

References

Lin, C. P., 1999, Teaching assessment in art courses: Using the Department of Art and Craft Education at Taipei Municipal Teachers College as an example. *Journal of Taipei Municipal Teachers College*, 30, 451-474.

Ou, C. H. 2002. *Educational Testing and Evaluation*. Taipei: Psychological Publishing Co., Ltd.

Lu, J. H., Chen, M. Y., & Wu, Y. Y., 2005, An Action Research of a Performance Assessment Design on Assessing Students' Ability of Conducting Experiments. *Contemporary Educational Research Quarterly*, 13, 4, 95-132.

Chang, L. L., 2002a, The promised land of assessment reform, virtual or real: The operation and criteria for performance assessment. *Journal of Education Research*, 93, 76-86.

Chang, L. L., 2002b, Establishing the validity of performance assessment based on the significance of scores. *Journal of Education Research*, 98, 76-86.

Tsai, D. C., 1996, Using group decision-making techniques to assess designs. Proceedings of the 2nd Conference on the Research and Practice of Design and Management. pp. 1-8.

Yen, J., 1997, Applying structured assessment methods to the assessment of visual communication designs. Proceedings of Visual Communication Design: Education, Culture, and Technology. Taipei: AsiaPac Books. pp. 231-236.

Li, D. W., 1986, *Measurement and Evaluation in Technical and Vocational Education*. Taipei: San Min Books Co., Ltd.

Beywl W. and Speer S., 2004, Data- and literature-based reflections on western European evaluation standards and practices, New Directions for Evaluation, No. 104, pp. 43-54.

Wiggins, G., 1993, Assessment: Authenticity, context, and validity, Phi Delta Kappa, pp. 200-214.

Airasian, P. W., 1996, Assessment in the classroom, McGraw-Hill, New York.

Fitzpatrick, R. and Morrison, E. J., 1971, Performance and product evaluation, In Thorndike, R. L. (Ed.) , Educational measurement, American Council on Education, Washington, DC, pp. 237-270.

Aschbacher, P. R., 1991, Performance assessment: state activity, interest, and concerns, Applied Measurement in Education, Vol. 4, No. 4, pp. 275-288.

Herman, J. L., Aschbacher, P. R., and Winters, L., 1990, Issues in developing alternative assessments, Paper presented at the annual meeting of the California Educational Research Association, Chicago.

Dorn, C., 1999, Mind in art: Cognitive foundations in art education, Mahway, Lawrence Erubaum Associates Publishers, New York.

Chen, Y. F. and Martin, M. A., 2000, Using performance assessment and portfolio assessment together in the elementary classroom, Reading Improvement, No. 37, pp.32-38.

Wolansky, W., 1985, Evaluating student performance in vocational education, The Iowa State University Press, Ames, IA.

Yunghans, M., 1981, A pull-out program for gifted elementary students, School Arts, Vol. 80, No. 8, pp. 50-51.

Khattri, N., Reeve, A. L., and Kane, M. B., 1998, Principles and practices of performance assessment, Lawrence Erlbaum, Mahwah, NJ.

Baker, E. L., Abedi, J. (1996). Dimensionality and Generalizability of Domain-Independent Performance Assessments. Journal of Education Research, 89(4), 197-205

Roeber E. D. (1990). Performance assessment: A national perspective. Policy Briefs Numbers 10 & 11. Special Double Issue. (ERIC Document Reproduction Service No. ED 370980)

Wiggins, G. (1993). Assessment: Authenticity, context, and validity. Phi Delta Kappan, November, 200-214.

Gronlund, N.E. (1993). How to make achievement tests and assessments (5nd ed.).Boston: Allyn & Bacon.

Linn, R. L., Baker, E. L., & Dunbar, S. B.（1991）. Complex, performance-based assessment: Expectations and validation criteria. Educational Research, 20(8), 15-21.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational Measurement, 13 –104.New York: Macmillan.

Messick, S. (1992). Validity of test interpretation and use. In M. C. Alkin (Ed.), Encyclopedia of educational research (6th ed., Vol. 4, pp. 1487-1495). New York: Macmillan.

Dunbar, S. B., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. Applied Measurement in Education, 4, 289-304

Jiang, Y. H., Smith, P., & Nichols, P. (1997). Error sources influencing performance assessment reliability or generaliziability: A meta-analysis. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Popham, W. J., 1995, Classroom assessment: What teachers need to know, Allyn and Bacon, Boston.

Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Investigating effects of mobile learning in familiar authentic environment on learning achievement and cognitive load

Rustam SHADIEV[a], Wu-Yuin HWANG[b], Yueh-Min HUANG[a] & Tzu-Yu Liu[c]
[a]*Department of Engineering Science, National Cheng Kung University, Taiwan*
[b]*Graduate Institute of Network Learning Technology, National Central University, Taiwan*
[c]*Bei-Zheng Junior High School, Taiwan*

Abstract: This study designed English as foreign language learning activities in a familiar authentic environment with mobile technological support. Students learned at school and then applied new knowledge to solve daily life problems by taking pictures of learning objects, describing them, and sharing their homework with peers. This study carried out two experiments in which 59 junior high school students participated. One class with 28 students served as the control group in the experiment 1 and as the experimental group in the experiment 2. The other class with 31 students served as the experimental group in the experiment 1 and as the control group in the experiment 2. Control students studied and completed each learning activity with traditional textbooks while the experimental group studied with an electronic textbook and a learning system installed in the tablet PCs. This study aimed to investigate effects of mobile learning in familiar authentic environment on learning achievement and cognitive load. This study discusses results, research findings, and implications along with conclusions and several suggestions for future development and research.

