ICCE2012

The 20th International Conference on Computers in Education

26 - 30 November 2012 Singapore

Proceedings of the 20th International Conference on Computers in Education ICCE 2012

Editors

Gautam BISWAS Lung-Hsiang WONG Tsukasa HIRASHIMA Wenli CHEN

NIE NATIONAL INSTITUTE OF EDUCATION SINGAPORE

An Institute of NANYANG TECHNOLOGICAL UNIVERSITY

APSCE

Learning Sciences Lab

LEARNING SCIENCES AND TECHNOLOGIES

Proceedings of the 20th International Conference on Computers in Education ICCE 2012

November 26 2012 - November 30 2012
Singapore

ISBN 978-981-07-4649-0

Publisher:
National Institute of Education,
1, Nanyang Walk,
Singapore 637616.

ORGANIZATION

Conference Organization

Conference Chair:
Tsukasa HIRASHIMA, Hiroshima University, Japan

International Program Coordination Committee:

Chair:
Gautam BISWAS, Vanderbilt University, USA

Co-Chair:
Lung Hsiang WONG, Nanyang Technological University, Singapore

Education Program Committee:
Chen-Chung LIU, National Central University (Chair)
Gautam BISWAS, Vanderbilt University, USA
Ben CHANG, National Chiayi University, Taiwan
Weiqin CHEN, University of Bergen, Oslo and Akershus University, Norway
KINSHUK, Athabasca University, Canada
Hiroaki OGATA, Tokushima University, Japan
Su Luan WONG, Universiti Putra Malaysia, Malaysia

Workshop/ Tutorial/ Interactive-Event Coordination Co-Chairs:
Ben CHANG, National Chiayi University, Taiwan
Seng Chee TAN, Nanyang Technological University, Singapore
Tatsunori, MATSUI, Waseda University, Japan

Doctoral Student Consortium (DSC) Coordination Co-Chair:
Ching Sing CHAI, Nanyang Technological University, Singapore

Poster Coordination Co-Chairs:
Jimmy LEE, The Chinese University of Hong Kong, Hong Kong
Doris CHOY, Nanyang Technological University, Singapore

Local Organizing Committee:
(all are affiliated to Nanyang Technological University, Singapore)

Chair:
Wenli CHEN

Co-Chair:
Lung Hsiang WONG
Advisors:
Seng Chee TAN
Victor CHEN
Daniel TAN

Conference Secretariat:
Shirley SEET

Sub-Committee Co-Chairs:
Mingfong JAN
Alice KHOO
Mi Song KIM
Cheryl LEE
June LEE
Rose LIANG
Chee Kit LOOI
Shirlene TANG
Angela WONG
Philip WONG

Other Sub-Committee Members:
Yan Sam CHEE
Young Hoan CHO
Doris CHOY
Louis GOH
Ahmed HILMY
Azilawati Bte JAMALUDIN
Beaumie KIM
Faith KOH
Hon Jia KOH
Joyce KOH
Sastha KUMAR LAKSHMI
Jen-Yi LEE
Yanqiu LI
Dalvin Jit Kaur SIDHU
Nurazirah Binte MISDANI
Wei Leng NEO
Zi Jia NG
Darren NONIS

John SOO
Vincent SUNG
Denise TAN
Jacqueline TAN
Yuh Huann TAN
Eileen TEO
Yancy TOH
Li-Yi WANG
Yun WEN
Huay Lit WOO
Wenting XIE
Eric YEO

Consultants:
Thepchai SUPNITHI, National Electronics and Computer Technology Center, Thailand
Su Luan WONG, Universiti Putra Malaysia, Malaysia

Special Interest Groups (SIG) Chairs:
Akihiro KASHIHARA, The University of Electro-Communications, Japan (C1)
Chee-Kit LOOI, Nanyang Technological University, Singapore (C2)
Kiyoshi NAKABAYASHI, Chiba Institute of Technology, Japan (C3)
Lung Hsiang WONG, Nanyang Technological University, Singapore (C4)
Ben CHANG, National Chiayi University, Taiwan (C5)
Yu-Ju LAN, National Taiwan Normal University, Taiwan (C6)
Siu Cheung KONG, The Hong Kong Institute of Education, Hong Kong (C7)
Su Luan WONG, Universiti Putra Malaysia, Malaysia (DICTAP)

Program Committee for C1

ICCE Conference on Artificial Intelligence in Education/Intelligent Tutoring System (AIED/ITS) and Adaptive Learning

Program Co-chairs

Akihiro KASHIHARA, The University of Electro-Communications, Japan (Executive PC-Chair)
Gordon McCALLA, University of Saskatcehwan, Canada
Thepchai SUPNITHI, National Electronics and Computer Technology Center, Thailand
Eduardo GUZMAN, University of Malaga, Spain

PC Members

Akira TAKEUCHI, Kyushu Institute of Technology, Japan
Amali WEERASINGHE, University of Canterbury, New Zealand
Ben DU BOULAY, University of Sussex, United Kimgdon
Chad LANE, University of South California, USA
Chao-Lin LIU, National Chengchi University, Taiwan
Chas MURRAY, Carnegie Learning, USA
Chih-Yueh CHOU, Yuan Ze University, Taiwan
Chi-Jen LIN, Fo Guang University, Taiwan
Cristina CONATI, University of British Columbia.
Diego ZAPATA-RIVERA, Educational Testing Service, USA
Dietrich ALBERT, Graz University of Technology, Austria
George GHINEA, Brunel University, UK
Helen PAIN, University of Edinburgh, United Kingdom
Jitti NIRAMITRANON, Kasetsart University,Thailand
Jose-Luis PEREZ-DE-LA-CRUZ, Universidad de Malaga, Spain
Kazuaki KOJIMA, Waseda University, Japan
Kazuhisa MIWA, Nagoya University, Japan
Krittaya LEELAWONG, Mahidol University International College, Thailand
Martin WOLPERS, Fraunhofer-Institut FIT, Germany
Marut BURANARACH, NECTEC, Thailand
Mercedes RODRIGO, Ateneo de Manila University, Philippines
Moffat MATHEWS, University of Canterbury, New Zealand
Monica TRELLA, Universidad de Malaga, Spain
Nilufar BAGHAEI, Unitec, New Zealand
Noboru MATSUDA, Carnegie Mellon University, USA
Noriyuki MATSUDA, Wakayama University, Japan
Norwati MUSTAPHA, Universiti Putra Malaysia, Malaysia
Onjira SITTHISAK, Thaksin University , Thailand
Riichiro MIZOGUCHI, Osaka University, Japan
Sagaya AMALATHAS, University of Canterbury, New Zealand
Sergey SOSNOVSKY, DFKI, Germany
Sridhar IYER, Indian Institute of Technology, India
Shinobu HASEGAWA, Japan Advanced Institute of Science and Technology, Japan
Siu Cheung KONG, The Hong Kong Institute of Education, Hong Kong
Susan BULL, University of Birmingham, United Kingdom

Tanya MITROVIC, University of Canterbury, New Zealand
Tatsuhiro KONISHI, Shizuoka University, Japan
Tatsunori MATSUI, Waseda Univeristy, Japan
Thanaruk TEERAMUNKONG, Sirindhorn International Institute of Technology, Thailand
Tomoko KOJIRI, Nagoya University, Japan
Tomoya HORIGUCHI, Kobe University, Japan
Tsukasa HIRASHIMA, Hiroshima University, Japan
Ulrich HOPPE, University of Duisburg-Essen
Vilas WUWONGSE, Asian Institute of Technology, Thailand
Virach SORNLERTLAMVANICH, NECTEC, Thailand
Yong Se KIM, Sungkyunkwan University, Korea
Yusuke HAYASHI, Nagoya University, Japan

Program Committee for C3
ICCE Conference on Advanced Learning Technologies, Open Contents, and Standards

Program Co-chairs
Yusuke HAYASHI, Hiroshima University, Japan (Executive PC-Chair)
Tatsuhiro KONISHI, Shizuoka University, Japan
Manuel Caeiro Rodríguez, University of Vigo, Spain
Tore HOEL, Oslo and Akershus University College of Applied Sciences, Norway

PC Members
Kumiko AOKI, The Open University of Japan, Japan
David ASIRVATHAM, University of Malaya, Malaysia
Chih-Kai CHANG, National University of Tainan, Taiwan
Irene CHEN, National Changhua University of Education, Taiwan
Gwo-Dong CHEN, National Central University, Taiwan
Tsung-Yen CHUANG, National University of Tainan, Taiwan
Jean-Noël COLIN, University of Namur, Belgium
Michael DERNTL, RWTH Aachen University, Germany
Giovanni FULANTELL, Italian National Research Council, Italy
Seungyeon HAN, Hanyang Cyber University, South Korea
Jeff Cheng-Hsu HUANG, Hwa-Hsia Institute of Technology, Taiwan
Ray Yueh-Min HUANG, National Cheng-Kung University, Taiwan
Ren-Hung HWANG, National Chung-Cheng University, Taiwan
Rozhan M. IDRUS, Universiti Sains Malaysia, Malaysia
Yih-Ruey JUANG, Jinwen University of Science and Technology, Taiwan
Abtar KAUR, Open University Malaysia, Malaysia
Hidenobu KUNICHIKA, Kyushu Institute of Technology, Japan
Lam-for KWOK, City University of Hong Kong, Hong Kong
James LAFFEY, University of Missouri, Columbia, USA
Chien-Sing LEE, National Central University, Taiwan
Jerry LEESON. Education Services Australia, Australia
Ching-Jung LIAO, Chung Yuan Christian University, Taiwan
Kin Chew LIM, SIM University, Singapore
Jon MASON, Queensland University of Technology, Australia
Yukihiro MATSUBARA, Hiroshima City University, Japan
Kenji MATSUURA, Tokushima University, Japan
Kiyoshi NAKABAYASHI, Chiba Institute of Technology, Japan
Kuo-Liang OU, National Hsinchu University of Education, Taiwan
Daniel R. REHAK, Advanced Distributed Learning Initiative, USA
Hitoshi SASAKI, Takushoku University, Japan
Ju-Ling SHIH, National University of Tainan, Taiwan
Onjira SITTHISAK, Thaksin University, Thailand
Christian M. STRACKE, University of Duisburg-Essen, Germany
Nguyen VIET HA, Vietnam National University, Vietnam
Maggie M. WANG, Hong Kong, The University of Hong Kong, Hong Kong
Hiroyoshi WATANABE, Teikyo University, Japan

Program Committee for C4

ICCE Conference on Classroom, Ubiquitous, and Mobile Technologies Enhanced Learning (CUMTEL)

Program Co-chairs

Tzu-Chien LIU, National Central University, Taiwan (Executive PC-Chair)
Barbara WASSON, University of Bergen, Norway
Gustavo ZURITA, University of Chile, Chile
Hyo-Jeong SO, Nanyang Technological University, Singapore

PC Members

Nelson BALOIAN, University of Chile, Chile
Ivica BOTICKI, University of Zagreb, Croatia
Tessy CERRATTO PARGMAN, Stockholm University, Sweden
Maiga CHANG, Athabasca University, Canada
Gwo-Dong CHEN, National Central University, Taiwan
Sherry CHEN, National Central University, Taiwan
Tzung-Shi CHEN, National University of Tainan, Taiwan
Shu-Chen CHENG, Southern Taiwan University, Taiwan
Carol CHU, Soochow University, Taiwan
Daniel CHURCHILL, Hong Kong University, Hong Kong
Monica DIVITINI, Norwegian University of Science and Technology, Norway
Moushir ELBISHOUTY, The University of Tokushima, Japan
Haiguang FANG, Capital Normal University, China
Serge GARLATTI, Telecom Bretagne, France
Daniela GIORDANO, University of Catania, Italy
Christian GLAHN, Swiss Federal Institute of Technology Zurich, Switzerland
Xiaoqing GU, East China Normal University, China
Luis GUERRERO, University of Costa Rica, Costa Rica
Yoshiaki HADA, The Open University of Japan, Japan
Ulrich HOPPE, University Duisburg-Essen Faculty, Germany
Hsien-Sheng HSIAO, National Taiwan Normal University, Taiwan
Ray HUANG, National Cheng Kung University, Taiwan
Ronghuai HUANG, Beijing Normal University, China
Gwo-Jen HWANG, National Taiwan University of Science and Technology, Taiwan
Tomoo INOUE, University of Tsukuba, Japan
Mingfong JAN, Nanyang Technological University, Singapore
Yih-Ruey JUANG, Jinwen University of Science and Technology, Taiwan
Hiroshi KATO, The Open University of Japan, Japan
Matthew KAM, Carnegie Mellon University, USA
KINSHUK, Athabasca University, Canada
Siu-Cheung KONG, The Hong Kong Institute of Education, Hong Kong
Chih-Hung LAI, National Dong Hwa University, Taiwan
Yu-Ju LAN, National Taiwan Normal University, Taiwan
Oscar LIN, Athabasca University, Canada
Robin LIN, National Hsinchu University of Education, Taiwan
Chen-Chung LIU, National Central University, Taiwan

Program Committee for C5
ICCE Conference on Game and Toy Enhanced Learning and Society (GTEL&S)

Program Co-chairs
Yam San Chee, Nanyang Technological University, Singapore (PC Executive Chair)
Thorkild Hanghøj, Aarhus University, Denmark
Ana Paiva, Technical University of Lisbon, Portugal
Masanori Sugimoto, University of Tokyo, Japan

PC Members
Thomas APPERLEY, University of Melbourne, Australia
Hans Christian ARNSETH, University of Oslo, Norway
Fernández-Manjón BALTASAR, Universidad Complutense de Madrid, Spain
Catherine BEAVIS, Griffith University, Australia
Tak-Wai CHAN, National Central University, Taiwan
Ben CHANG, National Chiayi University, Taiwan
Maiga CHANG, Athabasca University, Canada
Ming-Puu CHEN, National Taiwan Normal University, Taiwan
Michael DEZUANNI, Queensland University of Technology, Australia
Morris S. Y. JONG, The Chinese University of Hong Kong, Hong Kong
Matthew KAM, Carnegie Mellon University, USA
KINSHUK, Athabasca University, Canada
Chien-Sing LEE, National Central University, Taiwan
Fong Lok LEE, The Chinese University of Hong Kong, Hong Kong
Jimmy LEE, The Chinese University of Hong Kong, Hong Kong
Peter LEONG, Singapore Polytechnic, Singapore
Hiroyuki MITSUHARA, Tokushima University, Japan
Martin OLIVER, University of London, UK
Rikke MAGNUSSEN, University of Aalborg, Denmark
Demetrios G SAMPSON, University of Piraeus, Greece
Junjie SHANG, Peking University, People’s Republic of China
Gunilla SVINGBY, Malmö University, Sweden
Hiroyuki TARUMI, Kagawa University, Japan
Christopher WALSH, The Open University, UK

Program Committee for C6
ICCE Conference on Technology Enhanced Language Learning (TELL)

Program Co-chairs
Jie-Chi Yang, National Central University, Taiwan (Executive PC-Chair)
Jozef Colpaert, University of Antwerp, Belgium
Wen-Chi Vivian Wu, Providence University, Taiwan
Yu-Ju Lan, National Taiwan Normal University, Taiwan

Advisory Board Members
Jack Burston, University of Cyprus, Cyprus
Nian-Shing Chen, National Sun Yat-sen University, Taiwan
Philip L. Hubbard, Stanford University, USA
Glenn Stockwell, Waseda University, Japan

PC Members
Emerita Bañados, Universidad de Concepción, Chile
Jean-Claude Bertin, Université du Havre, Le Havre, France
Chih-Kai Chang, National University of Tainan, Taiwan
Mei-Mei Chang, National Pingtung University of Science and Technology, Taiwan
Pei-Chin Chang, National Taiwan Normal University, Taiwan
Yu-Chuan Chao, Providence University, Taiwan
Chi-Fen Chen, National Kaohsiung First University of Science and Technology, Taiwan
Ching-Huei Chen, National Changhua University of Education, Taiwan
Howard Hao-Jan Chen, National Taiwan Normal University, Taiwan
Weiqin Chen, University of Bergen, Norway
Hsueh-Hua Chuang, National Sun Yat-sen University, Taiwan
Shu-Mei Chwo, Hungkuang University, Taiwan
David Coniam, The Chinese University of Hong Kong, Hong Kong
Yoshiko Goda, Kumamoto University, Japan
Ying-Hua Guan, National Taiwan Normal University, Taiwan
Mirjam Hauck, The Open University, UK
Ching-Kun Hsu, National University of Tainan, Taiwan
Heng-Tsung Danny Huang, National Sun Yat-sen University, Taiwan
Hui-Wen Huang, National Sun Yat-sen University, Taiwan
Ray Yueh-Min Huang, National Cheng Kung University, Taiwan
Ronghuai Huang, Beijing Normal University, China
Hsiu-Ting Hung, National Kaohsiung First University of Science and Technology, Taiwan
Shao-Ting Hung, National Kaohsiung First University of Science and Technology, Taiwan
Wu-Yuin Hwang, National Central University, Taiwan
Kristi Jauregi, Utrecht University, The Netherlands
Qun Jin, Waseda University, Japan
Agnes Kukulska-Hulme, The Open University, UK
Chih-Hung Lai, National DongHwa University, Taiwan
Chun Lai, University of Hong Kong, Hong Kong
Hsing-Chin Lee, National Taipei College of Business, Taiwan
Mark J.W. Lee, Charles Sturt University and University of New England, Australia

Program Committee for C7
ICCE Conference on Technology, Pedagogy and Education

Program Co-chairs
Wenli CHEN, Nanyang Technological University, Singapore (Executive PC-Chair)
Pilar LACASA, University of Alcala, Spain
Tak Lam Ivan WONG, The Hong Kong Institute of Education, Hong Kong
Yanjie SONG, The Hong Kong Institute of Education, Hong Kong

PC Members
Zoraini Wati ABAS, Open University Malaysia, Malaysia
Kumiko AOKI, The Open University of Japan
Ching Sing CHAI, Nanyang Technological University, Singapore
Gwo-Dong CHEN, National Central University, Taiwan
Pearl CHEN, California State University, USA
Weiqin CHEN, University of Bergen, Norway
Wing Sum CHEUNG, Nanyang Technological University, Singapore
Young Hoan CHO, Nanyang Technological University, Singapore
Doris CHOY, Nanyang Technological University, Singapore
Daniel CHURCHILL, The University of Hong Kong, Hong Kong
Bronwen COWIE, University of Waikato, New Zealand
Shanti DIVAHARAN, Nanyang Technological University, Singapore
Robert Fox, The University of Hong Kong, Hong Kong
Khe Foon HEW, Nanyang Technological University, Singapore
Curtis HO, University of Hawaii at Manoa, USA
Kian Sam HONG, University Malaysia Sarawak, Malaysia
Gwo-Jen HWANG, National Taiwan University of Science and Technology, Taiwan
Mingfong JAN, Nanyang Technological University, Singapore
In-Su KIM, Korea National Sport University
Mi Song KIM, Nanyang Technological University, Singapore
Hwee Ling Joyce KOH, Nanyang Technological University, Singapore
Kwok-wing LAI, University of Otago, New Zealand
Yiu Chi LAI, The Hong Kong Institute of Education, Hong Kong
Chwee Beng LEE, University of Western Sydney, Australia
Insook LEE, Sejong University, Korea
Ji-Yeon LEE, Inha University, South Korea
Mimi Miyoung LEE, University of Houston, USA
Kai Ming LI, Hong Kong Institute of Education, Hong Kong
Rose Yee Hing LIANG, Nanyang Technological University, Singapore
Shu-Sheng LIAW, China Medical University, Taiwan
Chiu-Pin LIN, National HsinChu University of Education, Taiwan
Zhi Feng Liu, National Central University, Taiwan
Will MA, Hong Kong Shue Yan University, Hong Kong
Mai NEO, Multimedia University, Malaysia
Shaun NYKVIST, Queensland University of Technology, Australia
Wakio OYANAGI, Nara University of Education, Japan
Innwoo PARK, Korea University, Korea

Message from the Conference Chair, Programme Chair/Co-Chairs and Local Chair

It gives us great pleasure to extend our warmest welcome to all participants of the 2012 International Conference on Computers in Education (2012). This year, the conference is being held in Singapore, a vibrant multi-ethnic city-state located in the heart of Southeast Asia, and a leader in educational innovation. Building on the success of the previous conferences, the program aims to foster lively exchanges and global collaborations on understanding, interpreting, and applying the theories and practices in the field of computers in education.

We have a strong line-up of four keynote speakers from across the globe this year. This includes: 1) Yasmin Kafai from the University of Pennsylvania who will be addressing how youth learn through play amid navigating public digital spaces and making social connections; 2) Fu-Yun Yu from the National Cheng Kung University, Taiwan, who will be illustrating how pedagogy, theory, practice, technology and research interplay to create a learner-centered space exemplified by student question-generation; 3) Seng Chee Tan from the Nanyang Technological University, Singapore who will be anchoring his talk around the perspectives of knowledge creation theories and their implications for education; and 4) Marcelo Milrad from the Linnaeus University, Sweden who will be looking into the current perspectives and challenges revolving the innovation and sustainability of mobile learning.

In addition, we have three invited theme-based talks: (1) "Learning by Creating Cognitive Models" by Kazuhisa Miwa from the Nagoya University, Japan; (2) "Open Content for Deeper Inquiry" by Jon Mason from the Charles Darwin University, Australia and (3) "The No Significant Difference Syndrome and the Ecological Paradigm Shift" by Jozef Colpaert from the University of Antwerpen, Belgium. These presentations will provide further stimulus for us to re-think the role of technology in shaping mental models, cognitive reasoning, inquiry learning and the design of learning ecology so as to create alignment between the epistemological, methodological and ecological underpinnings of using technology for educational impact.

This large-scale conference is a culmination of the collective effort across departments, institutions, countries, and even continents. It is not an easy feat. With immense gratitude, we would like to thank all the kind individuals who have rendered their help in every possible way to make this conference a reality, and contributed in no small measure to its success. To recognize these efforts, the names of the conference organization members as well as paper reviewers are listed in the proceedings.

Lastly, no less deserving of thanks are our conference participants who have also contributed to the conference's success through their active participation in a variety of roles: authors and co-authors of papers and posters, session chairs, panelists, and as regular attendees. I hope all participants will have further opportunities to create new friendships and professional collaborations, and to explore Singapore. With its alluring fusion of cultures, interesting blend of tradition and modernity; as well as new iconic buildings rapidly changing the skyline of Singapore, there is bound to be something for everyone.

Thank you.

Tsukasa Hirashima
Conference Chair
(Japan)

Gautam Biswas
Program Chair
(USA)

Wenli Chen
Local Chair
(Singapore)

Lung-Hsiang Wong
Program Co-Chair / Local Co-Chair
(Singapore)

PREFACE

The 20th International Conference on Computers in Education (ICCE 2012) is a series of annual conferences encompassing a broad range of issues related to using information technology for education. The ICCE conference series is organized under the auspices of the Asia-Pacific Society for Computers in Education (APSCE). This year, ICCE 2012, being held from 26 November to 30 November 2012 at the National Institute for Education in Singapore, has 361 early registrants from 29 different countries. Like previous conferences in this series, ICCE 2012 is structured as a meta-conference to allow researchers in the Asia-Pacific region to connect with international research communities and with each other for the worldwide dissemination and sharing of research, development, and deployment ideas that span the field of Computers in Education. Seven interrelated sub-conferences on specialized themes, each organized by a program committee appointed by the respective special interest group (SIG – see http://www.apsce.net/SIGs.aspx), constitute the five-day Conference schedule. They are:

C1: ICCE Conference on Artificial Intelligence in Education/Intelligent Tutoring System (AIED/ITS) and Adaptive Learning

C2: ICCE Conference on Computer-supported Collaborative Learning (CSCL) and Learning Sciences

C3: ICCE Conference on Advanced Learning Technologies, Open Contents, and Standards

C4: ICCE Conference on Classroom, Ubiquitous, and Mobile Technologies Enhanced Learning (CUMTEL)

C5: ICCE Conference on Digital Game and Digital Toy Enhanced Learning and Society (GTEL&S)

C6: ICCE Conference on Technology Enhanced Language Learning (TELL)

C7: ICCE Conference on Technology, Pedagogy and Education (TPE)

The Program Committee is comprised of a strong team that includes the Conference chair, the Program Coordination Chair and co-Chair, seven executive sub Conference Chairs and 287 experts in the field of Computers in Education from 33 different countries. Two former ICCE conference chairs have played the role of consultants in overseeing the conference organization process.

In total, the conference received a total of 247 papers (170 full, 54 short, and 23 posters) from 31 different countries (Table 1 provides the submissions by country).

Table 1: Distribution of Paper Submissions for ICCE 2012

Country	Submissions	Country	Submissions
Algeria	1	Malaysia	6
Australia	5	The Netherlands	1
Austria	1	New Zealand	3
Bhutan	1	Norway	2
Brazil	2	The Philippines	2
Canada	5	Samoa	1
Chile	1	Singapore	23
China	9	South Korea	5

Croatia	1	Spain	7
Finland	2	Sweden	6
Germany	2	Switzerland	3
Hong Kong	12	Taiwan	74
India	1	Thailand	1
Indonesia	2	United Kingdom	3
Japan	60	United States	3
Kuwait	2		

All papers were subjected to a rigorous review process by at least 3 reviewers from the respective sub conference program committees. After a discussion period within the individual program committees led by the sub conference Executive Co-Chairs and Co-Chairs, recommendations were made to the Coordination Committee Chair and co-Chair. They made sure that the review process for all sub conferences maintained the highest standards. This resulted in 43 full, 81 short, and 59 poster acceptances across all of the sub conferences. The overall acceptance rate for full papers is 25.3%, and the complete statistics of paper acceptances is shown in Table 2 below.

Table 2: Results of the Overall Reviewing process for ICCE 2012

Submissions / Results			Full papers	Short papers	Posters
			170	54	23
Accepted	Full Papers	43	43 (25.3%)		
	Short Papers	81	61	20 (37.0%)	
	Posters	59	28	20	11 (47.8%)

The acceptance rate for the full papers in the individual sub conferences closely mirrored the overall acceptance rate. This is a testimony to the continued maintenance of the quality of presentations in our conference. The number of submissions and the full paper acceptance rate for each sub conference is summarized in Table 3 below.

Table 3: Breakdown of Acceptance Rates by Sub conference

Sub Conference Papers	C1	C2	C3	C4	C5	C6	C7
	accepted (submitted)						
Full	5(27)	7(28)	7(23)	6(24)	5(12)	5(23)	8(32)
Short	1(5)	2(6)	1(3)	3(9)	1(4)	5(12)	7(15)
Poster	0(1)	1(2)	1(4)	3(3)	1(3)	4(6)	1(4)

Table 4 summarizes the total number of papers in different categories accepted by the seven sub conferences.

Table 4: Accepted Papers and Posters by Sub conference

Sub Conference Papers	C1	C2	C3	C4	C5	C6	C7	Total
Full	5	7	7	6	5	5	8	43
Short	15	7	9	11	5	14	20	81
Poster	8	6	5	12	6	15	7	59

A new and novel feature in ICCE 2012 is the presentation of some of the posters as e-posters using PCs connected to LCD panels. This should increase interactivity between the poster presenters and the audience, and also give the poster presenters to make demonstrations of their software during the poster session. This year 19 of the 79 accepted posters will be presented as e-posters, and the rest will be presented in the traditional poster board format. The poster sessions include 22 presentations for the Work in Progress Poster (WIPP) session, and four posters showcasing ICCE research in Indonesia, the venue for ICCE 2013.

Last, the main conference schedule includes the all-important keynote speakers: (1) Professor Yasmin Kafai from the University of Pennsylvania, USA ("Connecting Play: Learning across People, Practices, and Spaces," representing sub conference C5), (2) Professor Fu-Yun Yu from the National Chengkung University, Taiwan ("Learner-Centered Pedagogy + Adaptable and Scaffolded Learning Space Design—Online Student Question-Generation," representing sub conference C7), (3) Professor Seng Chee Tan from Nanyang Technological University, Singapore ("Perspectives of Knowledge Creation Theories and their Implications for Education," representing sub conference C2), and (4) Professor Marcelo Milrad from Linneaus University, Sweden ("Innovation and Sustainability in Mobile Learning: Current Perspectives and Challenges," representing sub conference C4); and the theme-based invited speakers: (1) Professor Kazuhisa Miwa from Nagoya University, Japan ("Learning by Creating Cognitive Models," representing sub conference C1), (2) Mr. Jon Mason from Charles Darwin University, Australia ("Opening Content for Deeper Inquiry," representing sub conference C3), and (3) Professor Jozef Colpaert from the University of Antwerpen, Belgium ("The No Significant Difference Syndrome and the Ecological Paradigm Shift," representing sub conference C6). In addition, there will be three panel sessions: (1) "Deep Scaling: Using Pedagogy Not Technology as the Basis To Scale Up School Technology Usage" (moderator: Professor Elliot Soloway from the University of Michigan), (2) "Opportunities for Publication in the RPTEL Journal" (moderator: Professor Siu Cheung Kong from the Hong Kong Institute of Education), and (3) "Practice-driven Research, Teacher Professional Development and Policy of ICT in Education" (moderator: Professor Siu Cheung Kong from the Hong Kong Institute of Education).

The first two days of the conference are devoted to pre-conference events. This year they include 14 workshops, three interactive events, one tutorial, and the Doctoral Student Consortium, which will include nine pre-doctoral student presentations followed by mentoring activities conducted by top-notch researchers. The Workshop papers are published in separate proceedings with its own ISBN number.

We would like to thank everyone who has been involved directly or indirectly in making these proceedings come to fruition, and we hope a resounding success. We have to start with all of the paper authors; we acknowledge their exciting research contributions and are delighted that they chose ICCE 2012 as the conference at which to present their work. In conjunction we have to thank the IPC and the Executive Chairs for all of the sub conferences. We gave them a lot of autonomy in making decisions, and selecting papers, and, as you will see from the proceedings they were thorough in their reviewing and selection process. We have to thank our keynote and invited speakers for graciously accepting our invitations and for their willingness to participate in all activities of the conference. Many thanks also to the panel organizers – the panels give the conference a unique flavor. Of course, the biggest thanks go to the people who have to sweat it out, and work the hardest – the Local Organization Committee. They took on the difficult work of making sure all of the papers were submitted on time, and were properly formatted for inclusion in the proceedings.

Thank you all for your commitment and hard work toward making ICCE 2012 a success. We hope that you will find the conference presentations to be insightful, interesting, and thought provoking. Please partake in the rich research atmosphere of the conference, absorb the deep insights you can gain by interacting with colleagues, observe the advancements and innovations made in Singapore educational system, and most of all enjoy the bustling cosmopolitan city and the vibrant multicultural experiences around you in Singapore.

Conference Chair
Tsukasa HIRASHIMA, Hiroshima University, Japan

Program Coordination Chairs
Gautam BISWAS, Vanderbilt University, USA (Chair)
Lung Hsiang Wong, Nanyang Technological University, Singapore (Co-Chair)

Local Organizing Committee Chair
Wenli CHEN, Nanyang Technological University, Singapore

KEYNOTE SPEAKERS

KEYNOTE SPEAKER: C5

Connecting Play: Learning across People, Practices, and Spaces

Professor Yasmin B. Kafai
Graduate School of Education, University of Pennsylvania, USA

Abstract

Play in virtual worlds provides new opportunities for social development, academic learning, and creative expression. In this talk, I propose a multi-modal examination of play to understand how youth navigated the digital public, made social connections, and learned by using their “networking residues”, the traces they leave in form of interactions, clicks, chat, messages, and designs. Following the digital footprints of youth in Whyville.net, we tracked movements and interactions of over 500 tweens for a period of six months, coupled with observations captured in video records, field notes, and interviews with some of the players in after-school clubs and science classes. The discussion will focus on the design and research of new learning analytics and opportunities in networked communities.

KEYNOTE SPEAKER: C7

Learner-Centered Pedagogy + Adaptable and Scaffolded Learning Space Design—Online Student Question-Generation

Professor Fu-Yun Yu
Institute of Education, National Cheng Kung University, Taiwan

Abstract

In response to contemporary educational approaches to teaching and learning and a drive toward learner-centered pedagogy that accentuates the concepts and importance of learners being engaged in personal knowledge construction and meaningful learning, student question-generation has gained increasing attention in the academic arena over recent decades. While the theoretical foundations of student question-generation are sound, and empirical evidence supporting its positive effects for student cognitive and affective development is solid, there are several impending issues that need to be addressed to allow its immediate adoption and long-term sustainability and scalability within existing educational systems. In this talk, first of all, how student question-generation fits within the learner-centered pedagogy umbrella, how the responsibility for learning is relayed, how the learning process and focus of learning is changed, and how the purpose and processes of evaluation is unfolded using the student question-generation approach is explained. Secondly, issues on the following will be covered: how student question-generation is facilitated, enhanced and transformed with the support of technology, how different online scaffolds are integrated to support student question-generation, and how technology enables adaptable and multiple learning spaces to be built. Finally, studies illustrating how pedagogy, theory, practice and research interplay to enlighten one another for creating a learner-centered learning space are presented.

KEYNOTE SPEAKER: C2

Perspectives of Knowledge Creation Theories and their Implications for Education

Professor Seng Chee Tan

Learning Sciences and Technologies Academic Group, National Institute of Education, Nanyang Technological University, Singapore

Abstract

The advent of the Knowledge Age has profound impact on various sectors in modern societies, including the education. The past decades saw an emergence of advocates for knowledge creation practices among school leaders and educators. This clarion call for changes in education is driven by the urgent needs to prepare students for new challenges in the 21st Century. From the learning sciences perspective, Paavola, Lipponen, and Hakkarainen (2004) extended the acquisition and participation metaphors of learning to include knowledge creation metaphor of learning. They considered three influential models of innovative knowledge creation: the Knowledge building pedagogy (Scardamalia & Bereiter, 2006), the organizational knowledge creation model (Nonaka & Takeuchi, 1995), and the expansive learning approach (Yrjö Engeström, 1999). Paavola, Lipponen, and Hakkarainen suggest that from a knowledge creation perspective, learning involves a "collaborative effort directed toward developing some mediated artifacts, broadly defined as including knowledge, ideas, practices, and material or conceptual artifacts" (, pp. 569-570). This talk presents a deeper analysis of the various theories of knowledge creation to reveal their invariant characteristics and differences. Rather than seeking a single unified theory, a multi-dimensional, multi-level framework is proposed for knowledge creation in education. This framework could serve as a guide for considering knowledge creation approach for education through various phases in our life.

KEYNOTE SPEAKER: C4

Innovation and Sustainability in Mobile Learning: Current Perspectives and Challenges

Professor Marcelo Milrad
Department of Computer Science, School of Computer Science, Physics and Mathematics, Linnaeus University, Sweden

Abstract

The emergent research field of mobile learning has gained much attention in recent years. Since the late 90's, a substantial number of research projects on mobile learning have been conducted in both formal and informal learning settings. The analysis of the research literature in this field indicate that much of the global research efforts have concentrated on the design and development of technological solutions to support learning about a particular subject matter. The outcomes of these research activities have contributed to a refined conceptualization of learning with mobile technologies in schools, museums and in field trips. However, fewer are the efforts reporting on the problems and limitations while introducing and adopting mobile technologies in schools or, on the different types of barriers or constraints the use of mobile devices faces in different educational settings.

In this talk I will address and discuss those aspects related to innovation and sustainability in mobile learning. I will first introduce this research field in terms of perspectives and challenges in order to understand why it matters to study the sustainability of innovations in schools. Thereafter, I will turn the focus to a couple of our on-going research projects in order to provide a critical analysis on the barriers and constraints experienced by both researchers and teachers while integrating mobile devices and applications in elementary and secondary schools. Towards the end, I will discuss factors, stakeholders and lines of action identified when attempting to introduce and sustain innovative educational practices in Swedish schools.

THEME-BASED INVITED SPEAKERS

INVITED SPEAKER: C1

Learning by Creating Cognitive Models

Professor Kazuhisa Miwa
Graduate School of Information Science, Nagoya University, Japan

Abstract

Computational cognitive modeling is one representative research method in cognitive science. In this talk, we propose an innovative instructional method for using cognitive modeling as a learning tool. In this context, we should distinguish learning “by” and “of” creating cognitive models (LbyCCM and LofCCM).

LofCCM is intended to learn how to build computational cognitive models in which modeling methods and computational programming are focused on. For LofCCM, we developed a web-based production system called DoCoPro that can be used anywhere and anytime in an environment connected to the Internet, and propose an instructional design, “learning through intermediate problems.” We report some case studies for demonstrating the usefulness of this framework.

For LbyCCM, we expect participants to learn the model-based thinking by creating cognitive models. In this case, the model-based thinking involves three components: theory based data interpretation, hypothetico-deductive thinking, and understanding based on mental models. In our class practices, participants were guided to interpret observed human performance data based on a psychological theory by comparing the data and the results of computer simulations performed by the model that they created by themselves (theory based data interpretation). They tried to investigate causal relations between cognitive mechanisms as a black box and the observed behavioral data by manipulating the model’s parameters (hypothetico-deductive thinking). They also learned to explain reasons of their own and others' behavior, i. e., learned to understand cognitive processes behind the behavior, by constructing a mental model underlying the behavior (understanding based on mental models).

INVITED SPEAKER: C3

Opening Content for Deeper Inquiry

Mr. Jon Mason

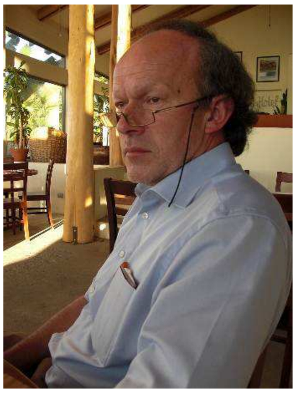
Centre for School Leadership, Learning & Development, Charles Darwin University, Australia

Abstract

Advancements in learning technologies are being driven from an increasing diversity of domains of practice and research. The “open” agenda – open architecture, open source, open standards, open access, open learning, open networks, open data, and open educational resources – is very much at the forefront of these advances for a growing international community of practice. While this agenda is valued highly in the education sector, openness is not the only driver of change or innovation with ICT. Social media continues to shape the nature of much engagement online and the late 20th century mantra that “content is king” is giving way to a fresh focus on so-called “21st century skills” and competencies where digital literacy is as important as critical thinking and problem solving. Meanwhile, discourses on sense-making and developments in knowledge management and knowledge-sharing infrastructures continue to inform the theory and practice of e-learning. This presentation acknowledges these trends and a broad range of narratives that track the evolution of e-learning as a means of contextualising a frontier ready for further technological innovation: the stimulation and support of questioning online. In particular, research into *why*-questioning is highlighted. Why? Because the semantics involved typically involve ambiguity, dialog or further inquiry. More specifically, investigation into *why*-questioning reveals that the object it seeks is *explanatory content* – and content that can be characterized as such presents a number of challenges for learning technology design.

INVITED SPEAKER: C6

The No Significant Difference Syndrome and the Ecological Paradigm Shift



Professor Josef Colpaert
Institute for Language and Communication, University of Antwerpen, Belgium

Abstract

"No technology carries on itself any inherent, measurable and generalizable learning effect. The only added value of a particular technology lies in its potential contribution to the added value of the entire learning environment as ecology." To what extent can this presenter's thirty years of cognitive and emotional friction in education, as teacher, system developer, educational designer and researcher, be formulated as arguments in such a way that they would exceed the level of idiosyncratic and haphazard observations, and that they could be used as a premises in some kind of scientific reasoning?

Inspired by the research methods in the Social Sciences, the educational research community has indeed imposed itself a stronger emphasis on *evidence* through adequate research design, focusing on methods for data gathering and analysis, and on properties such as representativeness, validity, reliability, and significance. Publishers will increasingly expect data to be linked to publications and researchers themselves become more and more in favor of an Open Data approach. The claim that a statement is true, or at least an accepted finding, should thus be substantiated by evidence. On the other hand, the extent to which this statement *really* contributes to the research community from an epistemological point of view is another matter. The word 'really' pertains to persuasive language use and should be avoided in scholarly language use. So should non-substantiated statements on aspects which we 'feel' are amenable to improvement: the amount of replication research, the generalizability of findings to other learning contexts, and the strength of the rationale behind research topic and methodology.

In recent years, the educational research community has also shifted more attention to the practitioners in the field: Practitioner-Led Research, Action Research, Reflective

Practice, Exploratory Practice, Scholarship of Teaching and Learning (SOTL), Teaching as a Design Science, Design-Based Research (DBR), just to name of few movements. These movements specify different roles for the practitioner, the researcher, the research question, the planned intervention, and for theory. This presenter's primary concern here is the mere application of theory onto practice. Empirical evidence may contribute to theory, but the question is what should happen with this theory before it is being applied to practice. The intervention itself mostly belongs to the traditional research methods. A third important phenomenon we will discuss is the recognition of the importance of the learning ecology, which puts the learning and teaching process in a broader context. However, there are not enough operational methodological frameworks for designing learning environments and for defining the role of technology in a systematic way.

These three concerns - epistemological, methodological and ecological - will serve in this presentation as beacons in our attempt to show the need for educational engineering as activity and that it can be considered as a research method. Engineering in our view, as a caveat, is less related to technology than to a specific strategy, way of thinking and contributing to our community.