Keywords: Mobile learning, familiar authentic environment, learning achievement, cognitive load

1. Introduction

The rapid advancement of information and communication technology created new opportunities in the design of the instruction (Hwang et al., 2014; Shadiev, Hwang, Huang, & Yang, 2014). According to the related studies, for instance, mobile technology provides such advantages as learning anywhere and anytime, and how one likes (Hwang, Chen, Shadiev, Huang, & Chen, 2012). Chang, Tseng, and Tseng (2011) suggested that mobile technology creates an authentic learning environment in a real-world context in which learners adapt learning content to the context they find themselves in. Furthermore, mobile technology can be used as a cognitive tool to aid the learning by decreasing the cognitive load (Hwang, Wu, Zhuang, & Huang, 2013). Therefore, mobile-assisted learning has been successfully implemented in many studies to manage cognitive load and facilitate learning; for example, in a social science course (Hwang et al., 2013) or foreign language learning (Chen, Hsieh, & Kinshuk, 2008; Chang et al., 2011).

The literature review of this study revealed several issues that they need to be addressed appropriately. First, several studies were carried out to reduce cognitive load in English learning; however, focus of most of them was limited to vocabulary learning and reading and listening comprehension (Chang et al., 2011; Chen et al., 2008). Therefore, cognitive load associated with speaking and writing skills, was overlooked in related studies. Second, familiar context that creates advantages in comprehension, recall, and cognitive load and which can be found in authentic environment, was not considered in most related studies. An authentic learning environment was created in a local temple (Hwang et al., 2013), in a zoo (Chang et al., 2011), or a classroom Chen et al. (2008) by employing mobile technology but learning context there was not so familiar to students. Third, in most previous related studies, active application of knowledge did not get sufficient attention (Chang et al., 2011; Chen et al., 2008; Hwang et al., 2013). In these studies, a mobile learning system provided learning content, guided students to the learning targets, displayed questions, and students were asked to answer them. Such learning process facilitates only low level cognitive processes (i.e. information recall).

To address these issues, this study designed various learning activities supported by a mobile technology. In learning activities, students could learn in class and then freely apply what they learned to solve daily life problems in a familiar authentic environment (e.g. their home or a local convenience store). In this study, students took pictures of learning objects in a familiar authentic environment and described them by using English as a foreign language. This study aimed to investigate how mobile learning in familiar authentic environment effects learning achievement and cognitive load.

2. Cognitive load

Mayer and Moreno (2003) and Paas et al. (2003) argued that cognitive load is a central consideration in the design of the instruction. The reason is that a learner has limited cognitive capacity to accommodate demands imposed by learning task, due to working memory (Paas et al., 2003). According to Mayer and Moreno (2003) and Paas et al. (2003), learning performance can be negatively affected when cognitive load exceeds the limit of cognitive capacity. Three types of cognitive load were distinguished in related literature: intrinsic, extraneous, and germane (Brunken, Plass, & Leutner, 2003; Sweller, Van Merrienboer, & Paas, 1998). Intrinsic load is determined by the inherent nature of learning material, learners' expertise and an interaction between them; that is, the amount of information units that a learner needs to hold in working memory to comprehend the information. Extraneous load is referred to cognitive load caused by the format and manner in which information is presented and by working memory requirements required to do the instructional activities. Extraneous load can be imposed by improper instructional design. Germane load is determined by learners' effort to process and comprehend the learning material. This load is also associated with motivation and interest. Germane load is induced by appropriate instructional design and it can enhance learning.

3. Managing cognitive load in authentic learning environment

It is suggested that learning activities in authentic environment are likely to facilitate students' cognitive activity and conceptual change (Shadiev, Hwang, & Huang, in press). Furthermore, in authentic environment, cognitive tools, such as mobile multimedia applications, can aid the learning by decreasing the cognitive load. Hwang et al. (2013) introduced a mobile learning system to aid sixth grade students' local culture learning during field trip in the social science course. With the system, students accessed physical and virtual resources in authentic environment; the system presented the learning tasks, guided students to visit the real-world learning targets for exploration, and provided them with supplementary materials via the mobile devices. The effects of this approach on students' cognitive load and learning achievements were investigated. Results showed that students who learned with technological approach had better learning achievement and less cognitive load than those who learned with the traditional approach. Hwang et al. (2013) suggested that the mobile learning approach has positive effects on students' local culture learning.

Chen et al. (2008) explored how short-term memory and content representation type affect language learning in mobile learning environment. EFL university students were divided into verbal and visual short-term memory group and all of them learned 24 English words (with written annotation and pictorial annotation) delivered by Short Message Service or Multimedia Message Service to their mobile devices. Chen et al. (2008) found that providing vocabulary with pictorial annotation is helpful to learners with lower verbal and higher visual ability. The reason is that these learners find it easier to learn words presented in a visual rather than in a verbal form. However, providing vocabulary with both written and pictorial annotation can also help learners with both high verbal and high visual abilities. Chen et al. (2008) concluded that providing the basic learning material is more helpful to learners with low verbal and visual abilities as too much information may produce a high cognitive load and shorten concentration time.

Chang et al. (2011) examined the effects of English proficiency (low vs. high) and material presentation mode (single channel vs. dual channel) on English listening comprehension and cognitive load in a ubiquitous learning environment. In an experimental learning activity, university students studied zoo animals by using PDA. The system guided students to target animal areas and then displayed related material (i.e. text) and played audio guide (spoken messages). Students in a single channel group learned through spoken messages only, whereas students in a dual channel group learned by text and spoken messages. Results of the study revealed that high and low English proficiency learners in the dual channel group had better English listening comprehension than learners in the single channel group. Low English proficiency learners in the dual channel group possessed significantly lower extraneous load than those in the single channel group. Chang et al. (2011) concluded that dual channel presentation mode leads to an increased depth of information processing and different input modes reinforce one another.

4. Method

Two experiments were carried out in this study. A total of 59 junior high school students participated in two experiments. One class with 28 students served as the control group in the experiment 1 and as the experimental group in the experiment 2. The other class with 31 students served as the experimental group in the experiment 1 and as the control group in the experiment 2. Most students in both groups were thirteen years old with four to six years' experience of using computers and less than one to three years' experience of using tablet PCs.