TABLE OF CONTENTS

C2: ICCE Conference on Computer-supported Collaborative Learning (CSCL) and Learning Sciences

Full Paper

Short Paper

Poster

C3: ICCE Conference on Advanced Learning Technologies, Open Contents, and Standards

Full Paper

Short Paper

Poster

C5: ICCE Conference on Digital Game and Digital Toy Enhanced Learning and Society (GTEL&S)

Full Paper

Short Paper

Poster

Poster

C7: ICCE Conference on Technology, Pedagogy and Education (TPE)

Full Paper

Short Paper

Poster

Special Poster Session on Technology-Enhanced Learning in Indonesia

Development of a Dictogloss System Oriented for Focus on Form

Makoto KONDO[a*], Ryo SANO[b], Asanori TASHIRO[a], Yasuhiro NOGUCHI[a], Satoru KOGURE[a], Tatsuhiro KONISHI[a] & Yukihiro ITOH[c]
[a]*Faculty of Informatics, Shizuoka University, Japan*
[b]*Graduate School of Informatics, Shizuoka University, Japan*
[c]*Shizuoka University, Japna*
*mkondo@inf.shizuoka.ac.jp

Abstract: In this paper, we discuss how to develop a dictogloss system oriented for focus-on-form (FonF) instruction. Dictogloss is a multiple skills collaborative activity in which learners listen to a text and try to reconstruct the text in collaboration with their partners. Focus on form is a pedagogical approach aiming at improving learners' ability to produce grammatically correct sentences. A dictogloss system incorporating FonF instruction would therefore be of great help for improving learners' correct understanding and production of their target languages. We discuss what functions are necessary for a dictogloss system and how to implement the functions. Preliminary evaluation of our system shows that the system can act as a dictogloss partner for second/foreign language learners.

Keywords: Dictogloss, focus on form, second/foreign language education

Introduction

Dictogloss is a multiple skills collaborative activity proposed by Wajnyrb [15]. In a dictogloss activity, a teacher reads a short text to learners and the learners try to reconstruct the contents of the text. Dictogloss is similar to a traditional dictation activity in that both activities require learners to reproduce what they have listened to. Dictogloss however differs from dictation in some important aspects. While dictation requires learners to replicate a dictated text word for word, dictogloss encourages learners to use their own linguistic and grammatical knowledge to produce a parallel text. In other words, learners can make use of different linguistic forms from the original text as far as the reproduced text is grammatical and has the same contents as the original one. Another important difference between those activities is that dictogloss requires learners to reconstruct the original contents with their peers. Dictogloss therefore promotes collaboration among learners, which contrasts sharply with a traditional dictation activity, in which each learner is required to complete the task by him/herself.

Although there are several ways to put dictogloss into practice, a typical dictogloss activity involves the following three stages according to Izumi [3].

(1) A teacher reads a short text to learners twice at a natural speed (or plays CD-recorded sound of the text). On the first listening, the learners are instructed to grasp the gist of the text but not allowed to take any notes. On the second listening, they are instructed to jot down key words/phrases in the text.

(2) The learners make small groups and reconstruct the original contents by sharing and referring to their notes. They are instructed to speak in the target language.

(3) The reconstructed text of each group is compared with the original text by focusing on grammatical and semantic aspects of the differences.

Let us refer to those stages as (1) dictation stage, (2) reconstruction stage, and (3) analysis and correction stage, respectively.

A number of studies support the use of dictogloss in second/foreign language education [8,9,11,12,13]. Dictogloss has at least two very important properties. One is that learners use all four language skills (listening, speaking, reading, and writing) in order to complete a dictogloss activity. In the dictation stage, learners listen to their teacher read a text. In the reconstruction stage, they speak to their groupmates in the target language, and write a reconstructed version of the text. In the analysis and correction stage, they must read the original text.

Another important property is that dictogloss provides learners with opportunity to reflect on their use of the target language. In order to complete a dictogloss task, learners must reconstruct the contents of the original text. In the reconstruction, they talk about the language of the text they are reconstructing. Kowal and Swain [6,7] have found that dictogloss activities have elicited metalinguistic talks from learners; that is, learners talk about form-meaning/form-function relations in their target language.

The second property is of particular importance from the viewpoint of a pedagogical approach called focus on form (FonF). FonF has attracted much attention because it could solve a potential problem of another pervasively adopted approach called communicative approach (CA) [1]. While FonF aims at improving learners' ability to produce grammatically correct sentences, the CA puts a higher priority on conveying a speaker's intention than on making grammatically correct utterances. The CA therefore has a risk that learners would acquire incorrect grammatical rules for their target languages. FonF-based dictogloss is an effective activity to promote learners' correct understanding and production of their target languages.

Since dictogloss is a collaborative activity, it requires the presence of a partner. If a language education system plays a role of a dictogloss partner, it would be of great help to second/foreign language learners. In sections 1, 2 and 3, we discuss what functions are necessary in each of the three stages of dictogloss activities, respectively. We also discuss how we have implemented each of the functions. Section 4 shows the result of preliminary evaluations and Section 5 gives the summary of what has been achieved and some remaining issues for the future work.

1. Functions for Dictation Stage

A dictogloss system must be able to play the sound of a dictogloss text in the dictation stage. There are several ways to implement this function. The simplest way is to have sound files of dictogloss texts recorded by native speakers of target languages. A technically more interesting/challenging way is to automatically generate sound from dictogloss texts. The latter approach is beyond the scope of this paper and we have chosen the former approach.

In the dictation stage, learners listen to a dictogloss text a designated number of times; that is, they should not be able to listen to the text as many times as they like. At the same time, rewind and fast-forwarding functions should not be included in the system. Accordingly, our system has dictogloss texts and corresponding sound files; the GUI for the dictation stage only has a "play" button that can be used only a designated number of times.

One can add a note-taking function to the system; however, effectiveness of this function heavily depends on learners' typing skills. We have therefore omitted the note-taking function from our system and decided to let learners freely take notes on a sheet of paper.

2. Functions for Reconstruction Stage

In the reconstruction stage, the system must play a role of a dictogloss partner. At least two functions are necessary for the system: a reconstruction function and a dialog function. The reconstruction function generates a dictogloss partner's answer to a given dictogloss task. Let us call it a system answer. The system answer is shared by learners to complete the task. The dialog function allows the learners to collaborate with the system in reconstructing the contents of the original text. Section 2.1 and section 2.2 describes how we have designed the reconstruction function and the dialog function, respectively.

2.1 Reconstruction Function

Ideally, we want learners to notice errors in their answers by themselves through the reconstruction process with the system. In order to help them notice their errors, system answers should be generated in such a way that learners easily notice their errors by comparing their answers with the system answers. The system therefore should be capable of generating different system answers according to which part of learners' answers involves errors.

In order for the system to behave differently in accordance with learners' answers, the system must be capable of detecting errors in learners' answers. Notice that a simple matching between learners' answers and the original text does not suffice for this purpose. This is because natural language generally has more than one way to express a single semantic content. Learners' answers may be perfectly grammatical and semantically well-formed even if surface forms in the answers are quite different from the ones in the original text. In other words, the system must be able to perform semantic comparison between learners' answers and the original text. Section 2.1.1 explains how to detect errors based on semantic comparison. Section 2.1.2 describes how to generate different system answers in accordance with learners' errors.

2.1.1 Error Detection based on Semantic Comparison

We have employed JDT semantic representation [2,14], which enables the system to conduct semantic comparison. In the JDT semantic representation, meanings of content words (verbs, nouns, etc.) are represented by concept frames containing attribute-value pairs, and meanings of function words (case particles, auxiliary verbs, etc.) are represented as attributes or markers attached to frames. Dependency relations between content words are represented by pointers which link attribute values to the concept frames denoting the values. Fig. 1 shows an example semantic representation for *Hoteru-o sagashi-te* (*Find [me] a hotel*), where markers are given in square brackets. (Unlike English, Japanese allows phonetically null subjects/objects. We put them into square brackets in English translation.) For the sake of simple illustration, we omit irrelevant details throughout this paper.

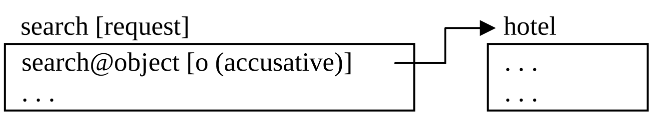


Fig. 1: Semantic Representation of *Hoteru-o sagashi-te* (*Find [me] a hotel*)

The JDT semantic representation enables semantic comparison based on attribute-value pairs. If two sentences have the same semantic content, they have the same set of attribute-value pairs in their JDT semantic representations. Consequently the JDT semantic representation enables the system to perform semantic comparison between two sentences with different syntactic structure by comparing the attribute-value pairs contained in each of the semantic representations.

Kondo et al. [5] have extended the JDT semantic representation to develop what they call situation knowledge. Situation knowledge is a set of JDT semantic representations in which concept frames denoting the same concept are integrated into one frame. Fig. 2 shows the situation knowledge associated with two sentences: Tokyo-no hoteru-ni tomari-tai ([I] want to stay at a hotel in Tokyo) and Yasui hoteru-o sagashite (Find [me] a cheap hotel). In Fig. 2, the meaning of yasui (cheap) is represented by the rate-possession frame, and the value of the rate-possession@object attribute, "- (minus)", is transferred to the value of the same attribute in the hotel frame based on the fact that they are the same attribute.

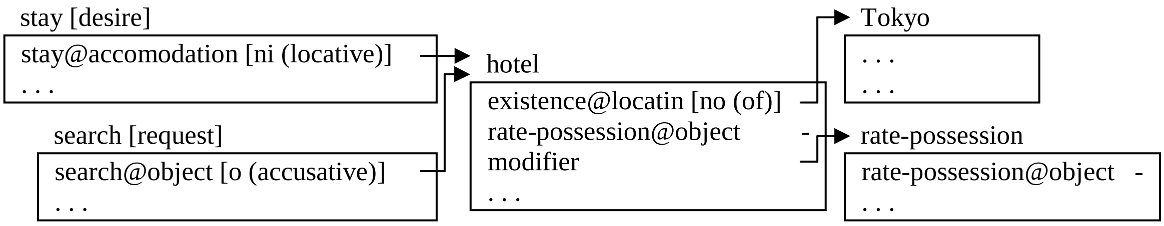


Fig. 2: Sample Situation Knowledge

We have adopted this extension of JDT semantic representation in order to represent the semantic contents of the original dictogloss text. By storing the original semantic contents in this manner, the system correctly matches learners' answers like *Yasui hoteru-ni tomari-tai* (*[I] want to stay at a cheap hotel*) and *Tokyo-no hoteru-o sagashi-te* (*Find [me] a hotel in Tokyo*) with the original semantic contents as well as those sentences in the original text.

In order to implement the reconstruction function, the system must be able to detect errors in learners' answers since the system should behave differently according to errors in learners' answers. Kondo et al. [5] and Sano et al. [10] have developed error judgment technique based on the JDT semantic representation and its extension discussed above. We have employed the technique in our system and the system can detects the following four types of errors.

(1) Erroneous omission: Learners incorrectly omit necessary linguistic forms (e.g., omission of a case particle).

(2) Erroneous addition: Learners incorrectly add unnecessary linguistic forms (e.g., addition of the past verbal suffix "-*ta*" when simple present tense is appropriate).

(3) Confusion of different linguistic forms: Learners mistake a linguistic form for another form (e.g., confusion of a progressive "verb-*teiru*" form and a perfective "verb-*tearu*" form).

(4) Incorrect word order

Basic idea behind the technique is that errors in learners' answers would result in difference between their semantic representations and the semantic representations of the original text. If a learner fails to refer to some entity, event, or property of an entity/event, the semantic representation of the learner's answer does not have the corresponding concept frame. If a learner fails to reconstruct the meaning denoted by a function word, it results in the semantic representation in which the corresponding marker is absent. If a learner's answer involves erroneous addition, the semantic representation of the learner's answer contains the corresponding concept frames/markers, which is absent in the semantic representation of the original text does not. Confusion errors are detected by combination of erroneous omission and addition. If erroneous omission and addition are detected simultaneously in the same position, the sentence should involve confusion of the omitted form and the added form. Incorrect word order is also detected by the combination of erroneous omission and addition because this type of error results in the semantic representation in which erroneous omission of a form is detected in one position and erroneous addition of the same form is simultaneously detected in another position.

2.1.2 Generation of the System Answers

While the system should generate different answers depending on which part of learners' answers involves errors, it should also behave differently according to types of forms involving errors. Since FonF instruction focuses on a few linguistic forms in a lesson, the system should behave differently according to whether errors are involved in focused forms or not. In addition to distinction between focused forms and non-focused forms, we divide non-focused forms into three categories: key words/phrases, FonF forms, and other forms. Accordingly, we divide forms in a given dictogloss text into four categories: (1) forms focused in a lesson employing the dictogloss text (focused forms), (2) key words/phrases in the text, (3) forms suitable for FonF instruction (FonF forms), and (4) other forms. Since the purpose of a FonF-based dictogloss is to improve learners' grammatical correctness in using focused forms, these are the forms which should be given the highest priority. The second priority should be given to key words/phrases in a text. Those are important in understanding semantic contents of the text. Since the very first proposal of dictogloss [15], it has been assumed that learners should be prepared for the vocabulary in a dictogloss text. Kondo et al. [4] have selected 159 FonF forms for FonF instruction. Since these forms are suitable for FonF instruction, they receive the third priority.

We have implemented the following rules for generation of system answers according to whether a learner's answer involves errors of each category mentioned above.

(1a) If a learner correctly uses a focused form, the system generates an answer involving erroneous omission of the focused form.

(1b) If a learner incorrectly uses a focused form, the system generates an answer involving a confusion error of the focused form.

(2a) If a learner correctly uses a key word/phrase, the system generates an answer involving erroneous omission of the key word/phrase.

(2b) If a learner incorrectly uses a key word/phrase, the system generates an answer involving the correct use of the key word/phrase.

(3a) If a learner correctly uses a FonF form, the system generates an answer involving the correct use of the FonF form.

(3b) If a learner incorrectly uses a FonF form, the system generates an answer involving a confusion error of the FonF form.

(4a) If a learner correctly uses one of other forms, the system generates an answer involving the correct use of the form.

(4b) If a learner incorrectly uses one of other forms, the system generates an answer involving the correct use of the form.

In (1a,b), the system always generates an erroneous answer in order to induce as much interaction between the system and the learner as possible. This is because focused forms have the highest priority in a given lesson and it is desirable to induce as much discussion on them as possible. In (2a,b), the system generates an erroneous answer if and only if the learner' answer is correct; otherwise, it generates a correct answer. This is because an erroneous system answer would induce interaction on the key word/phrase if the learner correctly uses it. At the same time, a learner's erroneous use of a key word/phrase suggests that the learner needs some help for their correct use; hence, the system should generate a correct answer. In (3a,b), the system generates an erroneous answer if and only if the learner's answer involves an error, because discussion on non-focused FonF forms is unnecessary if the learner correctly uses them. In (4a, b), the system always generates a correct answer because forms in this group have the lowest priority and discussion on them should be made as little as possible.

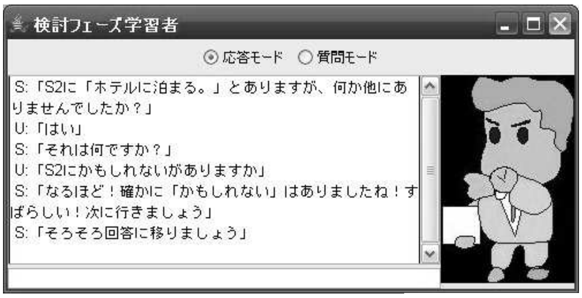
2.2 Dialog Function

In the reconstruction stage, the system and the learner collaborate to make the final answer. A learner engages in a dialog with the system through keyboard input. A leaner makes his/her question to the system or answer to the system's question using predetermined templates. Currently, the system has one question template and two answer templates. The question template is "Does S*n* have *a form*?", where "S*n*" stands for "sentence *n*" in the answer (each sentence in the answer is given a unique sentence number), and "*a form*" is replaced by any linguistic form of the learner's choice. As for answer templates, a learner can use a simple "Yes/No" answer.

The collaboration process varies depending on which participant makes an error. We must therefore consider the following four cases: (1) both a learner's answer and the system answer are correct, (2) neither a learner's answer nor the system answer is correct, (3) a learner's answer is correct whereas the system answer is not, and (4) a learner's answer is incorrect while the system answer is correct. The first case does not require any further collaboration. The forth case does not require further collaboration, either. This is because the system shows a correct answer and we can expect that the learner would notice his/her error by comparing his/her answer with the system answer. Accordingly, let us examine the remaining two cases.

When neither a learner's answer nor the system answer is correct, the system ask the user whether the sentence should have any other form: "Your Sn is '. . . (the learner's answer)'. Does it have anything else?" If the learner correctly points out the correct form by using the question template ("Does Sn have the correct form?"), the system thanks the learner and proceeds to the next sentence. If the learner's answer is "No", the system proceeds to the next sentence after expressing a doubt about the learner's answer: "Something might be wrong with the answer, but let's go on to the next sentence" or "There might be something else, but let's go on to the next sentence".

When a learner's answer is correct and the system answer is not, we can expect that the learner would point out the error in the system answer. If the learner points out the correct form by using the question template, the system thanks the learner and proceeds to the next sentence. If a learner does not initiate a dialog, the system asks the learner whether the correct form should be included in the answer. If the learner's answer to that question is positive, the system thanks the learner and proceeds to the next sentence; otherwise, the system expresses a doubt about the answer and go on to the next sentence. This strategy is implemented for focused forms alone. This case occurs if and only if the form in question is a focused form or a key word/phrase. In the latter case, a learner knows that the key word/phrase must be used somewhere in the answer. Since the system's incorrect answer in this case involves erroneous omission of the key word/phrase, we can reasonably expect that the learner would initiate a dialog. Fig. 3 shows a dialog example. In Fig. 3, the learner's answer does not involve the focused form "kamoshirenai (might)". The system detects the omission error since the marker associated with the form is absent in the semantic representation of the learner's answer. The system then produces the question "Does it have anything else?" in the first line of the example. Since the learner correctly points out, in his/her second utterance, that the answer should involve the focused form, the system thanks the learner and proceeds to the next sentence.



[Translation] (S: system, U: user/learner)
S: Your S2 is "*Hoteru-ni tomaru* (I will stay at the hotel)". Does it have anything else?
U: Yes.
S: What is it?
U: Does S2 have *kamoshirenai* (might)?
S: Yes! It certainly has *kamoshirenai*. Great! Let's proceed to the next.
S: Let's submit the final answer, shall we?

Fig. 3: Example Dialog

3. Functions for Analysis and Correction Stage

In the analysis and correction stage, the system shows the learner's final answer and the correct answer. Focused forms are highlighted in the correct answer so that the learner can confirm his/her use of the focused forms. If a learner makes an error with regard to focused forms, the system shows grammatical explanations about the forms. Four types of errors discussed in section 2.1.1 are also highlighted by giving them different font colors, and the system shows what type of error is involved in each of the sentences.

4. Preliminary Evaluation

We have implemented the functions for dictogloss activities discussed in sections 1, 2 and 3, and conducted two preliminary evaluations: an evaluation for generation of the system answers and an evaluation for the dialog function.

In evaluating the answer generation, we manually made dictogloss texts which include focused forms, key words/phrases, FonF forms and other forms. In addition, we manually made test sentences that involve erroneous omission, erroneous addition and confusion errors. Incorrect word order is not examined in this evaluation because the JDT system fails in morphological analysis of many sentences involving incorrect word order. The target of evaluation is characterized by whether a test sentence involves an error and what type of form involves an error. Since there are four types of forms, we examined eight cases: correct/incorrect use of focused forms, correct/incorrect use of key words/phrases, correct/incorrect use of FonF forms and correct/incorrect use of other forms. We examined each of the eight cases and confirmed that the system correctly generates correct/incorrect answers in accordance with the generation rules discussed in section 2.1.2.

In evaluating the dialog function, we fed the system with the above eight types of test sentences, and confirmed that the system generated appropriate questions to the learner in accordance with the dialog strategy discussed in section 2.2. The system also responded to the learner's answer to the system's question as desired. We also confirmed that the system replied to a question originally made by the learner.

5. Concluding Remarks

We have developed a dictogloss system oriented for FonF instruction. The system plays the sound of a dictogloss text in the dictation stage. A learner can listen to the sound only a designated number of times. In the reconstruction stage, the system generates its own answer in order to collaborate with the learner. The system answers vary depending on whether the learner's answer involves an error and what type of form is erroneous. The system also engages in a dialog with the learner in the reconstruction stage. The system makes questions on the answer if necessary. The learner can also make his/her own question to the system using the question template. The system changes its response to the learner's question according to whether the learner reaches the correct answer. In the analysis and correction stage, the system shows the correct answer and highlights errors in their final answer.

Needless to say, there are a lot of tasks to be completed in the future work. Since the evaluation discussed in this paper is a preliminary small-scale evaluation, we need a larger-scale evaluation with respect to both the size and variety of test sentences. One of the most important tasks to be completed is improvement of the dialog function. The current system accepts questions from a learner by using the question template. If the system takes full advantage of the JDT dialog system and accepts free input from a learner, it would surely improve the usability of the system.

Acknowledgements

This work was supported by Grant-Aid for Scientific Research (C) (21520398).

References

[1] Doughty, C., & Williams, J. (Eds.). (1998). *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.

[2] Ikegaya, Y., Noguchi, Y., Kogure, S., Konishi, T., Kondo, M., Asoh, H., Takagi, A., & Itoh, Y. (2005). Integration of dependency analysis with semantic analysis referring to the context. *Proceedings of PACLIC 19*, 107-118.

[3] Izumi, S. (2009). *"Focus on Form"-o Toriireta Atarashii Eigo Kyoiku*. Tokyo: Taishukan.

[4] Kondo, M., Daicho, Y., Sano, R., Noguchi, Y., Kogure, S., Konishi, T., & Itoh, Y. (2010). Form-wise error detection in a FonF-based language education system. *Proceedings of ICCE 2010*, 9-16.

[5] Kondo, M., Kure, U., Daicho, Y., Kogure, S., Konishi, T., & Itoh, Y. (2009). Error judgment in a language education system oriented for focus on form. *Proceedings of ICCE 2009*, 43-50.

[6] Kowal, M. and Swain, M. (1994). Using collaborative language production tasks to promote students' awareness. *Language Awareness*. *3*, 73-93.

[7] Kowal, M. and Swain, M. (1997). From semantic to syntactic processing: How can we promote metalinguistic awareness in the French immersion classroom? In R. K. Johnson, & M. Swain (Eds.), *Immersion Education: International Perspectives* (pp. 284-309). Cambridge: Cambridge University Press.

[8] Lim, W. L., & Jacobs, G. M. (2001). *An Analysis of Students' Dyadic Interaction on a Dictogloss Task*. ERIC Document Reproduction Service No. ED 456 649.

[9] Nabei, T. (1996). Dictogloss: Is it an effective language learning task? *Working Papers in Educational Linguistics*, *12*(1), 59-74.

[10] Sano, R., Noguchi, Y., Kogure, S., Konishi, T., Kondo, M., & Itoh, Y. (20109. Extension of error judgment component in educational dialog system oriented for focus on form. *Proceedings of the 35th Annual Conference of JSiSE*, 28-G1-2, 485-486.

[11] Storch, N. (1998). A classroom-based study: Insights from a collaborative reconstruction task. *ELT Journal*, *52*(4), 291-300.

[12] Swain, M., & Miccoli, L. S. (1994). Learning in a content-based, collaboratively structured course: The experience of an adult ESL learner. *TESL Canada Journal*. *12*(1), 15-28.

[13] Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*, 320-337.

[14] Takagi, A., Asoh, H., Itoh, Y., Kondo, M., & Kobayashi, I. (2006). Semantic representation for understanding meaning based on correspondence between meanings. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *10*, 876-912.

[15] Wajnryb, R. (1990). *Grammar Dictation*. Oxford: Oxford University Press.

Data-Driven Misconception Discovery in Constraint-based Intelligent Tutoring Systems

Myse ELMADANI[a*], Moffat MATHEWS[b] & Antonija MITROVIC[c]
[abc]*Intelligent Computer Tutoring Group, University of Canterbury, New Zealand*
*myse.elmadani@pg.canterbury.ac.nz

Abstract: Students often have misconceptions in the domain they are studying. Misconception identification is a difficult task but allows teachers to create strategies to appropriately address misconceptions held by students. This project investigates a data-driven technique to discover students' misconceptions in interactions with constraint-based Intelligent Tutoring Systems (ITSs). This analysis has not previously been done. EER-Tutor is one such constraint-based ITS, which teaches conceptual database design using Enhanced Entity-Relationship (EER) data modelling. As with any ITS, a lot of data about each student's interaction within EER-Tutor are available: as individual student models, containing constraint histories, and logs, containing detailed information about each student action. This work can be extended to other ITSs and their relevant domains.

Keywords: Misconceptions, data mining, constraint-based modeling

Introduction

It is common for students to make errors while solving problems. Intelligent Tutoring Systems (ITSs) usually respond very well to such errors by identifying the exact error and providing appropriate instructional feedback at varying levels, from error flags to bottoming out and providing the solution. Some of these ITSs also provide hints on what to do next in a particular step or problem in the student's solution. Both Cognitive Tutors [8] and Constraint-based tutors [10] have domain knowledge represented at an atomic level to give the most precise and helpful feedback on the incorrect step or solution state. Domain knowledge is represented in Cognitive Tutors as production rules or as constraints in Constraint-Based Modelling (CBM) tutors.

Making an error, however, could mean that the student has either got something genuinely incorrect (e.g. in the case of a novice who has faulty or incomplete knowledge) or they have made a slip (i.e. entered an incorrect answer accidentally). Long-term modelling of the student could show whether such cases are slips or genuine errors. In this paper, we are only concerned with genuine errors.

Some students fairly consistently make errors in certain larger portions of the domain (at a higher conceptual level). Glancing at the individual constraints this might not be obvious because of the low level of granularity at which constraints represent domain knowledge. Each constraint covers only one aspect of a domain concept. Such a fine level of granularity is necessary in order to provide very specific feedback on the error, so that the student can correct it. However, students do have misconceptions that cover several constraints. In constraint-based tutors, constraints are independent of each other, but subsets of constraints do cover the same domain concept. For example, student *A* violates constraint 27 every time they violate constraint 349 and 243. Even though our ITSs do a relatively good job in correcting these errors eventually via their atomic feedback, it is still pedagogically interesting to find out why these errors occur frequently together. Finding out these reasons might also help us instantiate different pedagogical strategies in these cases.

With genuine errors, students often have misconceptions at a conceptual level rather than just at a step or constraint level. These misconceptions may be due to the student having incorrect or incomplete domain knowledge involving parts of several concepts. Having misconceptions could result in the student making a number of uniquely different but similarly related errors while solving problems in that part of the domain.

Misconceptions and domain concepts are more abstract than the production rules or constraints. Many constraints, therefore, form part of a concept and a misconception results in many constraints (related to that concept) being violated. Often misconceptions are fudged within buggy rules or constraint feedback as the ITS author assumes that there must be a misconception occurring here if this step/state was done incorrectly.

It would be very useful if we could identify common misconceptions, not just at the rule/constraint level, but at a broader conceptual level, empirically, and have pedagogical strategies to deal with these known misconceptions. Normally, this is difficult, requiring human domain experts to manually observe large numbers of students working on tasks within that domain [5]. During the identification process, experts attempt to deduce misconceptions that students might have about that part of the domain by: using their expertise (e.g. knowing what is difficult in the domain); their account of each student's task-specific behaviour (e.g. errors the student made); and if available, introspective notes from the student. Once these misconceptions are identified, teaching strategies can be better aligned with those knowledge areas that are incorrect or lacking. The automation of misconception discovery would not only be of benefit for both teachers and students but would provide another level of adaptivity in ITSs.

The data-driven technique trialled in [5] provides a semi-automated method for misconception discovery. The study consisted of identifying domain misconceptions through the analysis of students' answers to multiple-choice test questions. Our project modifies the technique to elicit misconceptions in student knowledge in constraint-based tutors with more complex, ill-defined tasks. To show how this technique could be used, we implemented it in the domain of Enhanced Entity-Relationship (EER) data modelling [4] as taught to students through EER-Tutor [16, 17, 20]. This research investigates whether a list of misconceptions can be created; one that will reflect how constraints are actually used by students. The data-driven technique can then be applied to other ITSs. Having this new information about common misconceptions, the ITS will not only be able to identify that a student has a misconception and provide additional misconception-specific guidance (more than what a rule or constraint might offer) but also offers other possible applications. An example application is to create novel tasks for students, like the dynamic generation of erroneous solutions for students to correct. Such erroneous solutions are currently being created manually in other projects in the Intelligent Computer Tutoring Group.

In this paper, we discuss whether eliciting misconceptions in this manner is even possible in a constraint-based tutor within an ill-defined task (EER-Modelling) by data mining student models and logs from EER-Tutor.

1. EER-Tutor

EER-Tutor is a constraint-based ITS which provides an intelligent learning environment for students to practise and learn database design using the EER data model. Currently, EER-Tutor has 57 problems, where each subsequent problem increases in difficultly such that problem 1 is the easiest and problem 57 is the most difficult. Users create EER schemas satisfying a set of requirements which are checked for constraint violations on submission

[16, 17]. EER-Tutor records detailed session information, including each student's attempt at each problem and its outcome, as well as the history of each constraint [18, 20].

A constraint is an ordered pair <Cr, Cs>, where Cr is the relevance condition and Cs is the satisfaction condition [11]. EER-Tutor has 225 constraints which evaluate the student's solution for semantic and syntactic correctness. When checking a student's solution, the ITS uses the relevance condition to check whether a constraint is relevant for this particular solution. If it is relevant, then it checks whether the student's solution adheres to the satisfaction condition. The constraint is violated otherwise, indicating an error in the submitted schema. Appropriate corrective steps are then taken by the ITS. To learn from their errors requires that students be able to detect their errors before taking corrective steps [12]. On the basis of violated constraints, the system generates feedback, which allows the student to correct their errors.

We have large datasets of anonymised interaction data from our tutors as they are used by students worldwide. These datasets allow us to conduct data mining based research. Data collected is not from controlled experiments; the students used EER-Tutor over the Web, in a way that suited them.

2. Related Work

Educational data mining is a growing field that focuses on the analysis of large datasets of student-computer interaction logs to answer educational research questions [3]; "an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in" [7].

One such way of exploring data is to use association rules. An *association rule* represents relationships between different attribute-value pairs; stated simply, if there is an association between two attribute-value pairs (X,x) and (Y,y), then if X=x then it must also hold that Y=y. Association rule inference algorithms take a set of uniquely identifiable transactions, with each transaction being a set of attribute-value pairs that occur together [5]. An *itemset* is the set of attribute-value pairs or items. An itemset is *large* if it appears at least as many times in the transaction dataset as required by the predefined *minimum support* value [1]. Non-large itemsets are discarded and do not appear in the output association rules. Each association rule also has a *support*, a number of transactions containing the itemset and a *confidence*, the number of data instances it correctly predicts.

Current work in using a data-driven technique for semi-automatic misconception elicitation in a domain is described in [5]. This research involved the discovery of potential misconceptions by identifying the most frequent associations among incorrect answers in student solutions to multiple-choice questions. The hypothesis was that incorrect answers reflect misconceptions held by the student. The result of mining is a set of incorrect choices selected most frequently by the students. These itemsets were ordered by frequency, with the most frequent answer associations corresponding to potential misconceptions. Domain experts were required at this stage to identify which potential misconceptions could in fact be considered misconceptions. They found that students had the misconception "misunderstanding of the difference between binary and text files" for example.

The elicited misconceptions form a misconception layer (Figure 1) in the evidence model proposed in [5]. The model specifies the relationships between individual tasks, assessment activities such as test questions, and their relevant domain-specific concepts and misconceptions. This knowledge aids in more personalised and specific tutoring, such as problem selection for students based on their known misconceptions about a domain. We therefore provide an adaptation of this technique described in [5] for use with constraint-based ITSs.

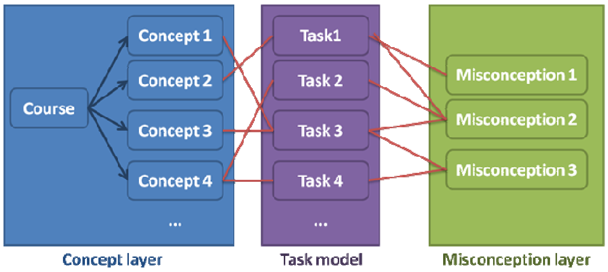


Figure 1. Relationships between misconceptions and domain concepts [5]

INFER* and MALGEN [15] are rule-based algorithms for identifying misconceptions. Using incorrect student actions, INFER* creates faulty rules. MALGEN conversely modifies existing operators to create faulty ones and tests them out. The suitability of rules generated by both algorithms is considered by human experts before inclusion in the bug library. ASSERT [2] was the first student modeling system to automatically create bug libraries. It uses theory refinement, which takes examples of student's behavior as input and modifies and creates rules if the behavior cannot be explained with the existing domain rules. MEDD [14] learns student and reference Prolog programs and uses similarity- and causality-based clustering of discrepancies between them. Each reference program has an associated error hierarchy, which is refined using the discrepancies.

3. Design and Implementation

Starting with all student models and logs collected by EER-Tutor from 2004 to 2006[1], we extracted the data about students who had attempted at least one problem. Here, an attempt is a student's solution submission, which could contain multiple steps. We decided to use all resulting 1135 student models and logs together for this project but different subsets can be processed separately in the future.

The logs were pre-processed (see Figure 2) to simplify information extraction. We extracted the required data from each pre-processed log and stored it in a database. The data is sparse because for each problem there is a number of relevant constraints, but not all constraints are relevant for every attempt at the problem. We output the data to sparse Attribute-Relation File Format (ARFF) files in order to carry out the data-mining process. There were 53,360 attempts at the various EER-Tutor problems in this data set.

RapidMiner (RM) is an open source data mining system [9]. RM provides an implementation of the Frequent Pattern-Growth (FP-Growth) algorithm [6], which generates frequent itemsets that can be used to generate association rules. The choice of algorithm differs from that in [5] because the use of Apriori with ARFF files did not allow us to treat irrelevant constraints as such within RM.

The process involved reading the ARFF files, converting all nominal values to binominal attributes and inputting all the data into the FP-Growth operator. This outputted all the frequent itemsets, which were inputs into the Create Association Rules operator to generate the association rules with a minimum confidence of 0.9. Initially the minimum support of FPGrowth was set to 0.25 as in [5], resulting in an average of 12.14 itemsets for each problem. In order to generate more itemsets, the minimum support was lowered to 0.1, with

[1] From Addison-Wesley's DatabasePlace Web portal: http://www.aw-bc.com/databaseplace/

an average of 159.2 itemsets generated per problem and ranging from 7 itemsets for problems 1, 3 and 12 to 2085 itemsets for problem 23.

Figure 2. A high-level outline of the data-driven technique described

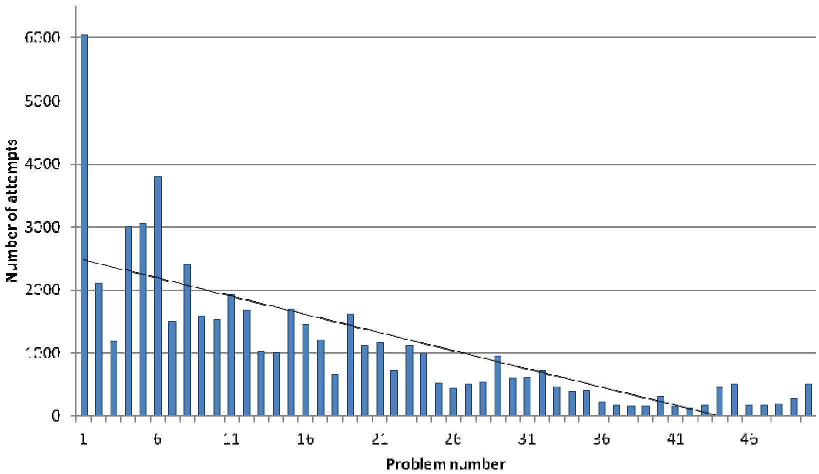


Figure 3. The number of attempts for each problem in EER-Tutor

As there was a varying number of attempts at each problem, ranging from 124 to 6045 (see Figure 3), the number of attempts required to satisfy the minimum support level also varied. A minimum support of 0.1 requires about 600 attempts in which a particular constraint is violated for problem 1, but only around 13 for problem 24, for example. Instead of controlling minimum support, RM allows us to specify a minimum number of itemsets to be generated. This returns the minimum number of itemsets with the highest support values, regardless of the minimum support value [13].

An alternative method of analysing the data set was to look at all attempts at all problems in the system at the same time, so an ARFF file containing all constraints violated during all attempts at all problems was generated as well. This was processed with the FP-Growth algorithm, with the minimum number of itemsets set to 500, resulting in the generation of 912 itemsets. Due to time constraints, the per-problem ARFF files were not reprocessed using this second mode of FP-Growth. It would be interesting to find out if there are any misconceptions that appear only in specific problems however.

The itemsets output formed a list of potential misconceptions that domain experts had to inspect in order to confirm that they were actual misconceptions. Domain experts inspected itemsets with the largest size first (see Table 1), in order of decreasing support, creating descriptions for each misconception. Both a general and a more specific description were generated for each. An example of such general descriptions is “wrong type” where the specific misconception is that “a regular entity is used instead of a weak entity”.

Table 1. The number of itemsets for all attempts at all problems

Itemset size	Number of itemsets
7	7
6	46
5	131
4	215
3	233

2	211
1	69

4. Results and Discussion

Misconception discovery is difficult, requiring human domain experts to observe large numbers of students in order to identify misconceptions [5]. By identifying which misconceptions are commonly held by students, they can be addressed in order to improve students' knowledge of a particular domain. The data-driven technique described in [5] has been adapted and implemented to identify domain misconceptions through the analysis of student-tutor interaction logs of a constraint-based ITS. Although our implementation used EER-Tutor, this method can be carried out with any constraint-based ITS.

There were a total of 912 itemsets generated by the FP-Growth algorithm, ranging in support from 0.221 to 0.01. This means that there were at least 11,793 attempts in which students violated a particular set of constraints in the first case and 534 in the latter. On average, a constraint appeared in 44 itemsets, with the highest occurrence being 238 for constraint 16. This constraint appeared as a single itemset with a support of 0.126, the equivalent of appearing in over 6700 attempts, and was found to correspond to the misconception related to the use of weak entities (with a label *weak entities are missing*).

Given that there is a total of 912 itemsets generated for all attempts at all problems, the process of labelling misconceptions is currently manual and time consuming. Because the student-ITS interaction logs are already stored and only new itemsets would need to be inspected for labelling in the future, we believe that this process is still manageable. We have started creating a hierarchy of misconceptions in the domain of EER modelling. This is beyond the scope of this paper but will be similar to the hierarchy in [19]. Because of time constraints for this project, domain experts did not inspect all itemsets. Some of the related misconceptions identified are shown in Table 2. Others included "Using simple attributes to represent regular entities" for example. The difficulty in creating the hierarchy is that violated constraints at the bottom of the misconception hierarchy could necessarily mean that the whole branch of the hierarchy is a misconception. The ITS could traverse up that branch until it finds the start of the misconception and take remedial action from that point onwards.

In some cases where two itemsets differed by a single violated constraint both corresponded to the same misconception. The most commonly occurring misconception is "using regular entities instead of weak entities". Because weak entities are introduced as early as problem 6, the results indicate that this could be a difficult concept for students to master. This misconception was identified for problem 6 in fact, with a support value of 0.306. The understanding of weak entities requires knowledge of the difference between partial and primary keys and between regular and identifying relationships among other concepts. Itemsets corresponding to this misconception could therefore be supersets of other itemsets, which correspond to related misconceptions that can also be addressed separately.

Thousands of association rules with confidence of at least 0.9 were generated. Such a rule is:

constraint21_B → constraint21_AR (confidence 0.976)

This means that when constraint 21_B (used a recursive relationship when one is not needed) is violated, we can be 97.6% confident that the constraint 21_AR (no role names defined for a recursive relationship) is also violated.

Table 2. One related list of identified misconceptions

Misconception	Support
Using regular relationships instead of identifying relationships	0.079
Using identifying relationships instead of regular relationships	0.038
Using regular entities instead of weak entities	0.023
Using weak entities instead of regular entities	0.015

Domain experts need to analyse the generated association rules to determine their usefulness. For example if the student violates constraints corresponding to the misconception that all regular entities require one key amongst the set of keys that is a primary key (i.e. *a regular entity is missing a primary key*), it may be that we could be 90% confident that they will hold the misconception that they are using primary keys to represent entities (i.e. *using a regular entity to represent a primary key*) and will violate the corresponding constraints. In this case, instead of giving atomic feedback relating to each constraint, we could give broader feedback that addresses a potential misconception by explaining both the meaning of entities and primary keys and their association.

Since we were initially analysing the original data set on a per problem basis with a minimum support value of 0.1, we also looked at how many itemsets of various sizes were generated for all attempts at all problems with the same minimum support value. It may be that there are more itemsets for the per problem files (thousands of itemsets in total) as there are duplicates or because some misconceptions are problem specific. The latter case may be due to variance in relevant constraints per problem, an average of 95 across the problems. When eliciting misconceptions for each problem, a problem might be flagged for further investigation if many itemsets are generated extraordinarily; for example, it may have a context that is too unfamiliar to students. As an example, this is what probably occurs for problem 23, where 2085 itemsets or potential misconceptions are generated when minimum support is set to 0.1 despite there being 1,112 attempts at the problem. Problem 23 describes a surveillance system database and includes some specific terminology regarding surveillance systems, which may have confused students; however these would need to be investigated further on a case by case basis.

5. Conclusion and Future Work

Misconception identification is important as it potentially aids the construction of pedagogical strategies to enhance a student's learning. Being able to semi-automate such a difficult and time consuming task allows teachers to focus on enhancements such as customised feedback and problem selection focussing on remedying misconceptions specific to a particular student. Teaching tasks can therefore be better aligned with knowledge areas that are incorrect or lacking. In this paper, we have shown that it is possible to semi-automatically identify domain misconceptions by using a data-driven technique to analyse student-tutor interaction logs of constraint-based ITSs. Students using EER-Tutor have misconceptions about the use of weak entities for example.