The experimental procedure of this study is as follows. First, a pre-test was conducted before the experiment started. Then both classes had the same amount of hours of English course: three one-hour lessons a week. After lessons, students participated in learning activities to practice their skills and applied new knowledge in daily life situations. Lessons and learning activities taught in the two classes were guided by the same instructor and shared the same learning content. However, the control group studied and completed learning activities with traditional textbooks while the experimental group studied with an electronic textbook and learning system installed in the tablet PCs. Learning activities were three tasks; each lasted for two weeks. In the beginning of the experiment, every experimental student received tablet PC and students were taught how to use the e-textbook and system by the instructor. A post-test and cognitive load questionnaire survey took place in the end of the experiment with all students. Finally, interviews with experimental students were carried out one week later after the experiment.

This study designed learning activities that were focused on learning at school and applying knowledge learned in authentic environment outside of school with a wide range of daily life situations (e.g. at local convenience store or supermarket). Three topics from the textbook were chosen to design learning activities: (1) "Where Are You From?" (2) "You School Is Very Big," and (3) "Be Quiet and Sit Down, Please" for the first experiment and (1) "Which do you like – Healthy diet," (2) "How much / many do we need," and (3) "We were in different classes" for the second experiment. Learning activities were three tasks, and each corresponded to its topic. In each task, students were asked to take a picture of a learning object (e.g. a sign for Topic 3 of the first experiment or a meal for Topic 1 of the second experiment) and then to introduce and to describe it by using at least 6-10 sentences.

This study developed the learning system to support students to carry out the learning activity tasks. The following four main functions were designed in the system: (1) Annotating. Students could annotate important parts of learning material on tablet PCs. Besides, students could take photos and attach them to an annotation. (2) Recording. When students spoke out to describe a learning object, they could record their own voice and play it afterwards. Besides, students could record the instructor's lectures and listen to them. (3) Assistance. Students could get assistance from the system, such as (a) read text out loud (Text-to-Speech Recognition), (b) translate unfamiliar vocabulary and sentences (Translation), and (c) list of words in alphabetical order with their meaning and translation (Dictionary). (4) Sharing. Students could share their own annotations, photos, and audio recorded files with peers.

The following are research tools that were employed by this study. Students' prior knowledge was evaluated by a pre-test and students' learning achievement was measured by a post-test in two experiments. The items of the pre-test and post-test for both experiments were similar in structure but different in content. Thirty items were included in each test: (1) Match English word with the correct Chinese meaning – eight items; (2) Write down the Chinese meaning of English word – six items; (3) Fill in the blank – ten items; (4) Write down: a) a question based on a sentence; b) negative sentence from given one; and c) translation of a sentence – 5 items; and (5) Write down about yourself when you were at the first grade of the elementary school, then write about yourself at the moment, and finally, compare the difference between when you were at elementary school and now – 1 item. Students' answers to the tests were scored on a 100-point scale.

Cognitive load questionnaire (Huang, Huang, Liu, & Tsai, 2013) was developed with seven items: (1) Learning these materials was easy, (2) Learning these materials did not require a lot of mental effort, (3) Completing learning activities was easy, (4) Completing learning activities did not require a lot of mental effort, (5) I was concentrated during learning, (6) My mood was joyful during learning, (7) My frustration was low during learning. Items 1 and 2 of the questionnaire measured intrinsic cognitive load, items 3 and 4 measured extraneous cognitive load, and items 5, 6 and 7 measured germane cognitive load. All 59 students were asked to respond to the questionnaire and 59 valid answer sheets were obtained. Responses to the items were scored using a five-point Likert scale, anchored by the end-points “strongly agree” (1) and “strongly disagree” (5). The internal consistency of the survey was tested by employing Cronbach α; the values exceeded 0.80 demonstrating satisfied reliability of the items.

One-on-one semi-structured interviews were conducted with randomly selected ten experimental students from each experiment. Interviews aimed to explore students’ learning experiences with the system and insights of their perceptions toward cognitive load. Each interview lasted for 20 minutes and students were asked the following questions: 1) Please describe your learning experience with the system; 2) Was the system useful for learning? If yes, please explain why.

5. Results and discussion

5.1 Effects of mobile learning in authentic environment on learning achievement

This study employed analysis of covariance to measure the difference in the learning achievement of students in the control and experimental students on the post-test with the pre-test as covariate. In the experiment 1, a significant difference was observed between the control (M=53.50, SD=13.21) and experimental group (M=65.45, SD=18.59) on the post-test, $F(1,56)=16.709$, $p=0.000$, partial eta-squared=0.236. In the experiment 2, the experimental group (M=70.32, SD=17.01) outperformed the control group (M=58.30, SD=22.67) on the post-test, $F(1, 56)= 20.345$, $p=0.000$, partial eta-squared=0.270.

Students were asked to introduce and describe some learning objects (e.g. signs and rules in the convenience store of their local community). The experimental students in two experiments completed assigned tasks better than the control students. This finding may suggest that learning activities in familiar authentic context supported by the system could facilitate students learning.

This study interviewed experimental students to provide subjective evidence that may support abovementioned objective evidence. In the interviews, students mentioned that learning activities could be completed more efficiently if using the system instead of traditional approach. Furthermore, the system enabled more effective practice of EFL skills. First, students took pictures of learning objects and recorded their own voice when describing learning objects. Students were fond of reviewing pictures and listening to their own recorded files, and if content quality of photos and recorded files was not satisfactory (e.g. mistakes in pronouncing some words), students would want to improve it. According to students, such learning behavior led to more frequent language practice as well as to better quality of language output. Similar reasons to using multimedia tools for language practice were reported in other research (Harmer, 2007; Hwang, Shadiev, & Huang, 2011; Hwang & Shadiev, 2014;). For example, students in the study of Hwang, Huang, Shadiev, Wu, and Chen (in press) and Hwang et al. (2011) took advantage of the technology in the same way of practicing the target language repeatedly and regularly. In the study of Harmer (2007), after students recorded their speeches, they listened to recordings, evaluated language performance, and monitored how much progress made. However, in contrast to other related research, this study focused not on learning the basic knowledge at school only, but the application of learned knowledge to solve wide range of real life problems in familiar authentic environment.