Although we have identified some misconceptions in order to determine that this could be done with constraint-based ITSs, the next step is to integrate this knowledge within the ITSs. For example, we can guide problem selection for students according to the specific misconceptions they hold. Further work is also required to make the data-driven technique more streamlined by investigating opportunities for automation in the developed system. In addition, the labeling of misconceptions and the creation of the misconception hierarchy is still a time-consuming and challenging job at this stage, and we can explore improvements such as developing tools to help domain experts. For example, a tool can be developed that displays only a subset of itemsets to the experts for inspection based on factors including similarity between itemsets. This naturally extends to providing some way of presenting the generated association rules to determine whether any are useful, perhaps in guiding feedback provided by the ITS.

Acknowledgements

This research would not have been possible without the support of other members of the Intelligent Computer Tutoring Group at the University of Canterbury, New Zealand.

References

[1] Agrawal, R., & Srikant, R. (1994) Fast algorithms for mining association rules. In Bocca, J.B., Jarke, M., Zaniolo, C. (eds.) *Proceedings of 20th Int. Conf. Very Large Data Bases VLDB*, *1215*, 487-499.

[2] Baffes P., & Mooney, R. (1996) Refinement-based Student Modeling and Automated Bug Library Construction. *J. of Artificial Intelligence in Education*, *7*(1), 75-116.

[3] Baker, R., & Yacef, K. (2009) The state of educational data mining in 2009: A review and future visions. *Educational Data Mining, 1*, 3-17.

[4] Elmasri, R., & Navathe, S. (2007) *Fundamentals of database systems*. Addison Wesley.

[5] Guzman, E., Conejo, R., Galvez, J. (2010) A data-driven technique for misconception elicitation. In De Bra, P., Kobsa, A., Chin, D. (eds.) *User Modeling, Adaptation, and Personalization, Lecture Notes in Computer Science*, *6075*, 243-254. Springer Berlin / Heidelberg.

[6] Han, J., Pei, J., Yin, Y., & Mao, R. (2004) Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Mining and Knowledge Discovery*, *8*, 53-87.

[7] International Working Group on Educational Data Mining (2009) Educational data mining. http://www.educationaldatamining.org/

[8] Koedinger, K., Anderson, J., Hadley, W., & Mark, M. (1997) Intelligent tutoring goes to school in the big city. *Artificial Intelligence in Education*, *8*, 30-43 (1997)

[9] Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., & Euler, T. (2006) Yale: rapid prototyping for complex data mining tasks. In: *Proc of the 12th ACM SIGKDD Int. Conf. Knowledge discovery and data mining*, 935-940. ACM, New York, NY, USA

[10] Mitrovic, A. (2012) Fifteen years of constraint-based tutors: what we have achieved and where we are going. *User Modeling and User-Adapted Interaction*, *22*(1-2), 39-72.

[11] Ohlsson, S. (1992) Constraint-based student modelling. *Artificial Intelligence in Education*, *3*(4), 429-447.

[12] Ohlsson, S. (1996) Learning from performance errors. *Psych. Review*, *103*, 241-262.

[13] Rapid-i: Rapidminer 4.4 class documentation. (2009) http://rapid-i.com/api/rapidminer4.4/com/rapidminer/operator/learner/associations/fpgrowth/FPGrowth.html

[14] Sison, R., Numao, M., & Shimura, M. (2000) Multistrategy Discovery and Detection of Novice Programmer Errors. *Machine Learning*, *38*, 157–180

[15] Sleeman, D., Hirsh, H., Ellery, I., & Kim, I. (1990) Extending domain theories: two case studies in student modeling. *Machine Learning*, *5*, 11–37.

[16] Suraweera, P., Mitrovic, A. (2002) Kermit: A constraint-based tutor for database modeling. In Cerri, S., Gouardres, G., Paraguacu, F. (eds.) *Proc. Intelligent Tutoring Systems, LNCS*, *2363*, 377-387. Springer Berlin / Heidelberg.

[17] Suraweera, P., & Mitrovic, A. (2004) An Intelligent Tutoring System for Entity Relationship Modelling. *Int. J. Artificial Intelligence in Education*, *14*(3-4), 375-417.

[18] Weerasinghe, A. & Mitrovic, A. (2005) Supporting deep learning in an open-ended domain. In Gh. Negoita, M., Reusch, B. (eds.) *Real World Applications of Computational Intelligence, Studies in Fuzziness and Soft Computing*, *179*, 4-75. Springer Berlin / Heidelberg.

[19] Weerasinghe, A., Mitrovic, A., Martin, B. (2008) A Preliminary Study of a General Model for Supporting Tutorial Dialogues. *16th International Conference on Computers in Education*, 125-132.

[20] Zakharov, K., Mitrovic, A., & Ohlsson, S. (2005) Feedback Micro-engineering in EER-Tutor. In C-K Looi, G. McCalla, B. Bredeweg, J. Breuker (Eds.) *Proc. Artificial Intelligence in Education*, IOS Press, 718-725, Amsterdam.

Experimental Evaluation of Kit-Build Concept Map for Science Classes in an Elementary School

Kouta SUGIHARA[a*], Takuya OSADA[a], Tsukasa HIRASHIMA[a], Hideo FUNAOI[b] & Shinsuke NAKATA[c]
[a]*Graduate School of Engineering, Hiroshima University, Japan*
[b]*Faculty of Education, Soka University, Japan*
[c]*Science Research Department, Attached Elementary School of Hiroshima University, Japan*
*sugihara @lel.hiroshima-u.ac.jp

Abstract: We have already proposed Kit-Build Concept Map (we call it as "KB map") that is a framework to realize automatic diagnosis of concept maps built by learners and to give feedback to their errors in the maps. In this paper, we report an experimental evaluation of KB map in science learning in an elementary school. Two classes of fourth grade students in an elementary school attended the experiment. "Waxing and Waning of the Moon" in science class was the topic of this experiment. The results suggested that the diagnosis and feedback with the KB map has positive effect to improve learner's understanding of this topic.

Keywords: Kit-Build, Concept Map, Automatic Diagnosis, Goal Map, Learner Map, Group Map, Segmentation and Construction Tasks

Introduction

Many investigations have reported that a concept map is a promising tool to promote learners learning effectively [1-6]. Automatic diagnosis of concept maps built by learners is one of the most important issues to realize useful interaction between learners and their teacher. We have already proposed Kit-Build Concept Map (KB map) as a framework to promote interaction between them based on automatic diagnosis of concept maps [7,8]. This paper reports the results of practical use of KB map for science learning in an elementary school.

In KB map, the task to make a concept map is divided into two sub-tasks: 1) "segmentation task" where parts of the concept map are extracted and 2) "structuring task" where the extracted parts (kit) are integrated into a map. In the framework of KB map, an ideal concept map (goal map) is prepared by a teacher or an expert as the goal of his/her lecture, and parts are generated by decomposing the goal map. The parts are provided to learners, and then the learners build concept maps (learner maps) by connecting the parts. Since the same parts are used both in the goal map and learner maps, it is easy to compare them and detect their differences. The differences from the goal map are errors in the learner map. This detection of the differences is the automatic diagnosis of KB map. Besides, by overlaying several learner maps of a group, a "group map" that expresses the understanding of the group is generated. By comparing the group map with the goal map, the group map can be also diagnosed.

Several researchers have tackled the problem of automatically diagnosing concept maps and interaction between teachers and learners through the concept maps. Basic approach to realize the diagnosis is to compare the learner's and the teacher's concept maps. Some of the investigations have addressed the automatic diagnosis and paid special attention to handle the cases where learners have misspelled a label of a concept or they have used a synonym or a related concept to the appropriate one, based on techniques of natural language processing and knowledge based reasoning [4, 5]. Betty's Brain [6] has ability to simulate a map built by a learner and the results of simulation promote the learner to improve the map. Although these researches have advance diagnosis abilities of concept maps, it is necessary to prepare knowledge base or simulation function for a subject domain. Such preparation is usually very difficult for usual teachers. Therefore, it is not easy for teachers to use the concept maps with such advance diagnosis. Then, the notations of the concept maps are usually complex because it is necessary to describe the map precisely in order to realize automatic diagnosis. Therefore, they often require high ability for learners.

KBmap is a kind of simplified concept map for a teacher, learners and a system. A teacher is required to describe a map that is composed of nodes and links with labels as the goal of his/her teaching to a class. The teacher is allowed to decide the labels of nodes and links depending on his/her teaching session. This means that the teacher is not required to keep specific notation and is allowed to use his/her word. Therefore, interpretation of the goal map depends on specific teaching/learning context. Although it is not easy to interpret the map in general, it can be expected that it is possible and useful for learners who learned in the specific context. As for diagnosis of a system, although meanings of the differences detected by comparing the goal map and the learner maps cannot be interpreted, each difference itself indicates a defect of his/her understandings and it is possible for the teacher and learners to find the meaning of the difference because they share the same context.

In order to examine whether KBmap works as expected, we used KBmap practically in science learning in an elementary school. The learning topic is "waxing and waning of the Moon" and subjects are fourth grade students in two classes. Through this practical use, we have confirmed that (1) teachers were able to build the goal map themselves (2) learners were able to build their map by using provided parts and they thought the building activity was useful for learning, and (3) feedback for learners by means of the group map was useful to improve their understanding.

1. KBmap System for "Waxing and Waning Moon"

In this section, we introduce implementation of KBmap system for subjects of "waxing and waning Moon".

1.1 KBmap System

We have already developed a system based on the framework explained in the previous section. This system is called as "KBmap System". It is a web application with two client systems: "KBmap Editor" and "KBmap Analyzer", and a server system: "KBmap DB". KBmap Editor provides an environment to make a goal map, a kit, and a learner map. This system has been implemented by Java (version 1.6). KBmap Analyzer has functions to gather learner maps online, generate a group map and diagnose the maps. This system has implemented by Flash and supports version Flash Player 10. KBmap DB has a function to store and share maps. This system was developed by Ruby (version 1.8.7) on Rails (version

1.2.3) and MySQL (version 5.1.30). Flow of teaching/learning with KBmap System is explained in the next section.

1.2 Flow of Teaching/learning with KBmap System

The flow of teaching/learning with KB map is composed of following four phases: (1) goal map building by a teacher as a part of preparation of learning material, (2) teaching with the learning material, (3) learner map building by learners as reflection of their learning, (4) diagnosis of learner maps by the teacher, and (4) feedback and complementary teaching to learners based on the results diagnosis. In the goal map building, a teacher or an author of a learning material is required to build a goal map as an ideal concept map that learners would make after their learning activity. Figure 1 shows a typical figure that is used to teach the cycle of “waxing and waning Moon”. Figure 2 shows the goal map that was made by a teacher in the practical use of KB map reported in this paper. In this map, "Waxing Moon" or "New Moon" corresponds to a node word and "Set" corresponds to a link word. A set of components (it is called "kit") is generated by decomposing the goal map, as shown in Figure 3. In the phase of leaner map building, a learner is required to build a concept map (learner map) with the kit. Figure 4 shows an example of learner map building by KBmap Editor. Because a map is composed by connecting links between nodes, all errors are detected as mistakes in link connection. In Figure 4, “rising” is not used in the learner map. We call this error as “leaving link”. Since “setting” link from “waxing moon” to “evening” in the learner map does not exist in the goal map, this error is called as “excessive link”. Then, “crossing the meridian” link from “waxing moon” to “evening”. We call them as “lacking link”.

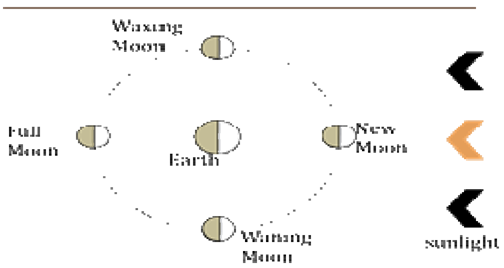


Figure 1: Waxing and Waning of the Moon

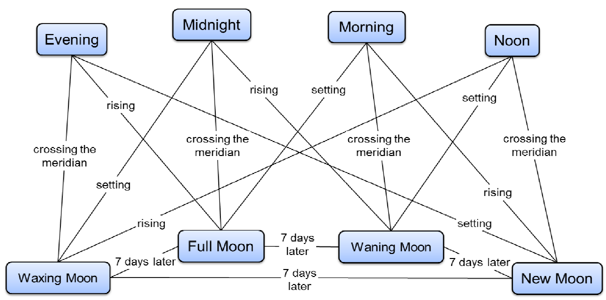


Figure 2: Goal Map

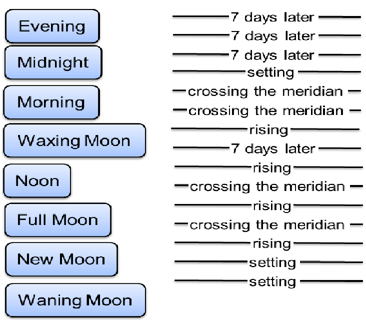


Figure 3: Kit of the Goal

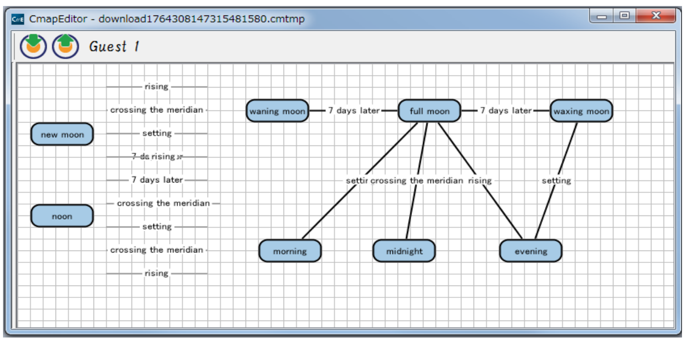


Figure 4: Learner Map building by KBmap

By overlaying several learner maps, it is easily to generate a group map describing understanding of a learner group. A link in the group map has a "overlaid degree" that is a ratio of learner maps including the link. By comparing the group map with the goal map, the weak points in understanding of the topic of the group are detected. Figure 5 shows group map where a link connected by many learners is marked with bold line. A thin line means that very few learner maps include the link. Thickness of link changes gradually by the number of students that connected the link. By comparing the group map with the goal map, “group difference map” is generated as show in Figure 6. In the difference group map, a solid link expresses a lacking link and a broken link expresses an excessive link or a leaving link. In this difference group map, usually many errors are detected. Therefore, KB map system has ability to arrange the overlaid degree for visualization of errors in the group map. In the case of Figure 6, there are four error links with high overlaid degree are shown. Bold solid link “rising” means that many learners couldn’t connect it to any nodes.

Diagnosis of learner maps is carried out as generation of the group map and group difference maps in the framework of KB map. Several types of feedback are able to realize based on the diagnosis. In the experimental use of KB map in this paper, a teacher detected high frequent mistakes from a group difference map and gave learners feedback to remedy the mistakes.

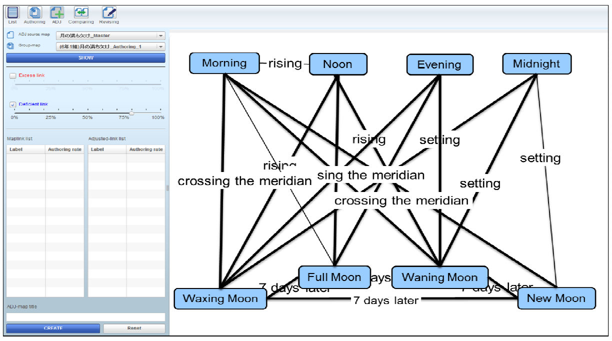


Figure 5: Group Map.

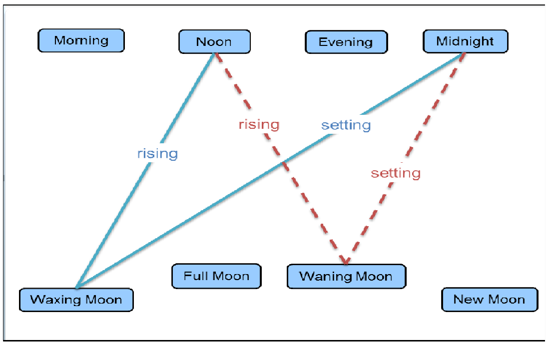


Figure 6: Group Difference Map

2. Experimental Use

2.1 Outline of the use

In this use, seventy-two 4th grade students (9 or 10 years old) in two classes attended. All of them had taken three lectures (45 minutes per a lecture) of “waxing and waning of the Moon” before two weeks of the use of KBmap system. A day before the use, they took a test for measuring their understanding about the topic for thirty minutes (this is a pre-test). At the first use of KB map, they build their respective maps of “waxing and waning of the Moon” with KBmap system for thirty minutes in a class after the introduction of the operation of KB map. The map build in this class in called “first map”. A week later of the use, a teacher in charge of the two classes gave a feedback lecture. For one class, the teacher gave feedback with KB map (details are explained in the next subsection). In contrast, for another class, the teacher gave feedback with the results of the pre-test with explanation of the questions and errors. This is a typical feedback for exercises or examinations in a usual class. Here, we call the former class “map feedback class” and the latter class “usual feedback class”. Then, a few days later, all of them built the maps of “waxing and waning of the Moon” again with KBmap system. This map is called “second map”. In the same time, they answered questionnaires. A post-test that was the same one with the pre-test was carried out two weeks later of the map use. All of the above lectures were conducted the teacher in charge of the two classes. In the classes of the use of KBmap system, four or five assistants (including authors) participated in the classes to help learners to use the system. In this use, another teacher also attended as an observer. The two teachers were science teachers who took charge of all science lectures of the elementary school.

2.2 Maps Built in the Use

Figure 2 shows a goal map prepared by the teacher. This goal map was prepared by a teacher in charge of the subjects and another science teacher accepted it. Although the responsible teacher said to prepare the goal map was not easy. he also commented that the preparation of the goal map was reasonable task as a part of preparation of learning materials. Moreover, the teacher made the goal map commented that to make the map was useful to arrange the learning material. Then, another teacher commented that to check the goal map was useful to understand the lecture. Figure 4 shows examples of learner map. A group map is shown in Figure 5. In the feedback lecture with KB map, the teacher used a group difference map as shown in Figure 6. Because the group difference map itself is too complex to show learners directly, gradual visualization function is used. At first, the teacher took on the highest ratio link and showed the learners only it. Then, the teacher requested the learners to think about whether the link was correct or not. In this case, "rising" link between "midnight" and "waning moon" was the excess link with the highest frequency. After gathering several opinions from learners, the teacher explained correct answer. Then, the teacher moved to the second excess link. Figure 6 shows in a snapshot of this gradual visualization of the group difference map. In this figure, four links with the highest ratio are shown. They are as follows: "waxing moon-rises-noon", "waxing moon–sets–midnight", "waning moon-rises-noon", "waning moon–sets–midnight". As the ratio, more than 20% learners wrongly connected the four links.

2.3 Analysis of Map Scores and Test Scores

We have categorized learners into "map feedback" group and "usual feedback" group. The map scores were obtained before and after the feedback lecture. Former one is the "first map score" and the latter one is the "second map score". The learners are also categorized into "high score" group and "low score" group by the average score of the first maps. The test scores were obtained as the pre-test and post-test. As for the test scores, the learners are also categorized into "high score" group and "low score" group by the average score of the pre-test. The test was composed of eleven questions that could not be solved only by the information explicitly described in the goal map. The test is a little advanced but not special one. Similar type of test is often used to confirm understanding of "waxing and waning the Moon" in elementary schools.

2.3.1 Map Scores

A score of each learner map is calculated by counting correctly connected links by comparing the goal map. In this paper, the score is expressed as ratio of correct links, so full mark is "1". Because of a few absent learners in the map building classes, 67 pairs of maps scores (first & second) were used in this analysis. Their average scores are shown in Table 1. Average first map scores of the whole class is 0.65 (SD = 0.29). Changes of map scores in each category of learners between the first map and the second map are shown in Figure 7.

Table 1: Map Scores

	First Map Score	Second Map Score
Map Feedback (n = 35)	0.71(SD = 0.27)	0.89 (0.16)
High Score (n = 21)	0.89 (0.10)	0.93 (0.12)
Low Score (n = 14)	0.43 (0.19)	0.82 (0.18)
Usual Feedback (n = 32)	0.60 (0.30)	0.67 (0.30)
High Score (n = 12)	0.94 (0.07)	0.90 (0.14)
Low Score (n = 20)	0.39 (0.16)	0.53 (0.29)

The results of map scores were analyzed with a three-way 2 (map feedback or usual feedback) x 2 (high score group or low score group) x 2 (first map or second map) mixed ANOVA, multiple comparison was made using Ryan's method. As the results, the secondary interaction between map/usual, high/low and first/second was marginally significant ($F(1,63) = 3.471$, $p < 0.1$). As for the simple-simple main effect of "map/usual feedback" factor, there is a significant difference in the maps scores in "low score & second map" ($F(1,126)=21.9$, $p < 0.001$) although there are not significant differences in "high score" and "low score & first map". As for the simple-simple main effect of "low/high score" factor, there are significant differences in "usual feedback" and "map feedback & first map" (for all of them, $p < 0.001$), and there is marginal difference in "map feedback & second map" (p = < 0.1). As for the simple-simple main effect of "first/second map" factor, there are significant differences in "low score" (for all of them, $p < 0.001$), and then, there are no significant differences in "high score".

As the results in statistically, the low score learners both in "map feedback" and "usual feedback" improved their maps although the high score learners didn't. Besides, "map feedback" was more effective than "usual feedback" to improve their map scores. Effect size between the first map score and the second map score in "map feedback & low score" condition is $d = 3.05$ (extra-large) and effect size between the first map score and the second map score in "usual feedback & low score" is $d = 0.87$(large). Then, the effect size between the second map scores of "map feedback & low score" and "usual feedback & low score" is $d = 1.74$ (extra-large).

2.3.2 Test Scores

The pre-test and post-test are the same one and its full mark is eleven. The average scores and standard deviations are shown in Table 2. Because of a few absent learners in the map building classes, 71 pairs of test scores (pre and post) were used in this analysis. Average of pre-test score is 6.3 (*SD* = 2.4). Changes of test scores in each category of learners between the pre-test and post-test are shown in Figure 8.

Table 2: Test Scores

	Pre-Test	Post-Test
Map Feedback (*n* = 35)	5.94 (*SD* = 2.12)	9.00 (2.12)
High Score (*n* = 12)	8.32 (0.85)	8.25 (2.55)
Low Score (*n* = 23)	4.70 (1.27)	9.39 (1.74)
Usual Feedback (*n* = 36)	6.7 (2.7)	7.3 (2.3)
High Score (*n* = 22)	8.41 (1.40)	8.41 (1.83)
Low Score (*n* = 14)	4.00 (1.77)	5.43 (1.72)

The results of map scores were analyzed with a three-way 2 (map feedback or usual feedback) x 2 (high score group or low score group) x 2 (pre-test score or post-test score) mixed ANOVA, multiple comparison was made using Ryan's method. As the results, the secondary interaction between map/usual, high/low and pre/post was significant ($F(1,67) = 11.379$, $p < 0.01$). As for the simple-simple main effect of "map/usual feedback" factor, there is a significant difference in the test scores in "map feedback & low score" ($F(1,134)=43.68$, $p < 0.001$) although there are not significant differences in "usual feedback" and "map feedback & high score". As for the simple-simple main effect of "low/high score" factor, there are significant differences in "usual feedback" and "map feedback & pre-test" (for all of them, $p < 0.001$), and there is marginal difference in "map feedback & post-test" ($p < 0.1$). As for the simple-simple main effect of "pre/post test" factor, there are significant differences in "map feedback & low score" ($p < 0.001$) and "usual feedback and low score" (p < 0.01). Then, there are no significant differences in "high score".

As the results, the low score learners both in "map feedback" and "usual feedback" improved their test scores although high score learner didn't. Besides, "map feedback" was more effective than "usual feedback" to improve their test score. Effect size between the pre-test score and the post-test score in "map feedback & low score" condition is d = 3.09 (extra-large) and effect size between the pre-test score and the post-test score in "usual feedback & low score" condition is d = 0.82(large). Then, the effect size between the post-test scores of "map feedback & low score" and "usual feedback & low score" is d = 3.06 (extra-large).

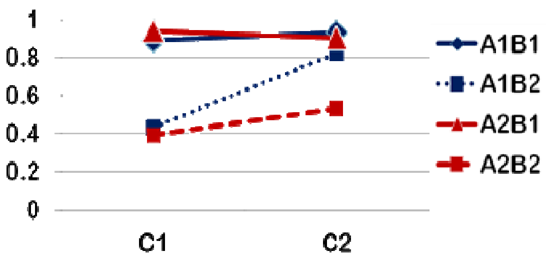


Figure 7: Map Scores

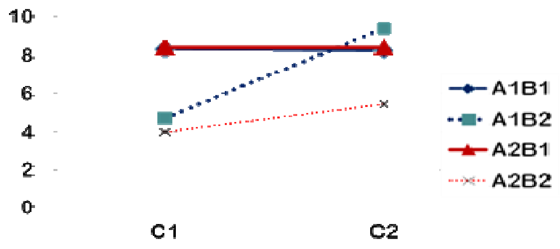


Figure 8: Test Scores

{A1: Map feedback, A2: Usual feedback, B1: High score, B2: Low score, C1: First/Pre, C2: Second/ Pre}

2.3.3 Correlation between the Map and Test Scores.

In order to evaluate the relation between the map scores and the test scores, we carried out Pearson correlation test on 137 pairs of "first map score & pre-test score" or "second map score & post-test score". As the results, there is a statistically significant positive correlation ($r = 0.45$, $p < 0.001$).

2.4 Questionnaires

Table 3 shows the results of a questionnaire for learners. As a result, most of the students agreed that map building is a useful and enjoyable.

Table 3: Results of Questionnaires for Learners

	4	3	2	1
(1) Did you enjoy map building?	65	6	0	0
(2) Was map building useful to learn the Moon?	55	15	1	0
(3) Do you like to use the map for other subjects	63	8	0	0

4 = strong agree, 3 = agree, 2 = disagree, 1 = strong disagree

Table 4 shows the results of a questionnaire for the two science teachers. One is a teacher in charge of the two classes, and the other attended all map building classes as an observer. The results of the analysis were shared before to answer this questionnaire. As a result, the teachers accepted KB map as a useful tool for learning.

Table 4: Results of Questionnaires for Teachers

	4	3	2	1
(1) KB Map Building was useful for learning	2	0	0	0
(2) Prepared Kit positively effect for learning	1	1	0	0
(3) Map feedback is more useful to correct mistakes than usual feedback	2	0	0	0

(4) Map feedback is more useful to motive learners than usual feedback	1	1	0	0
(5) KB map is better than scratch-build concept map in the learning of the Moon.	2	0	0	0

4 = strong agree, 3 = agree, 2 = disagree, 1 = strong disagree

2.5 Consideration

The KB map visualized basic knowledge of the learning topic and all parts were taught in the lecture explicitly. Therefore, building the map would be a reasonable and promising activity to confirm what the learners learned by themselves. This would be a reason that KB map was accepted for most of the learners and teachers as useful learning tool. In the map feedback, errors of learners were indicated in the map. Because map expression is near to learner's inner knowledge expression, the indication would be more direct one and promote learner's knowledge correction. Because the learners of the low score group received more indications of their errors, they could improve their knowledge more than the learners of high score group.

As for high score learners who could not improve their scores, the two teachers has guessed that KB map would give positive effect for their understanding even though it couldn't be measured. To find a method to measure the effect for high score learners is one of the most important issues for our research group.

3. Conclusion

Through this experimental use, we have confirmed that (1) teachers could build the goal map following their teaching, (2) the learners could build KB maps, (3) the learners and teachers accepted KB map as a useful tool for learning, and (4) diagnosis and feedback with KB map were effective to support the learning. Based on these results, we have judged that KB map worked as expected in this learning topics. Because the two teachers also have accepted the usefulness of KBmap, we are collaboratively working together about KB map in order to extend the target domain, find convenient & effective way to use the results of diagnosis of learner maps, and investigate the details of the learning process with KB map.

References

[1] Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. *New York, NY:* Cambridge University Press.

[2] Barenholz, H., & Tamir, P. (1992). A comprehensive use of concept mapping in design, instruction and assessment. Science and Technological Education, 10, 37–52.

[3] Ozdemir, A. (2005). Analyzing Concept Maps as an Assessment (Evaluation) Tool in Teaching Mathematics. *Journal of Social Sciences,* 1(3)141-149.

[4] Chang, K., Sung, T., & Chen, S-F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning,* 17(1), 21-33.

[5] Gouli, E., Gogoulou, A., Papanikolaou, K. A., & Grigoriadou, M. (2005). Evaluating Learner's Knowledge Level on Concept Mapping Tasks. *Fifth IEEE International Conference on Advanced Learning Technologies,* 424-428.

[6] Krittaya Leelawong, Gautam Biswas. (2008). Designing Learning by Teaching Agents: The Betty's Brain System. *International Journal of Artificial Intelligence in Education 18* (2008) 181-208.

[7] Kazuya Yamasaki, Hiroyuki Fukuda, Tsukasa Hirashima and Hideo Funaoi: Kit-Build Concept Map and Its Preliminary Evaluation, Proc. of ICCE2010, pp.290-294(2010),

[8] Hirashima, T., Yamasaki, K., Fukuda, H., Funaoi, H. (2011). Kit-Build Concept Map for Automatic Diagnosis , Proc. of AIED2011, Lecture Notes in Computer Science, 2011, Volume 6738/2011, 466-468.

A Case Study of Learning by Problem-Posing in Introductory Phase of Arithmetic Word Problems

Sho YAMAMOTO[a*], Takehiro KANBE[a] , Yuta YOSHIDA[a] , Kazushige MAEDA[b] & 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: We have developed an interactive environment for learning by problem-posing targeting arithmetic word problems that can be solved either by addition or subtraction. In our previous work, the environment was used by third or fourth grade students who have already acquired ability to solve the targeted problems. Therefore, problem-posing was an additional practice for the students and the purpose of the learning with the environment was sophistication of their ability. In this paper, practical use of the environment for the first grade students is reported. Just after the classes of problem solving of the arithmetic word problems, the way of problem-posing itself was taught in classes, and then, the environment was used as exercise of the problem-posing. Through this practice, we have confirmed that (1) the first grade students were able to pose problems in the environment, and (2) the practice to pose problems improved their ability not only in problem-posing but also in problem-solving.

Keywords: Problem-posing, Sentence-integration, Teaching method, Problem structure, interactive learning environment

Introduction

Design and practical use of a teaching method in which learners learn problem structures of arithmetic word problems through problem-posing is described in this paper. Learning by problem-posing is well known as an important way to promote learners to master the use of solution methods [1, 2]. Several researchers have already suggested that understanding the problem structure is important to solve arithmetic word problems and poor problem solvers often fail to elicit problem structures from the problems [3-5]. We have continuously investigated technology-enhanced learning by problem-posing in arithmetic word problems and practically used a developed learning environment (we called the environment as “MONSAKUN” (problem-posing boy in Japanese)) for fourth and third grade students in an elementary school [6, 7]. In these projects, although we have defined problem structures of several types of arithmetic word problems, the structures were only used by MONSAKUN in order to diagnose learners’ problem-posing or design a series of problem-posing exercise.

Based on the practical uses of MONSAKUN for four years, we (including responsible teach of the elementary school where we have used MONSAKUN) planed to teach the problem structures used in MONSAKUN and to carry out exercises of problem-posing with MONSAKUN. The participants were the first grade students just after they learnt problem-solving of the word problems. Through the practical use, we have confirmed that (1) the first grade students were able to pose problems in the environment, and (2) the practice to pose problems improved their ability not only problem-posing but also problem-solving.

In previous researches of MONSAKUN, students who have already acquired ability to solve the targeted problems were subjects. Therefore, problem-posing was an advance practice for the students and the purpose of the learning with the environment was sophistication of their ability. In contrast, the subjects of this practice were the first grade students who had classes of problem solving of the arithmetic word problems just before this practice. In this practice, the way of problem-posing itself was taught by a teacher, and then, problem-posing with MONSAKUN was used as exercise to operate the structures.

In this paper, in the next sections, the problem structures and their classification are explained. Teaching method of the problem structures with MONSAKUN used in this practice is explained. MONSAKUN used in this practice were able to be used with a tablet PC and wireless LAN. We call this version as "MONSAKUN Touch". This improvement is indispensable to realize this teaching method in a usual classroom. This improvement is also explained in Section 2. In Section 3, procedure of practical use of MONSAKUN Touch and analysis of the results are described.

1. Problem Structures Used in MONSAKUN

1.1 Problem Structures

We have proposed a model to describe problem structures of arithmetic word problems that are solved by one operation of either addition or subtraction. From viewpoint of calculation, the word problems include two given numbers and one required number. By operating with the two given numbers, the required number is derived. In the model, therefore, a word problem is composed of three sentences, that is, two sentences express two given numbers and one sentence expresses one required number. Then, every word problem has a cover story. In many investigation of arithmetic word problems indicated that there are four types of cover story, (1) increase-change, (2) decrease-change, (3) combine, and (4) compare [8]. These cover stories express one numerical relation between two numbers. The relation corresponds to an operation, that is, addition or subtraction. For example, in increase-change story, there is one number at first, and then, a number that is added to the first number is shown. At last, the number after the addition is shown. For example, in case that "Tom has 3 pencils" is the first sentence, "Tom buys 2 pencils" is the second sentence, and "Tom has 5 pencils" is the last sentence, the three sentences form one increase-change story. Then, the numerical relation in the cover story is 3+2=5.

In MONSAKUN, we have expressed each type of story by using two "existence sentences" (corresponding to the first and the last sentence in the above example) and one relation sentence (to the second sentence). We call the series of sentences as "cover story". A problem is specified by the location of required number. In case of the above example, there are three problems is included in the cover story. For example, if the number included in the first sentence is set to required number, a problem that includes numerical relation expressed as "?+2=5" can be generated. The problem, then, can be solved by "5-2". The equation expressing the numerical relation included in the problem is called "story operation structure", and the calculation that is used to derive the required number is called "calculation operation structure".

In MONSAKUN, a problem is composed of three sentences. Then, a problem is categorized by (a) cover story, (b) story operation structure, and (c) calculation operation structure. Figure 1 shows several existence sentences and relation sentences. By using the Sentence-1, Sentence-5 and Sentence-3 in this order, a problem that is "combine" in cover story, “5+?=8” in story operation structure, and “8-5” in calculation operation structure is generated.

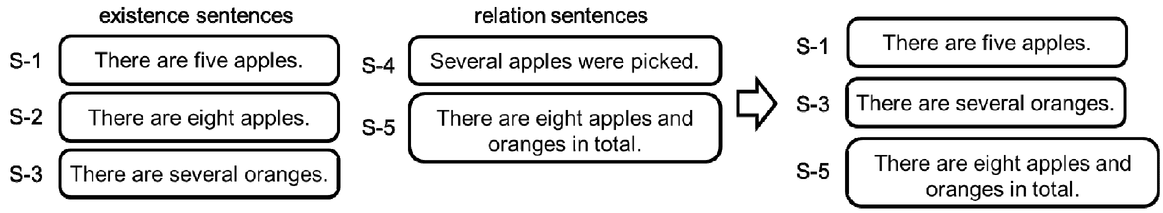


Figure 1. Order of Simple Sentence (Combine Problem)

1.2 Task Model of Problem-Posing

We have already proposed a task model of problem-posing based on the expression of the problem structure [9]. The model is shown in Figure 2. Problem-posing task is divided into four tasks (1) decision of calculation operation structure, (2) decision of story operation structure, (3) decision of cover story and (4) decision of three sentences. A learner should complete these tasks to pose a problem correctly though the execution procedure of the tasks is not decided in the model. So it is important to understand the task model of problem posing for problem solving because this task model means the relations between the problem structures to constitute a problem. In MONSAKKUN, difficulties of problem-posing are interpreted in the task model.

If operator (+ or -) of calculation operation structure is the same one with the calculation operation structure, understanding the cover story is almost same as solving the problem. We call such a problem as “forward-thinking problem”. Then, if operator (+ or -) of calculation operation structure is not the same one with the calculation operation structure, it is necessary to transform the story operation structure to calculation operation structure after understanding the cover story. We call such a problem as “reverse-thinking problem. Because the learner is more required to comprehend the relations between two structures, the reverse thinking problem is more difficult than forward thinking problem.

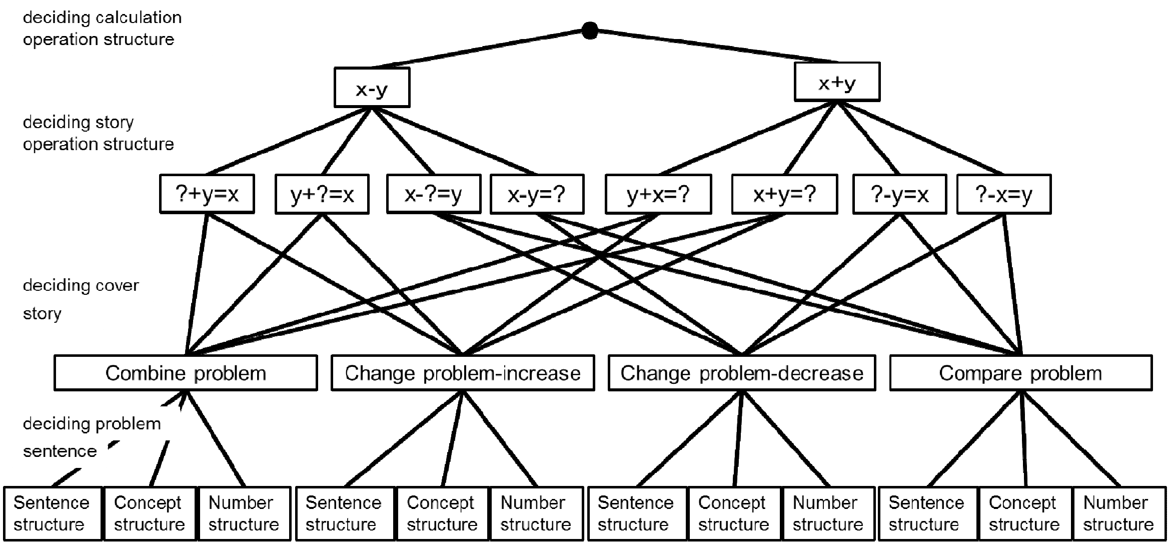


Figure 2. The Model of Problem-posing

2. Teaching Method with MONSAKUN

2.1 Teaching by a Teacher

We have designed a teaching method based on the problem, and we have developed MONSAKUN based on this structure. The teacher teaches the problem structure on the black board by using several sentence cards that are parts of problems. The teacher also prepares several cover story name cards, story operation cards, and calculation operation cards. These cards are provided to the learners as a request of problem posing. For example, the teacher requested learners to pose problems that can be solved by a specific calculation operation. The teacher lets the students pose the problem which will be solved by the prepared calculation expression and story by selecting several sentence cards and arranging them in a proper order.

In this process, the teacher teaches the students the problem structures described in Section 2. This teaching consists of five contents; (1) simple sentence is composed of an object or event, countable attribute and a value of the attribute, (2) problem is composed of two existence sentences and one relation sentence, (3) cover story, (4) calculation operation structure, and (5) story operation structure. (1) and (2) correspond with the problem sentences of the model shown in Figure 2. Firstly, the teacher presents one simple sentence card to students from prepared cards, and he/she teaches the elements of simple sentence. This teaches contents (1). Secondly, the teacher presents the students to one simple sentence card from prepared cards one after another. Then, the learner answers whether presented simple sentence card is necessary to pose a problem or not. They are also made to answer about the reason why a card is necessary. They are also made to answer about the reason why a card is necessary, and the teacher explains a problem structure based on their answer. Through this teaching, the students understand the following; (2) problem is composed of two existence sentences and one relation sentence, (3) a sentence representing each story, a relations among them, and proper order of simple sentences in each story, (4) calculation expression to represent a story directly, and (5) calculation expression to find an answer. The relations among these structures are also taught.

2.2 Interactive Environment for Learning by Problem-posing as Sentence Integration

We have used MONSAKUN at an elementary school for four years. However, MONSAKUN could be used only in a computer room because previous version of MONSAKUN was implemented on the desktop PC platform. In this practice, the responsible teacher hoped to let learners not only to use MONSAKUN as exercise but also to receive lectures of problem structure as usual classes. Therefore, we have implemented MONSAKUN on tablet PC platform so that the teacher was able to use it in the usual classroom. We named it MONSAKUN Touch. In this learning environment, the learner selects the difficulty of problem-posing task before carrying out a problem-posing exercise. After that, a learner is presented the problem-posing area shown in Figure 3 to a learning environment. The area on the left side is problem composition area. At the top, a calculation expression and story is given. Several sentence cards are presented at the right side of the interface. The learner poses the problem by moving a simple sentence card with a finger and putting a card into blank. When a learner finished posing problem, he/she can push a diagnosis button under the problem composition area. Then the system diagnoses the combination of sentences, and shows the results of the diagnosis and message to help the learner's problem-posing on another window. Then the system diagnoses the combination of sentences, and gives messages to help the learner's problem-posing on another window. The messages composed of two kinds of indications, one is indication of correct or incorrect of the posed problem and the other is indication of wrong cards. Former indication is called Flag Feedback and the latter one is called Pointing Hint [10].

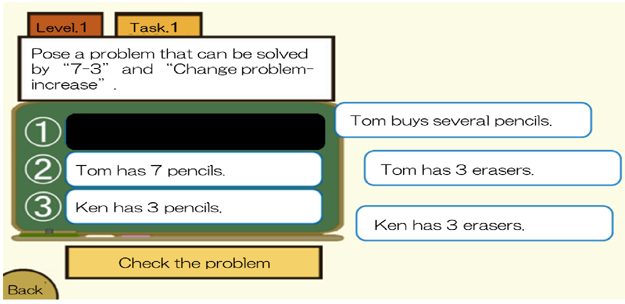


Figure 3. Interface of MONSAKUN

3. Practical Use of MONSAKUN Touch and Teaching Method

3.1 Procedure of Practical Use

The subjects of our experiment are 40 students in the first grade of an elementary school (one student absent from the pre-posttest and questionnaire). The arithmetic word problems are usually taught on the first grade of elementary schools, but the problem structures are not taught explicitly. In this practice, the problem structure of arithmetic word problems used in MONSAKUN Touch were taught explicitly and carried out problem-posing exercise with MONSAKUN Touch as exercise to operate the problem structure. This practice used nine lesson times (45 minutes per lesson, 3 weeks, 9 days). Students took the pretest before the period, and took a posttest and questionnaires after the period. Each test took 45 minutes. Problem-posing exercises divided into 6 levels. Contents of each level are shown in Table 1. The levels categorized by (1) forward-thinking or reverse-thinking, (2) story operation stricture given or calculation operation structure given, and (3) cover story. In a level, students were required to pose problems following provided story operation structure or calculation operation structure and cover story. Cover stories were excerpted from several textbooks. Also, if the student finishes problem-posing exercise in a level in a class, he/she repeats the same level exercise.