Second, students shared recorded files with peers. In this way, students could listen to peers’ recorded files (i.e. usually to those who study hard and perform well) to get inspirational ideas to complete their own assignments or to learn how peers accomplished assignments and to improve their own homework. Students could exchange meaningful comments through sharing. That is, some students gave reflective comments and suggestions to a peer who did not complete homework correctly. Besides, students’ comments were useful to revise or improve homework. Students highly thought of sharing mechanism of the system as they were able to learn from others, and then to locate and revise their own mistakes in homework. Hwang et al. (in press) and Hwang et al. (2011) argued that, with multimedia aids, students access more diverse learning objects and this may increase the richness of their language experience. They further suggested that sharing multimedia learning content with others not only increases practice opportunities but engages students in EFL contexts and allows their deeper reflection on learning content, discussion and collaboration.

Finally, students stated that the built-in dictionary was very handy when they needed to translate some unfamiliar vocabularies when completing assignments outside of school or at home. In this case, a dictionary could help to translate these words. Moreover, with a dictionary, students could find multiple meaning of a word and how it can be used in different contexts. Hulstijn and Laufer (2001) argued that the use of a dictionary positively affects vocabulary learning. Students look up target words in the dictionary during the reading session in order to find word meanings and to understand the main idea of texts. According to Hulstijn and Laufer (2001), those students who read foreign language texts and use a dictionary can understand texts better and remember more word meanings.

5.2 Effects of mobile learning in authentic environment on cognitive load

This study examined whether designed learning activities supported by the system brings extra cognitive load on students during learning. Therefore, cognitive load of the experimental and control students was measured and then compared by employing independent samples test. The means and standard deviations from the assessment with respect to seven items of the questionnaire and results of t-test are presented in Table 1. According to the table, the control students had higher cognitive load with regard to all seven items ($p<0.05$) than the experimental students. This finding suggests that learning activities supported by the system enabled students have less cognitive load compared to traditional learning setting.

Table 3: Cognitive load assessment and t-test results

Item	The experiment #1							The experiment #2						
	Control		Experimental		*t*	Sig. 2-tailed	MD	Control		Experimental		*t*	Sig. 2-tailed	MD
	M	SD	M	SD				M	SD	M	SD			
1.	1.90	0.94	1.48	0.51	2.097	0.042	0.413	2.10	0.76	1.50	0.64	3.247	0.002	0.600
2.	2.10	0.86	1.58	0.50	2.852	0.007	0.523	2.33	0.96	1.68	0.67	2.995	0.004	0.655
3.	2.07	0.84	1.61	0.50	2.535	0.015	0.456	2.27	0.91	1.61	0.74	3.025	0.004	0.660
4.	2.14	0.88	1.55	0.51	3.166	0.003	0.590	2.37	0.89	1.68	0.72	3.218	0.002	0.688
5.	2.21	0.82	1.77	0.43	2.544	0.015	0.433	2.50	0.68	1.68	0.72	4.452	0.000	0.821
6.	2.10	0.90	1.61	0.50	2.591	0.013	0.491	2.60	0.81	1.46	0.58	6.095	0.000	1.136
7.	2.03	0.87	1.58	0.50	2.464	0.018	0.454	2.37	0.72	1.54	0.64	4.647	0.000	0.831

This study further explored the changes of cognitive load in two groups of students across different approaches (learning activity with and without technological support). For example, the group A was the control group in the experiment 1 and used traditional method to complete the tasks; then, it became the experimental group in the experiment 2 and used technological support to complete the tasks. On the other hand, the group B was the experimental group in the experiment 1and it became the control group in the experiment 2. So this study investigated how students' cognitive load would change when approaches changed. Independent samples test was employed for such analysis. According to the results, there were significant differences in students' cognitive load between two different approaches. Compared to the control students, those who were in the experimental groups had lower load with respect to all seven items ($p=0.005$). This result confirms the previous finding of this study that learning activities supported by the system helped to reduce experimental students' cognitive load compared to control students.

One reason that may explain these two findings may be due to the nature of the learning material and activity for two groups. In this study, the learning materials and activities for the control and experimental students were identical apart from that the learning materials for the experimental students was in electronic form and they could take advantage of the system's functions mentioned earlier, such as (1) Annotating, (2) Recording, (3) Assistance, and (4) Sharing, to complete learning tasks.

The system functions helped to keep intrinsic load from being overloaded while experimental students learned with the electronic textbook. Students could annotate important parts of learning material (e.g. key concepts) by highlighting them or adding textual and multimedia explanation (e.g. a concept meaning and examples of its application in various contexts). Afterwards, these annotations helped students to find important concepts easily, to recall them, and to complete homework or to prepare for the exams. It is important to note that learning material and relevant annotations (i.e. text, photo, and audio) were located on the same screen. Students anchored their annotations to learning material which built a connection between the annotation and the learning material and gave students a clear picture of the whole learning scenario with an appropriate explanation of it. Mayer and Moreno (2003) called this form of presentation as integrated presentation; with such approach, learners would be more devoted to essential information processing, that is, more cognitive capacity would be activated. Apart from learning from the electronic textbooks, students could learn learning materials from peers' annotations. Studying shared annotations including photos, texts, and audios helped students to enhance their understanding of learning material, to get new ideas and inspiration, and to improve their own homework. Students could also get assistance from the system as Dictionary, translation or Text-to-Speech Recognition to find translation and correct pronunciation of unfamiliar vocabularies or sentences. If students needed to recall some important concepts taught by the instructor from previous classes, they could listen to recorded lectures. Related literature suggests that intrinsic load lies in the nature of learning material, learners' expertise and an interaction between them (Brunken et al., 2003; Sweller et al., 1998). It is argued that intrinsic load represents the amount of different types of information that students need to consider in acquiring new knowledge, i.e. how much information the working memory needs to deal with at the same time (Mayer and Moreno, 2003; Paas et al., 2003). Hwang et al. (2013) suggested that intrinsic cognitive load can be affected by the instructional learning material and students will be cognitively overloaded if the materials are poorly structured, difficult to read, or too complex. Based on the above-mentioned result and considering the fact that students in two groups learned with the same learning materials and tasks, this study suggests that the intrinsic cognitive load resulted from the method students used to access and to process the information. The interviews with students support this finding. Students mentioned in interviews that, compared to traditional approach, it was easier to learn with electronic textbooks and the system. Based on the abovementioned findings, this study concludes that, with the technological support, experimental students' intrinsic load was lower; it was easy for students, and it did not require a lot of mental effort to learn learning material.