In this practical use, students used the MONSAKUN Touch as an introduction of new level problem-posing (5 min) at the beginning of a class. The students, then, are taught the problem structures by the teacher on blackboard (35 min). Finally, they used the MONSAKUN Touch as confirmation of teaching (5 min). The teacher has taught the problem structures and its relations by using the teaching method explained in Section 2.

In pre- and post-test, we used the same problem solving test and problem-posing test. Problem solving test used to assess the students problem solving performance. In problem-posing test, the students are required to pose four problems by composing several sentence cards provided beforehand. This test is used to examine the student's problem-posing performance.

Table 1. Level that Implemented by MONSAKUN

Level	Number of task	Kinds of problem	Kinds of problem-posing task	Story structure
1	12	forward thinking problem	story operation structure	combine· increase· decrease· prepare
2	3	forward thinking problem	story operation structure	combine-increase
3	12	reverse thinking problem	story operation structure	combine· increase· decrease· prepare
4	3	reverse thinking problem	story operation structure	combine-increase
5	12	reverse thinking problem	calculation operation structure	combine· increase· decrease· prepare
6	12	random		

3.2 Analysis of Log Data, Questionnaire and Students Remark

Figure 4 is the rate of correct problems that were posed on MONSAKUN Touch in each class. Vertical axis shows the rate and number of correct problems. Horizontal axis shows the days of practical use and the level that correspond to it. And the number of students that finished posing problems in each level is shown in Table 2. The students performed level 1 and 2 during the 3rd day from the 1st day, level 3 and 4 during the 6th day from the 4th day, and then, level 5 at the 8th day. The teacher has taught the problem structure corresponding to level 5 in detail in the 7th day. The task in level 5 is very difficult for learners, because it requires them to pose reverse-thinking problems from calculation operation structure. Then, problem-posing with MONSAKUN was not carried out in the 7th day and took almost double times for the exercise on the 8th day. In Figure 4, a rate of correct problems increased between 1st and 2nd days and between 4th and 5th days. But a rate of correct problems decreased between 2nd and 3rd days and between 5th and 6th days because the students worked on the new problem-posing task respectively. A rate of correct problems decreased sharply 8th day because the students were required to pose reverse-thinking problems from calculation operation problems as a task in level 5. These results suggested that teaching method about the task to present story operation structure was effective for understanding of forward thinking problem and reverse thinking problem. But it is necessary for teaching method about the task to present calculation operation structure to be improved.

The results of the questionnaire are shown in Table 3. Almost all students agreed that problem-posing exercise by using MONSAKUN and effective to learn, but, we supposed, because of level 5, many students answered the problem-posing is difficult. The teacher agreed that it is easy to teach problem-posing using a tablet PC in the general classroom, and he said that he want to use the MONSAKUN in his class. But, also he suggested that it is necessary to improve the sentence of feedback and to expand the kinds of feedback.

Through this teaching method, the student was required to explain not only by using the block but also by using the problem structures and its relations. For example, the student said that, this problem story is increase so the order of the simple sentence card is decided. Also, they indicated the problem has story operation structure and calculation operation structure. They called story operation structure "story expression" and called calculation operation structure "calculation expression". These results suggested that this teaching method for problem structure using MONSAKUN Touch was effective to learn problem structures and its relation.

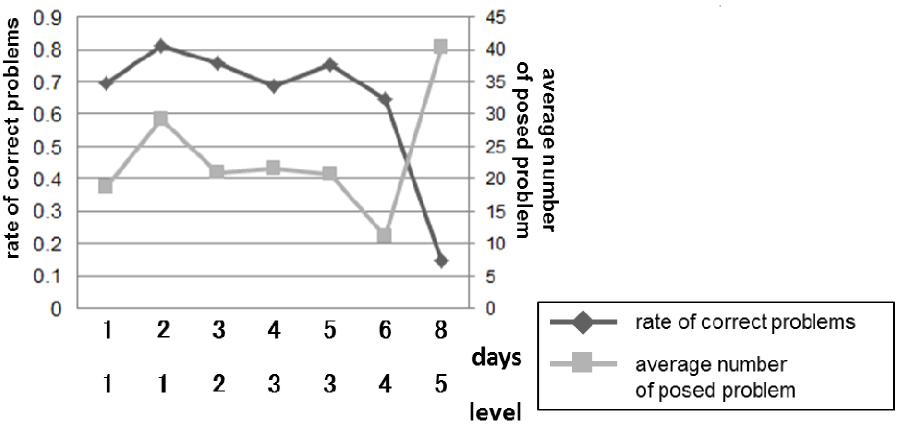


Figure 4. Rate of Correct Problems

Table 2. Number of Students that Finished Posing Problems

Level	1	2	3	4	5	6
Number of students	39	39	39	38	39	23
Number of not finished students	3	1	11	0	17	16

Table 3. Results of Questionnaires

	Strongly Agree	Agree	Disagree	Strongly Disagree
1.Do you enjoy posing problems in arthmetic?	35	3	0	0
2. Are arithmetic problems easy to pose?	8	7	19	4
4.Do you think that posing problems made it easier to solve problems?	20	17	1	0
7.Would you like to attend arithmetic classes where problem posing is used?	36	2	0	0

3.3 Analysis of the Pre- and Post-test

The results of pre- and post-test are shown in Table 4 and Table 5. And the scene of using MONSAKUN Touch is shown in Figure 5. The full marks of problem-posing test are 4. The problem-solving test is composed of 9 forward-thinking problems and 8 reverse-thinking problems. So, the full marks of problem-solving of forward-thinking problems are 9 and the full marks of problem-solving of revers-thinking problems are 8. In the scores of problem-solving test shown in Table 4, there was a significant difference in the scores between pretest and posttest of reverse thinking problems (two sided p-values from Wilcoxon matched-pairs signed-ranks test with correction for ties, p=.009), and effect size is medium ($|r|$=.45). These results suggested that explicit teaching of problem structures was effective to understand the reverse thinking problem. In problem-posing test, there was a significant difference in the between pre-test and post-test as for the number of correct problems at reverse thinking problems (two sided p-values from Wilcoxon matched-pairs signed-ranks test with correction for ties, p=.0006), and effect size is medium ($|r|$=.39). In contrast with this, the number of correct problems at forward thinking problems decreased. These results suggested that the students would be aware of the difference between the reverse thinking problems and forward thinking problems. Based on these results, we have judged that this teaching method with MONSAKUN Touch is a promising way to teach arithmetic word problems.

Table 4. Results of Problems Test (*1% significant)

		forward thinking problem	reverse thinking problem
pre-test	M	8.82	7.13*
	SD	0.6	0.65
post-test	M	8.71	7.66*
	SD	0.39	1.28

Table 5. Results of Problem-posing Test (*1% significant)

	Number of problem-posing	Number of correct problems		Number of wrong problems
		forward thinking problem	reverse thinking problem	
pre-test	3.72	1.54	0.74*	1.44
post-test	3.87	1.44	1.44*	1

Figure 5. Scene of Using MONSAKUN

4. Concluding Remarks

In this paper, we have described the practical use of the learning environment for the first grade students in introductory phase of arithmetic word problems. For using the interactive environment for learning by problem-posing in the general classroom, we have developed a learning environment can use on tablet PC, and designed the teaching method. Then, in this class, the teacher taught problem structures and its relations that were implemented in environment, and the environment was used as exercise of the problem-posing. Through this practice, we have confirmed that the first grade students were able to pose problems in the environment, and the teaching and using learning environment are improved their ability not only problem-posing but also problem-solving in the reverse thinking problem. Also, it is accepted by students and teachers as an effective teaching method. As our future works, monitoring of learners' problem-posing behavior and detection of their errors aiming at remedial feedback for their problem posing is one of the most important issues. Sophistication of the task model of problem-posing and evaluation of learning effect of the teaching method with MONSAKUN is also important future works.

References

[1] Polya, G. (1945). How to Solve It. Princeton University Press.

[2] NF Ellerton. (1986). Children's made-up mathematics problems: a new perspective on talented mathematicians. Educational Studies in Mathematics, Vol.17, 261-271.

[3] Brown, J.S., VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. Cognitive Science, 4, 379-426.

[4] Mayer, R. E. (1982). Memory for algebra story problem. Journal of Educational Psychology, 74, 199-216.

[5] Kintsch, W., Greeno, J.G. (1985). Understanding and Solving Word Arithmetic Problem. Psychological Review, 92-1, 109-129.

[6] Hirashima, T., Yokoyama, T., Okamoto, M., Takeuchi, A. (2006). A Computer-Based Environment for Learning by Problem-Posing as Sentence-Integration. Proc. ICCE2006, 127-130.

[7] Kurayama, M., Hirashima, T. (2010). Interactive Learning Environment Designed Based on A Task Model of Problem-Posing. ICCE2010, 98-99.

[8] 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.

[9] Hirashima, T., Kurayama, M. (2011). Learning by Problem-Posing for Reverse-Thinking Problems. AIED 2011, 123-130.

[10] VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M. (2005). The Andes Physics Tutoring System: Lessons Learned, International Journal of Artificial Intelligence in Education, 15(3), 1-47.

Ontological Descriptions for Eye Movement Data and Mental States in Taking Computer-based Multiple-Choice Tests

Keiichi MURAMATSU[a*], Kazuaki KOJIMA[a] & Tatsunori MATSUI[a]
[a]*Faculty of Human Sciences, Waseda University, Japan*
*kei-mura@ruri.waseda.jp

Abstract: Recently, the research in intelligent educational systems has much interest in exploring data from educational settings to understand learners behavior and mental states. To further understand learners based on such the data and sophisticate supportive interventions by educational systems, knowledge of relationships between learners' behavior and mental states must be shared. To propose a framework to share the knowledge, this study attempted ontological descriptions for learners' eye-movement and mental states in solving computer-based multiple-choice problems. The current study forms a technical basis for development of an IMS (Intelligent Mentoring System) in which an automatic mentoring function is implemented with ITS (Intelligent Tutoring System).

Keywords: Ontology, mental state, eye movement, computer-based test

1. Introduction

Recently, data such as operation logs, face images, eye movements and various physiological indices has been analyzed to recognize situation and state of learner who uses e-learning system. Broad-ranging analyses of the data can provide understandings of learners' mental states in addition to their knowledge structure. For example, affective states such as confidence or confusion [1, 7], unusual states such as impasses or illusion [2, 13] and subjective difficulties for problems [9] are tried to be recognized.

In these studies, information of eye movement is one important data generally used and is useful for realizing human mental processes. Due to higher sampling frequency than time interval of operation [9] and required time for learning [13], such data from eye movement enables to detect learners' mental states in more detail [6, 12]. On the contrary the disadvantage of focusing on that information is that device for the measurement of eye movement is not available in ordinary learning environment. However, knowledge of the eye movement is helpful in understanding of learners' mental states, when it is considered that fixation position can be estimated from face image obtained through stereo-camera [10] which has the potential to become a common device.

Knowledge of learners' behavior such as eye movement is helpful for building intelligent system which supports learning from both aspects of knowledge and mental states. We named such the system as Intelligent Mentoring System (IMS). One of its main characteristics is diagnostic function of learner model considering mental states of learners. Because mental states can instantly change in a short activity (e.g., solving of a single problem), IMS is required to monitor learners at all time and give feedback based on diagnosis. The IMS provides integrative learning-support including real-time estimation of learners' mental states and selection of ways to support learners, in addition to diagnosis of learners' knowledge structures and determination of teaching strategies provided by ITS (Intelligent Tutoring System). Figure 1 shows skeleton framework of IMS.

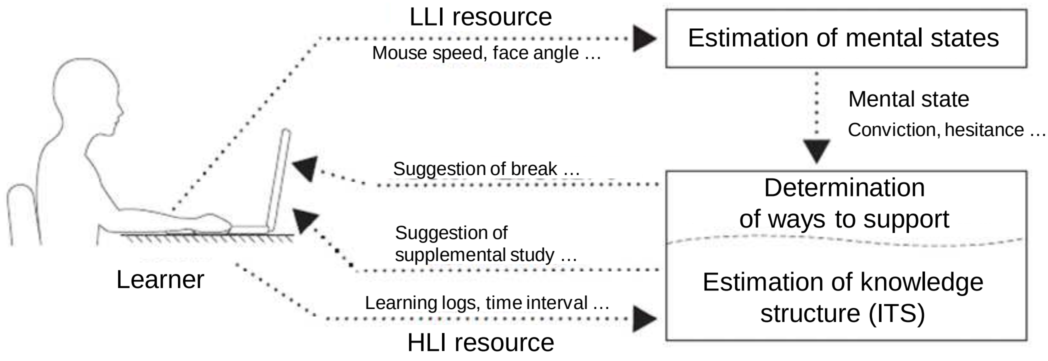


Figure 1. Skeleton framework of Intelligent Mentoring System

Interactions between user and system can be captured in levels of cognitive activity; high-level interactions (HLI) and low-level interactions (LLI). The HLI is explicitly accompanied with user's awareness and is consequently illustrated by data resource which is sampled in large grain sizes. On the other hand, LLI is not always accompanied with user's awareness and is consequently illustrated by data resource which is sampled in very smaller grain sizes. For example, learning logs, time interval of operation [9], and required time for learning [13] can be the former data resources, and moving speed of a mouse, face or posture angle of learner, and gaze position [10] or eye movement of learner can be the latter data resources. Focusing on LLI resources, IMS aims at estimation of learners' mental states from their unarticulated and semi-conscious behaviors.

In e-learning environment, operations of learners who take a multiple-choice test with mouse click are regarded as HLI resource, and behaviors of learners who gaze problem and selection statements are regarded as LLI resources. IMS function to determine ways to support learners is required to integrate information about knowledge structure estimated from HLI resources (this corresponds to function of ITS) and mental states estimated from LLI resources.

In the current study, our purpose is to develop framework to integrate the information and provide consistent description of knowledge implemented in IMS. We focus LLI resource and mental states of learners, because the knowledge about relationships among them is task-oriented and independent from knowledge structures of specific study domains. For this purpose, we build an ontology for description of relationship between eye movement of learner and associated mental states during taking multiple-choice tests as an example of the knowledge. First, we introduce experiment to obtain eye movement data. Second, we show ontological descriptions of multiple-choice test event in which eye movement of learner and associated mental states are positioned.

2. Gaze on Multiple-Choice Tests

2.1 Experiment to Obtain Eye Movement Data

We conducted an experiment to obtain eye movement data of learners during taking multiple-choice test and subjective measurement data about their mental states. Mental states of the learners are extracted from questionnaires. In this experiment, participants were asked to answer thirty four-choice tests which require only encyclopedic knowledge. To observe learners' behavior to read texts and retrieve answers, we created these tests which do not require learners of mental processing and reasoning.

In this experiment, the tests and questionnaires were displayed on a PC monitor full-screen by a program implemented by the second author. First, a button labeled "*proceed next*" was displayed in a position where a problem statement appeared. After clicking of a mouse button on the displayed button, a four-choice test was appeared (Figure 2). When the mouse cursor entered in areas of selection statement, the statement turned red. The red statement was selected as the response with clicking the mouse button on it. Second, questionnaires about the test were displayed after selection of the response. The first questionnaire asked whether the participants respond the selection that they had judged to be the answer. The second asked how the answer was familiar by selecting one of "I know it, I could answer without the choices (*recall*)", "I remembered the answer from the choice (*recognition*)", "I don't know, but guessed from the choices (*guesstimate*)", or "I had no idea (*no-idea*)". The third asked how the each selection was close to the answer. Every selection was evaluated with "it's definitely the answer/ not the answer", "it's probably the answer/ not the answer", or "I cannot judge".

Five undergraduate students participated in this experiment. The participant was seated in front of a desk where a PC monitor (the resolution was 1280 * 1024 pixels) was set up and was asked to answer four-choice tests with a mouse as quickly and correctly as possible. After answering of each test, s/he responded to the questionnaires. As practice task, s/he answered one test and its questionnaires prior to answering the thirty problems. During the task, each participant's eye movements were recorded with EMR-AT VOXER manufactured by NAC Image Technology Inc. Eye movement data included values of x and y coordinates on the screen (pixels) which are sampled at 60 frames/sec.

Which auditorium does
Waseda University have?

1. Okuma auditorium
2. Toyota auditorium
3. Yasuda auditorium
4. Kanematsu auditorium

Figure 2. Example of four-choice test displayed on a PC monitor

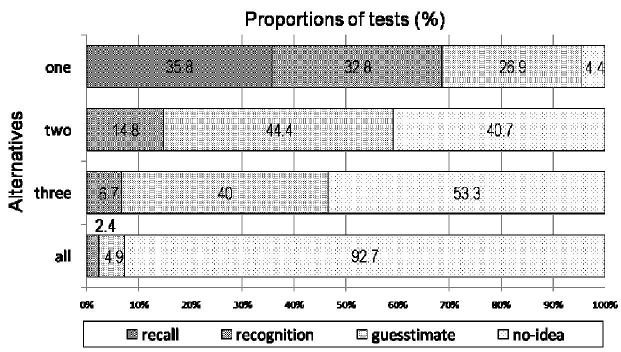


Figure 3. Proportions of tests in each category of alternatives

2.2 Categorization of Responses to Tests

According to the number of response definitely not to be the answer in the third questionnaire, possible selections in each test was decided as alternatives. For example, one alternative remains when three selections were evaluated not to be the answer, and all alternatives remain when nothing was evaluated not to be the answer. Categories of response to tests were decided by combining the alternatives and familiarities of the second questionnaire. Figure 3 shows the proportions of these categories.

However, eye movement data of two participants was removed in analysis because of deficits by mechanical problems. After error processing and smoothing, eye movement data for 58 tests used in analysis. Then the categories were made as follows; 14 tests were one/recall, 12 tests were one/recognition, 6 tests were one/guesstimate, 6 tests were two/guesstimate, 4 tests were two/no-idea and 16 tests were all/no-idea.

2.3 Gaze Patterns

We analyzed gaze patterns visualizing eye movement data. The problem statement was positioned at 96-256 pixels in y-axis direction and selection statement from 1 to 4 were positioned 352-440 pixels, 496-584 pixels, 640-728 pixels and 784-872 pixels respectively. Therefore, plotting values of y coordinate against frames indicates participants' gaze pattern during taking a test.

We found three typical gaze patterns. Figure 4 (a) shows an example of a pattern looking-ahead at selection statement. This pattern was observed in 15 tests and indicates that the participant looked-ahead at selection during reading problem statement. Figure 4 (b) shows an example of a pattern looking-back at problem statement. This pattern was observed in 16 tests and indicates that the participant looked-back at problem statement while s/he scans all selection statements. Figure 4 (c) shows an example of a pattern looking-back at selection statement. This pattern was observed in 11 tests and indicates that the participant looked-back at the previous selection statement while s/he scans all selection statements.

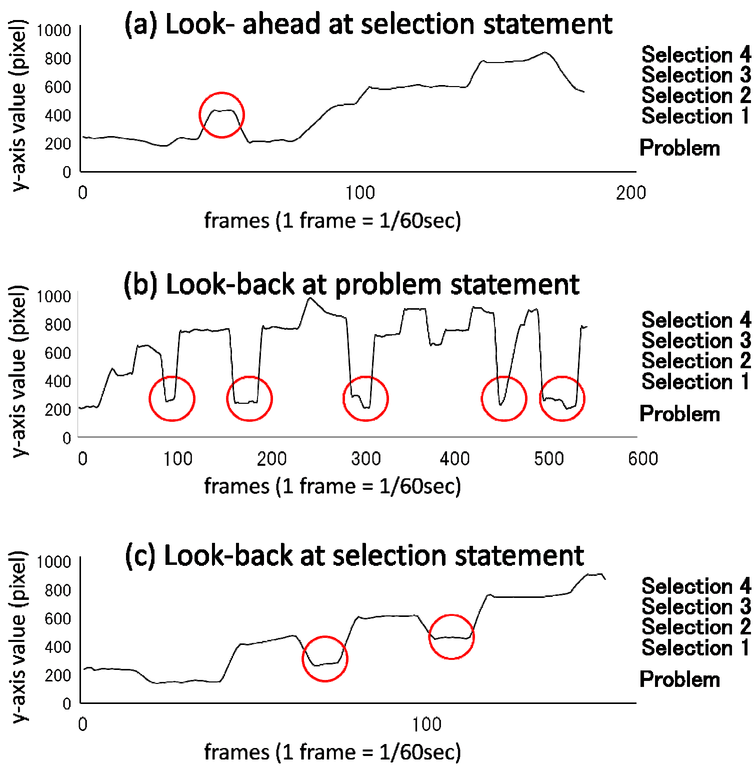


Figure 4. Typical gaze patterns

Table 1. Proportions of gaze patterns and categories of response

	look-ahead at selection	look-back at problem	look-back at selection	cases
one/recall	21% (3)	29% (4)	21% (3)	14
one/recognition	25% (3)	42% (5)	25% (3)	12
one/guesstimate	17% (1)	50% (3)	17% (1)	6
two/guesstimate	34% (2)	1 7%(1)	34% (2)	6
two/no-idea	0	0	0	4
all/no-idea	38% (6)	19% (3)	13% (2)	16

2.3 Relationships between Eye Movement and Mental States

Proportions of gaze patterns and categories of response are summarized as Table 1. Proportion and number of cases which are observed to be each gaze pattern are shown by row. The categories of familiarity based on the second questionnaire indicate learners' mental state called conviction, and the number of alternative is interpreted to lead learners' hesitance in selection.