It is suggested that extraneous load can be caused by improper instructional design. Thus, in order to reduce extraneous load, instructors need to organize, present and carry out learning information and activities appropriately. In this study, functions of the system helped to reduce extraneous load while students participated in learning activities. When they exposed themselves to the authentic learning environment outside of the school, taking pictures of learning objects and describing them with text or voice annotations worked like gathering their thoughts and then transforming them into artefacts. Hollan, Hutchins, and Kirsh (2000) called such transformation of thoughts as distributed cognition. Hollan and his colleagues argued that knowledge and cognition are not confined to an individual but distributed by placing memories, facts, or knowledge on the objects, individuals, and tools in the learning environment as a set of representations. Lu, Lai, and Law (2010) argued that technology plays an important role to handle intellectual tasks and to ease individual cognitive load. Later, when students were at home, they could have a more tranquil environment and study created artefacts (i.e. pictures and their textual and audio descriptions). These artefacts would help students to easily recall details of learning objects from authentic learning environment, to find out what they missed while completing the tasks, and what else can be improved in their homework. Otherwise, students had to hold a mental representation of the context in working memory over a period of time which was called representational holding (Mayer & Moreno, 2003). According to Mayer and Moreno (2003), when students attempt to engage in both information processing (i.e., selecting, organizing, and integrating learning material) and representational holding, cognitive overload occurs. Our finding is supported by the interview with students. Students claimed in interviews that, compared to traditional method, it was easier to participate in learning activities with electronic textbooks and the system. Students also mentioned that a familiar authentic environment helped them to recall some important vocabularies; the context was related to students' background and previous experiences and also very relevant to learning tasks. Therefore, this study may conclude that with technological support, experimental students' extraneous load in a familiar authentic environment was lower. Thus, completing learning activities with the support of the system a familiar authentic environment was easy and it did not require a lot of mental effort.

Related literature suggests that germane load is determined by appropriate instructional design and it can enhance learning. With the system support, this study attempted to direct students' attention to cognitive processes that are directly relevant to the learning material and tasks. The system functions enabled students to take pictures of learning objects in familiar authentic environment and then describe them with textual and voice annotations. Students could review their textual descriptions or listen to the audio recorded files afterwards. Students mentioned that completing the tasks in this way helped learning and made it more interesting. Furthermore, students claimed that learning in authentic familiar context and creating their own learning materials related to everyday life inspired students to engage in the materials and to try producing more meaningful output. Huang Chiu, Liu, and Chen (2011) suggested increasing students interest in learning and making them engaged in learning activities and tasks more by utilizing multimedia aids (e.g. pictures and audio). Furthermore, Caldwell (1998) argued that multimedia objects in learning stimulate students' imagination and helps bring out meaningful output. Some students, particularly of low ability, admitted that in the way they learn with the system, they could communicate in the target language with less anxiety of making mistakes. Chen and Chang (2009) argued that anxiety is a subjective feeling of worry, nervousness, or unease associated with arousal of the autonomic nervous system. Anxiety interferes with cognitive ability to absorb, process, and produce a foreign language. Furthermore, it negatively affects cognitive load. In contrast to traditional learning, experimental students learned with more confidence and their learning was more creative and enjoyable. The interview with students verified this finding. In interviews, students confirmed that learning content and activities in the electronic textbooks with a familiar authentic context were more interesting, fun, and engaging than that in traditional method. Therefore, this study may conclude that, with the support of the system, experimental students' germane load was higher in a familiar authentic environment. Experimental students concentrated more during learning in a familiar authentic environment; their mood was joyful and their frustration was low.

6. Conclusion

Two main findings were revealed in this study. First, the experimental students outperformed better than control students on post-test items in both experiments. Second, learning activities in a familiar authentic environment supported by the system enabled students to have less cognitive load compared to the learning setting without technological support. That is, the experimental students' intrinsic and extraneous load were reduced while germane load was increased compared to the control students.

Based on these findings, this study recommends educators to employ appropriate learning activity design and the system to facilitate students' learning achievement and their cognitive load management. In designing learning activities, the teacher has to consider how to make the best of the system to develop students' productive skills and managing cognitive load. For example, in this study, students took photos of learning objects and described them with text and voice. Photos and textual and audio descriptions were shared among students so that they could learn from each other and get some new ideas to improve their own homework. The system provided multiple channels for students to present their language output (i.e. taking pictures of learning objects and then describing them in written and oral ways) and gave students more opportunity to use the language. Thus, the teacher may organize learning activities in the way they were designed in this study. Furthermore, the teacher may encourage students to use the functions of the system, such as annotating, recording, assistance, and sharing, to reduce their intrinsic and extraneous loads and increase germane load. In this way, students can efficiently study learning material, complete the tasks, and enjoy the learning process at the same time. For example, students can take advantage of annotating to reflect on learning material and review reflections afterwards for better understanding of new concepts or for exam preparation. Students can also distribute their cognition to artifacts created in authentic environment with familiar context. Furthermore, the learning environment created by the system can reduce students' anxiety and help in giving meaningful output, especially low ability students.

There are several limitations found in the study that need to be considered. The first limitation concerns the relatively small sample size. The second limitation relates to short-term exposure of the technology to aid learning. For this reason, these findings cannot be generalized to a broader community based on this study alone or they have limited relevance to learning scenarios in which the technology long-term exposed in "real-world" conditions. These limitations will be addressed in a future study. In the future, our approach can be applied to other domains (e.g. Mathematics or Biology), and cognitive load can be measured objectively by observations of behavior or physiological conditions. The future study will also focus on how a familiar authentic context without mind tools can help to decrease cognitive load of students by comparing cognitive load of control and experimental groups.

References

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 53-61.

Caldwell, S. (1998) Going the distance with Online Instruction. *VSTE Journal*, *13*(1), 3–6.

Chang, C. C., Tseng, K. H., & Tseng, J. S. (2011). Is single or dual channel with different English proficiencies better for English listening comprehension, cognitive load and attitude in ubiquitous learning environment?. *Computers & Education*, *57*(4), 2313-2321.

Chen, I. J., & Chang, C. C. (2009). Cognitive load theory: An empirical study of anxiety and task performance in language learning. *Electronic Journal of Research in Educational Psychology*, *7*(2), 729-746.

Chen, N. S., Hsieh, S. W., & Kinshuk, A. (2008). Effects of short-term memory and content representation type on mobile language learning. *Language Learning and Technology*, *12*(3), 93-113.

Harmer, J. (2007). *The Practice of English Language Teaching*. London: Longman.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *7*(2), 174-196.

Huang, Y.M., Chiu, P.S., Liu, T.C., & Chen, T.S. (2011). The design and implementation of a meaningful learning-based evaluation method for ubiquitous learning. *Computers & Education*, *57*(4), 2291-2302.

Huang, Y. M., Huang, Y. M., Liu, C. H., & Tsai, C. C. (2013). Applying social tagging to manage cognitive load in a Web 2.0 self-learning environment. *Interactive Learning Environments*, *21*(3), 273-289.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language learning*, *51*(3), 539-558.

Hwang, G. J., Wu, P. H., Zhuang, Y. Y., & Huang, Y. M. (2013). Effects of the inquiry-based mobile learning model on the cognitive load and learning achievement of students. *Interactive Learning Environments*, *21*(4), 338-354.

Hwang, W.Y., Huang, Y.M., Shadiev, R., Wu, S.Y., Chen, S.L. (in press). Effects of using mobile devices on English listening diversity and speaking for EFL elementary students. *Australasian Journal of Educational Technology*.

Hwang, W.Y. & Shadiev, R. (2014). Cognitive Diffusion Model with User-Oriented Context-to-Text Recognition for Learning to Promote High Level Cognitive Processes. *Knowledge Management & E-Learning*, *6*(1), 30-48.

Hwang, W.Y., Shadiev, R., Hsu, J.L., Huang, Y.M., Hsu, G.L., Lin, Y.C. (2014). Effects of Storytelling to Facilitate EFL Speaking Using Web-Based Multimedia System. *Computer Assisted Language Learning*, (ahead-of-print), 1-27. DOI: 10.1080/09588221.2014.927367

Hwang, W.Y., Shadiev, R. & Huang, S.M. (2011). A study of a multimedia web annotation system and its effect on the EFL writing and speaking performance of junior high school students. *ReCALL*, *23*(2), 160–180.

Lu, J., Lai, M., & Law, N. (2010). Knowledge building in society 2.0: Challenges and opportunities. In *New Science of Learning* (pp. 553-567). Springer New York.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, *38*(1), 43-52.

Shadiev, R., Hwang, W.Y., & Huang, Y.M. (in press). A pilot study of facilitating cross-cultural understanding with project-based collaborative learning activity in online environment. *Australasian Journal of Educational Technology*.

Shadiev, R., Hwang, W.Y., Huang, Y.M., & Yang, Y.S. (2014). A study of using a multi-touch tabletop technology to facilitate collaboration, interaction, and awareness in co-located environment. *Behaviour & Information Technology*, (ahead-of-print), 1-12. DOI: 10.1080/0144929X.2014.942755

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296.

Applying Adaptive Hybrid Recommendation Technology for Searching Algorithm Learning Articles

Shu-Chen Cheng, Shih-Che Huang
Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology, Taiwan
kittyc@mail.stust.edu.tw

Abstract：In this generation, technology is developed in tremendous speed. The Information on the Internet is increasing at a high speed each day. People get plenty of information via the search engine and spend much time to filter out insignificant information at the same time. Therefore, this research system can filter all kinds of articles to exclude advertisement, news from network bookstore, or insignificant information related to keyword. After filtering, this system will gather up all useful articles and provide users to review. This study proposes the hybrid recommended system with multi-adaptive recommendation to learners. Hybrid recommendation is divided into two ways to recommend "Content-Based Recommendation" and "Collaborative Filtering Recommendation" articles. First, content-based recommendation is based on Term Frequency- Inverse Document Frequency to estimate the characteristic values of articles. Then, we set the weight of difficulty of keywords. After that, people can decide the level of article in the beginning and use it for reference. Besides, collaborative filtering recommendation is applied based on user abilities estimated by IQ tests, quizzes, online tests, ability certificates, and other exams. When the result of two users is similar and one of them thinks it's useful, the article will be automatically forwarded to the other.

Keyword: Content-Based Recommendation System, Filtering Recommendation, Term Frequency-Inverse Document Frequency.

1. Introduction

1.1 Background information

When searching for articles tend to obstruct because of too much information on the Internet. As shown in figure 1, users need to spend much time filtering information when they read it. In the search result, it also accompanies with irrelevant sites of learning such as: advertisement, information of bookstore…etc. Owing to this reason, this study uses the hybrid recommendation to assist users to filter the articles they don't need and learn efficiently.

Figure 1. search result by Google

1.2 Purpose of Research

This study will focus on hybrid methods recommended to build a system for recommended articles. The hybrid recommended system is combined with "Content-Based Recommendation" and Collaborative Filtering Recommendation." Content-Based Recommendation is to analyze the content of articles keyword. And another one Collaborative Filtering Recommendation selects the recommended articles by the ability of learners and peer recommendation.

2 Literature Review

The chapter is to explore how to use the hybrid recommended system. Then introduce the advantage and disadvantage when users use this system and give the advice in detail to improve it. [1]

2.1 Classification

The most important part of Content-Based Recommendation is the classification. We need to analyze the content of article, set keywords, and article categories.

2.1.1 N-Gram word segmentation

After selecting articles by Crawler on the Internet, we need to deal with the content of these articles first. How to analyze these articles? We use each word of the articles to do the cutting so that the article presents the smallest unit of word. Then by smallest word unit of different lengths to make up words, it can form words and sentences. N-Gram word segmentation is used at the beginning[2][3] of the data processing. (Table1)

Table1. length of word segmentation (example: string of Algorithm by Chinese characters)

N-Gram	results
2-Gram	演、算、法、演算、算法
3-Gram	演、算、法、演算、算法、演算法

2.1.2 Term Frequency-Inverse Document Frequency

After using the N-gram word segmentation, the system starts to determine the importance of each string in the article. Take Knapsack algorithm for example, "Knapsack" and "Algorithm" will be the most important word. When the strings of the importance in the article to be found out, these strings are used for the characteristic value of this article.

The characteristic value is determined by Term Frequency-Inverse Document Frequency (TF-IDF). [4][5]

2.2 Recommendation System

2.2.1 Content-Based Recommendation System

On the basis of Content-Based Recommendation System is to analyze the content rather than the user for evaluation. This system will calculate the ability of user's interest and then calculate the characteristic values of user. Therefore it can find the item user needs. In this study, the most important thing is to recommended articles through keywords and characteristic values. To achieve this recommended system, we need to use characteristic values to make up keyword and given the different scores by different rights then the result of recommendation will be come out.

2.2.2 Collaborative Filtering Recommendation System

CF system is abbreviated form of Collaborative Filtering Recommendation System. This method was successfully applied in a wide variety of fields, such as YouTube, Amazon, etc., are very famous examples. This system is based on the gathering user preferences, personal information, gender, browsing habits and other personal information [6], to make recommendation. Therefore, while user browsing merchandise or other things, CF system will recommend users the section plus the recommended information, not only can make commodities accelerated been viewed but also speed up information dissemination.

2.2.3 hybrid Recommendation System

Hybrid recommendation System recommended by Content-Based Recommendation System and Collaborative Filtering Recommendation system. We can use these two different systems which have different strength and shortages to improve recommendation system better. Therefore, hybrid will compensate for the shortcomings of their systems to each other. The advantage and disadvantage of two recommendation systems are shown in Table 2. [7][8]

Table 2. Comparisons between Content-Based Recommendation and Filtering Recommendation

	Content-Based Recommendation	Filtering Recommendation
reference value	contents	Information of user
advantage	Quick Recommendation	Find out the preference of user High-automated
shortage	Can’t find out the user’s preference Much time to do processing information Can’t be recommended automatically	New Item Sparsity Cold-start Scalability New user

Both of Content-Based Recommendation System and Collaborative Filtering Recommendation System all have drawbacks, so we combine[9] these two ways to be hybrid recommendation. (Table3)

Table3. hybrid recommendation way

method	hybrid recommendation way
Method of weighting	Take two different recommendation methods to calculate, and in accordance with the result of calculate to given different weighting.
Method of selection	If there is users’ information, use Collaborative Filtering Recommendation and Content-Based Recommendation. If not, use Content-Based Recommendation only.
Method of mixing	When both content have recommendation systems, use the two recommended ways to calculate, and then generate a recommendation.

3 Research Method

3.1 System Framework

This study is based on the self-developed Algorithm platform to design a hybrid recommendation system to give users search teaching articles. (Figure.2) We use the Algorithm database to do the recommendation. The hybrid recommendation is divided into two modules which are “Content-Based Recommendation module "and" collaborative filtering recommendation module". In the Content-Based Recommendation module, the article needs to through preprocessing module first and then use N-Gram for word segmentation. After that, the characteristic value of the article will be selected which will be the keyword to analyze the degree of the article difficulties and make recommendation. The Collaborative Filtering Recommendation use the calculated module to calculate the users’ ability. Then in accordance with the users’ ability to recommend or accordance with the recommendation module to allows users with the same ability to recommend the article and share with them. Therefore, it will have many different recommended articles to let users read.

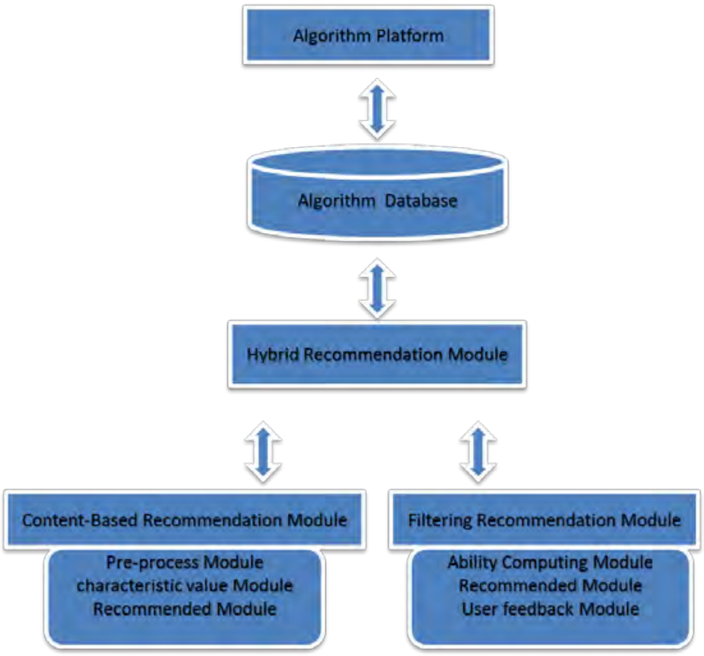


Figure.2 System Framework

3.2 Content-Based Articles Recommendation Method

When the analysis of the article that will be processed by the keyword of characteristic value and then in accordance with the value of calculation to recommend users. The weighting of keyword is the most important part of this method which is in accordance with given to those area of expertise; conducting analysis to improve ease of articles and recommended stability.

3.2.1 Articles Pre-process

Articles are pre-processed by N-Gram segmentation to handle the database articles. There are many Stop Words that needs to be removed. Chinese Stop word List is shown in Table 4.

Table4. Examples of Chinese Stop Words List

Stop Words
嗎（麼）、吧、呢、罷、呀、啊、啦、哇、哪、罷了、而已、我、你、您、我們、他們、是、全部、所有、不知道

3.2.2 Initial Item Bank

This system is to determine whether the article for Algorithm teaching articles and set up the degree of difficulties to achieve recommendation. It needs to set up positive, negative item bank, and weighting. The positive item bank put in the words represents Algorithm; however, the negative item bank represents non-Algorithm.

First of all, the initial item bank set up by the artificial, the positive item bak select the keyword that can represent Algorithm from the book of Algorithm. The more abstruse terms set up higher weighting that represents the more difficult article. On the contrary. The negative key words comes out from the advertisement of on-line bookstore. Therefore, the system can exclude those of words to choose Algorithm article. (Figure.3)

Forward keyword	Weights
Hoffmann	1
greedy	1
Binary	1
merge	1
Bubble	2
Prim	2
Kruskal	2

Negative keyword	Weights
The authors introduce	1
Library	1
Purchasing	1
Books	1

Figure.3 Positive and Negative Keywords and Weightings

3.2.3 Recommendation Module

When the item bank set up, then the keywords in the item bank will be compared and do the weighting accumulated. If the article in the positive item bank has higher weighting, then it can be used as recommended priorities. This module can be used while users searching the article in accordance with the recommended system.

3.3 Collaborative Filtering Recommendation Methods

The system, another recommendation System, is Collaborative Filtering Recommendation. Since the Algorithm belongs to professional in the field, not like English article which contains preference, and article style such as exercise, movies, and music. Therefore, in the study we propose a Peer Recommendation, which is based on user's ability and preference from groups. They can find appropriate article by sharing and watching the recommended article from the same ability users.

3.3.1 Recommendation by Users' Abilities

Although Algorithm is in the same article; however it can be different by users, articles and the degree of the articles. Sometimes it's too difficult to let beginning user or not good at logical thinking user to understand. On the other hand, if it's too easy to understand, it can't help user who has higher ability learning. This study will be in accordance with the user's ability to recommend the article (Recommend Articles), initially to collect the user's test scores, and recorded data repository, such as intelligence tests, proficiency examinations test program, license, Algorithm online quizzes, etc. After that, set the initial capacity and in accordance with article difficulty keyword weighting to classify and control user ability to give a different degree of difficulty of the article. [11][12]

3.3.2 Group Preferences Recommendation Method

When one of users think the article is helpful, the system will record it. Next time, if there is a user of same ability appears, the system will recommend the article to him. Therefore, user can search appropriate article easily.

3.4 Hybrid Recommendation Method

When user enters the keyword in the Content-Based Recommendation, it will search out the article. After that, the collaborative Filtering Recommendation will offer recommended article to the user. Because Algorithm articles are not much, when users all recommended an article, the article will be exposed highly. Because of this reason, this type of the article order will increase and be found easily.
If this article isn't related to the keyword, it will give the system in return. The system will reduce the article exposure to avoid system erroneous judgment that causing users to research on the system distress to make the system more perfect. If the feedback has found that the article isn't related to the

mathematics article or the article can't be connected, the system will reduce the recommendation. When it comes to an certain amount, the system will not recommend to the user and remove it.

4 Conclusions

In this study, we design a hybrid recommendation system which offers user to find appropriate articles according to their abilities, keywords, or groups recommendation articles so that it can achieve hybrid recommendation effect. Users can follow their own abilities or similar users' to read the recommended articles in accordance with collaborative filtering recommendation system. By the way, the feedback from user can increase the article exposure and improve the content-based recommendation. All of these systems can help users find appropriate algorithm articles and enhance learning abilities.

Acknowledgement

This study is supported in part by the Science Council of the Taiwan under Contract No. NSC 100-2511-S-218 -008 -MY3

References

[1]Adomavicius G. and Tuzhilin A. (2005), Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 6, pp.734-749

[2]Jerome R. Bellegarda(2000),Exploiting Latent Semantic Information in Statistical Language Modeling,Proceedings of the IEEE,Vol.88,No.8,pp.1279-1296.

[3]Qiang Yang and Zhang,H.H.(2003),Web-Log Mining for Predictive Web Caching,IEEE Transactions on Knowledge and Data Engineering,Vol.15,No.4,pp.1050-1053.

[4]Wu HC,Luk RWP,Wong KF,Kwok KL(2008),Interpreting tf-idf term weights as making relevance decisions,ACM Transactions on Information Systems,Vol.26,No.3,pp.13:1-13:36.

[5]Salton G,Buckley C(1988),Term-weighting approaches in automatic text retrieval,Information Processing & Management, Vol.24,No.5,1988, pp.513-523.

[6]Rokach L.and Kisilevich S.(2012),Initial Profile Generation in Recommender Systems Using Pairwise Comparison,IEEE Transactions on Systems,Man,and Cybernetics,Part C:Applications and Reviews,Vol.42,No.6,pp.1854-1859.

[7]Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989).User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Manag. Sci., vol. 35, no. 8, pp. 982–1003

[8]Ahmad A. Kardan and Mahnaz Ebrahimi (2013), A novel approach to hybrid recommendation systems based

on association rules mining for content recommendation in asynchronous discussion groups, Information Sciences, Vol. 219, pp.93-110.

[9]Robin Burke(2002),Hybrid Recommender Systems Survey and Experiments.,User Modeling and User-Adapted Interaction,Vol.12,No.4,pp.331-370.

[10]Wikipedia(n.d.),"Stop words",https://en.wikipedia.org/wiki/Stop_words.

[11]Srinivasan, A. (1985). Alternative measures of system effectiveness: Associations and implications. MIS quarterly, 9, 243-253.

[12]Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research.