According to this summary, the pattern of "look-ahead at selection" marked high value of proportion at two/guesstimate and all/no-idea. Thus, this pattern seems to be closely-linked to low conviction and high hesitance. The pattern of "look-back at problem" marked high value of proportion at one/recognition and one/guesstimate, so that this pattern seems to be closely-linked to middle conviction and low hesitance. Finally, the pattern of "look-back at selection" marked high value of proportion at only two/guesstimate, and so this pattern seems to be closely-linked to middle conviction and hesitance.

3. Ontological Descriptions

3.1 Ontological Engineering

Ontological engineering is one of methodologies to describe knowledge systematically. From knowledge-based viewpoint, "ontology is defined as a theory (system) of concepts/vocabulary used as building blocks of an information processing system" [4]. In ontology development environment Hozo[1], each node represents a whole-concept and has some slots which represent part-of or attribute-of relations. Hozo supports describing role concepts which represents a role depends on contents of each whole-concept. For example, a teacher role which is played by a human only in a context of school, and he does not play the role out of the school. In other words, every slot has a role under whole-concept implying a context. In the context, a class of instances which can play a role is defined by a class constraint, and it is called role-holder [3].

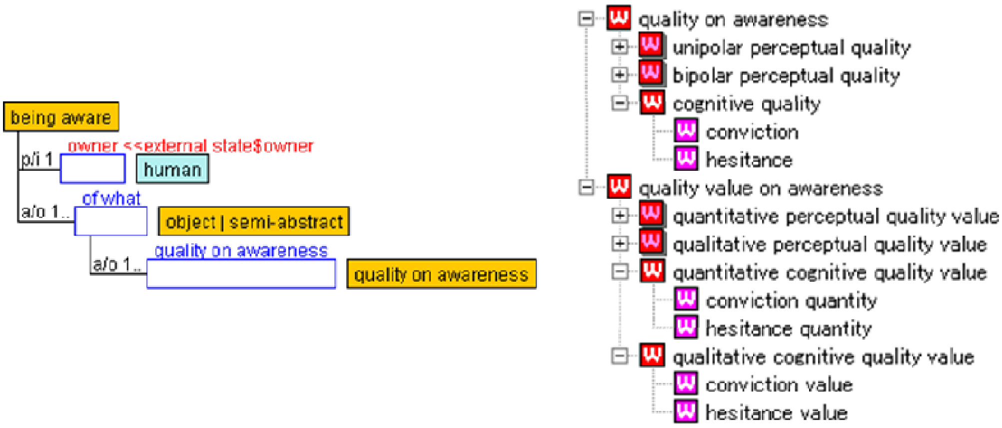
We referred a top-level ontology YAMATO[2] [6]. According to YAMATO, an entity is divided into three classes of a physical, abstract and semi-abstract. While instances of physical class need 3D space and time to exist, instances of abstract class needs neither of them. Instances of semi-abstract class need only time to exist, and the class contains mind, representation, a content and a representation form. Representation such as novels, poems, paintings, music, and symbols is distinguished from its proposition and form of representation [5]. Representation has a content slot which indicates a role played by a proposition, and has a form slot which indicates a role played by a representation form.

3.2 Attributes on awareness

We have partially expanded YAMATO to describe a subjective evaluation which is regarded as an expression of a psychological quantity. More precisely, it is defined as a representation of quality (defined in YAMATO) based on doer's awareness [8]. Doer's awareness is described in a state "being aware" (figure 5a). This is defined subclass of "external state" in YAMATO. Objects of awareness are represented by "of-what" role-holder played by a physical or a semi-abstract. Subslot of of-what is "cognitive quality"

[1] http://www.hozo.jp
[2] http://www.ei.sanken.osaka-u.ac.jp/hozo/ontolibrary/upperOnto.htm

played by "quality on awareness" represents a psychological quality which a doer subjectively feels.



(a) State of "being aware" (b) Hierarchy of quality and quality value on awareness

Figure 5. State of "being aware" and quality on awareness

Qualities on awareness and their values are sharply distinguished from physical qualities and their values defined in YAMATO. Figure 5b shows hierarchy of "quality on awareness" and "quality value on awareness". Learners' psychological quality such as conviction and hesitance are defined as subclass of "cognitive quality" under the quality on awareness. For example, "conviction" has two "referring to" slots; one is played by "conviction quantity" and the other is played by "conviction value". The conviction quality is subclass of "quantitative cognitive quality value", and the conviction value is subclass of "qualitative cognitive quality value". Both quantitative and qualitative cognitive quality values are defined under "quality value on awareness".

3.3 Multiple-choice test event

3.3.1 Description of learner's eye movement

To clarify relationships between learners' eye movement and mental states, we defined "multiple-choice test event" which represents that a learner takes multiple-choice test in an e-learning environment (figure 6). This is subclass of "extrinsic accomplishment" which is defined under "ordinary event" in YAMATO, and consists mainly of participants (objects) and constituting processes (actions). A learner and a learning material which participate in this event are defined as a role of participant inherited from "occurrent" defined in upper level. The learner role is played by a "human", and learning material is played by a "representing thing". The representing thing is a physical thing which mediates "representation" and is defined to have a "representation media" which is played by a "physical" thing in YAMATO. Thus, the learning material also consists of representations of a test and a representation media which is characterized by positions of a text and choices on the multiple-choice test.

A "constituted by" slot of multiple-choice test event, which is inherited from ordinary event, is played by "solve" action. In YAMATO, physical action is defined to be composed of "bodily motion". Thus, we defined "gaze" as a bodily motion and composed the solve action with gaze motions in the context of multiple-choice test. These slots represent "look-ahead at choice", "look-back at text" and "look-back at choice" as role-holders which are played by gaze motions. These slots also defined to have "fixation position" played by positions of a text and choices on the multiple-choice test.

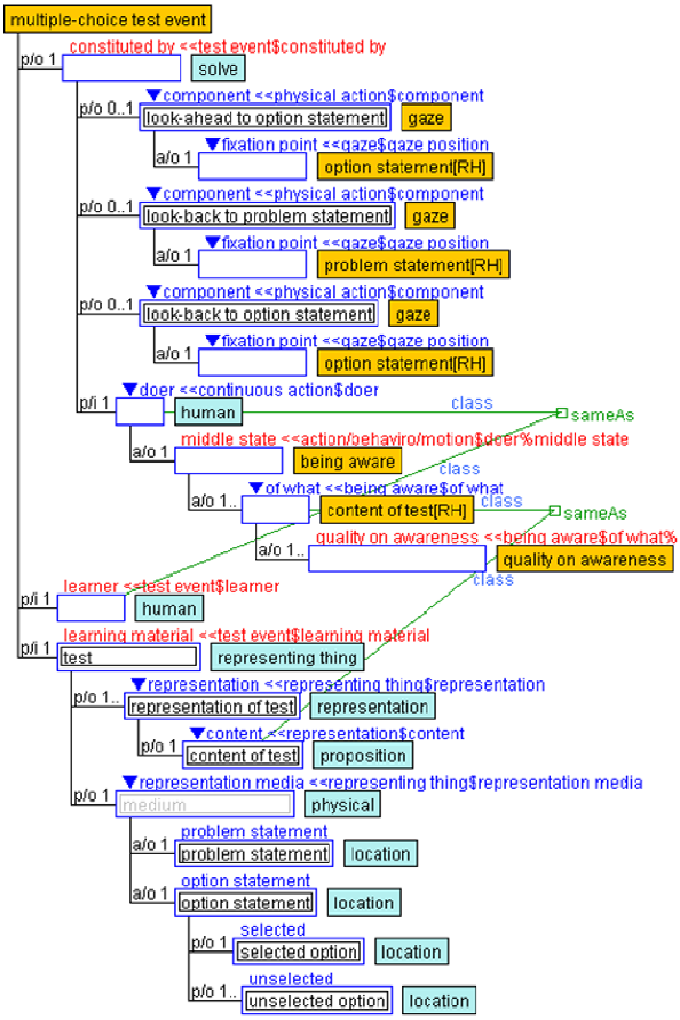


Figure 6. Description of multiple-choice test event

3.3.2 Description of learner's mental states

In the "constituted by" slot which is mentioned above, "doer" played by human represents a subject of the action to solve and motion to gaze. The doer's "middle state" is played by "being aware" and its "of what" slot represents content of consciousness. In this context, the content of consciousness is "content of test" which is a role-holder played by a proposition and is linked to subsubslot of the learning material with a "same as". Learners' mental states are described as "quality on awareness" role played by a "quality on awareness" which represents a psychological quality.

As mentioned above, conviction and hesitance are defined in the current study and can be player of the quality on awareness role. Thus, relationships between learner's eye movement and mental states are clarified in the multiple-choice test event.

4. Discussion and Conclusion

In the current study, we developed an ontology to provide consistent descriptions of knowledge for IMS. We described multiple-choice test event and positioned learner's eye movement and mental states in it. These descriptions can support to identify knowledge about relationships between eye movement and mental states which were obtained through empirical data. In a practical manner, when the gaze behaviors we found in section 2 are observed, each behavior is represented by setting cardinality of the role-holders played by gaze (figure 6) to 1. Along with this, learner’s mental states are represented by playing the quality on awareness role by conviction or hesitance. Hence, the ontological descriptions proposed in the current study form the outline of framework to manage knowledge implemented in IMS. In future work, we expand description of mental states and LLI resources such as easy/difficult, boring/interesting, confused/comprehending and tired/concentrating estimated facial image and face angle [11]. The main issues that must be addressed are how to clarify semantic constraints among various mental states and LLI resources.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 22300294.

References

[1] Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K. & Christopherson, R. (2009). Emotion Sensors Go To School. *Proceeding of the 14th International Conference on Artificial Intelligence in Education* (pp. 17-24). IOS Press.

[2] Horiguchi, Y., Kojima, K., & Matsui, T. (2009). A Study for Exploration of Relationships between Behaviors and Mental States of Learners For an Automatic Estimation System. *Proceedings of the 17th International Conference on Computers in Education* (pp. 173-176). Hong Kong: Asia-Pacific Society for Computers in Education.

[3] Kozaki, K., Kitamura, Y., Ikeda, M., & Mizoguchi, R. (2000). Development of an environment for building ontologies which is based on a fundamental consideration of ”relationship” and ”role”. *Proceedings of the Sixth Pacific Knowledge Acquisition Workshop* (pp. 205-221). Sydney: University of New South Wales.

[4] Mizoguchi, R., Vanwelkenhuysen. J., & Ikeda. M. (1995). Task ontology for reuse of problem solving knowledge. *Proceedings of Knowledge Building & Knowledge Sharing 1995* (pp. 46-59) .

[5] Mizoguchi, R. (2004). Tutorial on ontological engineering - Part 3: Advanced Course of Ontological Engineering. *New Generation Computing*. 22(2), pp. 193-220.

[6] Mizoguchi, R. (2010). YAMATO: Yet Another More Advanced Top-level Ontology, *Proceedings of the Sixth Australasian Ontology Workshop* (pp. 1-16). Sydney : ACS.

[7] Muldner, K., Christopherson, R., Atkinson, R. & Burleson W. (2009). Investigating the Utility of Eye-Tracking Information on Affect and Reasoning for User Modeling. *Proceedings of the 17th International Conference on User Modeling, Adaptation and Personalization* (pp. 138-149). Berlin: Springer.

[8] Muramatsu, K., Togawa, T., Kojima, K,. & Matsui, T. (2011). Proposal of A Framework to Share Knowledge About Consumer’s Impressions. *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence*. Vol. 1 (pp.388-393) SciTePress.

[9] Nakamura, Y., Akamatsu, N., Kuwabara, T., & Tamaki, M. (2002). Method for Detecting a Learner's Stalled Situation Using a Dispersion of Operation Time Intervals. *IEICE Transactions on Information and Systems*, J85-D (1), 79-90 (in Japanese).

[10] Nakamura, K., Kakusho, K., Murakami, M., & Minoh, M. (2010). Estimating Learners' Subjective Impressions of the Difficulty of Course Materials by Observing Their Faces in e-Learning. *IEICE Transactions on Information and Systems*, J93-D (5), 568-578 (in Japanese).

[11] Nosu, K. & Kurokawa, T. (2006) A Multi-Modal Emotion-Diagnosis System to Support e-Learning, *Proceedings of the First International Conferenceon Innovative Computing, Information and Control* Vol. 2 (pp. 274-278). IEEE Computer Society.

[12] Salvucci, D. D. (2001). An Integrated Model of Eye Movements and Visual Encoding, Journal of Cognitive Systems Research, 1(4), 201-220.

[13] Ueno, M. & Nagaoka, K. (2005). On-Line Analysis of e-Learning Time based on Gamma Distributions. *Proceeding of World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 3629-3637). AACE.

Development and Evaluation of an Intelligent Tutoring System for Teaching Natural Deduction

Kazuhisa Miwa [a*], Hitoshi Terai [a], Nana Kanzaki [a], and Ryuichi Nakaike[b]
[a] *Graduate School of Information Science, Nagoya University, Japan*
[b] *Graduate School of Education, Kyoto University, Japan*
*miwa@is.nagoya-u.ac.jp

Abstract: We present an intelligent tutoring system that teaches natural deduction to undergraduates. Our system was implemented on a client-server framework. An expert problem solver in the system provides basic instructional help, such as suggesting the use of a rule in the next step of solving a problem and indicating the inference drawn by applying the rule. Students learning with our tutoring system can vary the degree of help they receive (from low to high and vice versa). Empirical evaluation showed that the system enhanced the problem-solving performance of participants during the learning phase, and these performance gains were carried over to the post-test phase.

Keywords: Natural deduction, Help seeking, Client-server framework, Levels of support

1. Introduction

In this article, we present an intelligent tutoring system that teaches natural deduction (ND) to undergraduates. An ND system is a proof calculus in which logical reasoning is expressed by inference rules that are closely related to natural ways of reasoning (Barwise, 2003). Users of our tutoring system learn inference rules and strategies for applying the rules, such as strategies for inferring the proposition ¬Q => ¬P from the premise P => Q. ND is used in many universities' curricula for teaching the basics of logical thinking and formal reasoning. Figure 1 shows an example problem.

P → Q	Premise
| ¬Q	Premise
| | P	Premise
| | P → Q	Reiteration
| | Q	Conditional Elimi.
| | ¬Q	Reiteration
| ¬P	Negation Intro.
¬Q → ¬P	Conditional Intro.

Figure 1: An example proof of contraposition from P => Q to ¬Q => ¬P.

Traditionally, ND has been taught using paper and pencil. The instructor explains the functions of the inference rules and presents example problems to students, and the students learn the inference rules as they solve the problems. However, problem solving in ND is relatively complex, because it involves selecting a single appropriate rule from various candidate rules, applying the rule to the appropriate propositions, and drawing the inference that results from the rule's application. Students often face serious difficulties because they cannot determine the correct rules to use. They may stop learning because they have no idea about how to proceed. They need the instructor's help, but in a traditional classroom setting, which may have more than 100 students, individual tutoring is sometimes impossible.

We use an intelligent tutoring system in this setting. The system provides basic instructional help, such as suggesting a possible rule for the next step when solving a problem and indicating the inference that results from the application of that rule. To provide this help, the system is equipped with technical knowledge and inference mechanisms. The system has been designed as a complete problem solver, and it acts like an ND expert.

The system performs as a complete problem solver as an ND expert, but it has neither knowledge about nor strategies for the instructions. The system is able to solve problems, and it tells students what to do next on the basis of its own attempt to solve the problem. However, the system does not know how to help students learn effectively, that is, it does not know when, how, or to what degree it should offer help. Hint presentation is an essential factor in intelligent tutoring systems (Koedinger & Aleven,2007). In our system, users are responsible for the instructional management. Although our system has no management functions for controlling the degree of help students receive, students learning with our tutoring system can vary the degree of help they receive (from low to high and vice versa). Our research question is whether the system's mode of direct instruction, that is, simply telling students what to do next, is effective compared to the traditional paper and pencil method for teaching relatively complex tasks such as performing ND.

2. The Tutoring System

Our tutoring system consists of an expert problem solver and tutor terminals. The system does not have a database containing a set of ND problems and their solutions. Rather, the expert problem solver solves each problem on demand. It is designed as a production system and consists of a working memory whose layout is consistent with the structure of ND problems and production rules corresponding to the inference rules and strategies for solving ND problems.

Our system was implemented on a client-server framework. Miwa and his colleagues developed a web-based production system architecture called "DoCoPro" that enabled the system design (Miwa et al., in press). The expert problem solver is implemented on the server and performs the complex ND inferences. Client computers for the tutor terminals that are connected to the server perform simple interface processing. User learning processes are saved as log data on the server. Using this client-server framework, our system can operate in any educational environment, where various types of computers, e. g., high spec, poorly performing, and those on different types of operating systems are used.

Students working at tutor terminals can control the levels of help that they receive. The LOSs (levels of support) can be controlled with respect to both rule selection and rule application.

<u>LOSs for rule selection</u>:

- Level 3: The system presents applicable candidate rules and strategies and indicates the propositions to which those should be applied.
- Level 2: The system presents applicable candidate inference rules and strategies.
- Level 1: The system only presents a set of inference rules and strategies (no assistance).

LOSs for rule application:

- Level 2: The system generates a proposition that was inferred automatically.
- Level 1: The system presents partial information about an inferred proposition. Users provide the terms for the complete proposition that is to be inferred.

Figure 2 shows an example screenshot of the tutoring system. The system provides users with lists of inference rules and strategies. Once, the user selects one of the rules or strategies from the lists, then the system automatically runs the rule. The system presents the complete inference result or a template in which the result is partially blanked. The system supports students by offering help about the rule and strategy selection.

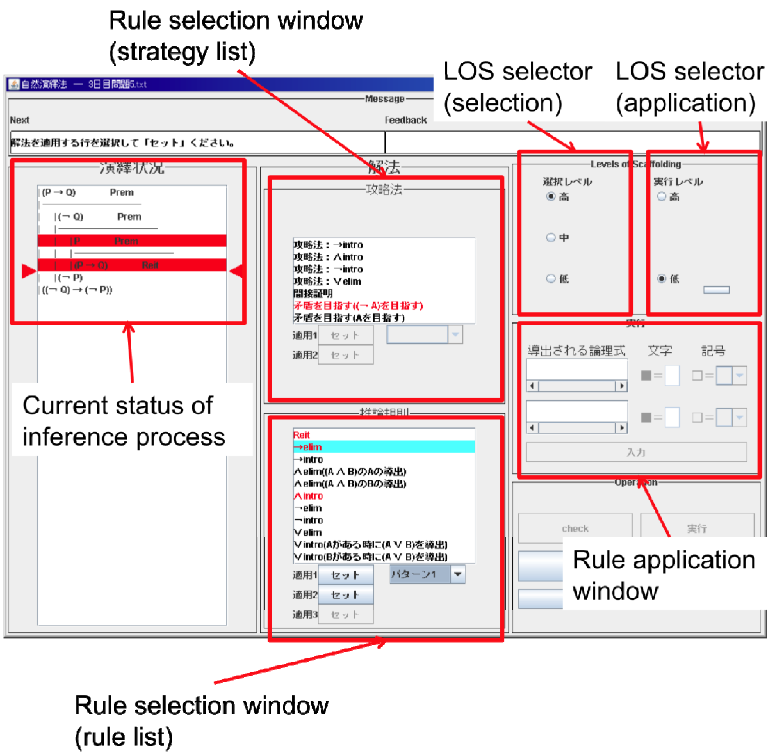


Figure 2: An example screenshot.

3. Evaluation

We verified that the system can successfully solve all 35 problems that are contained in a representative textbook (Todayama, 2000) that is used for ND education in Japan.

Our system is currently being used for teaching ND in university classes run by the first author. A class of 29 undergraduates in fall semester (2010) and 49 undergraduates in spring semester (2011) learned ND using our system, successfully acquiring the necessary skills for constructing ND proofs.

We designed and conducted an experiment to evaluate our system's usefulness. The purposes of the experiment were to verify whether learning with our tutoring system is an improvement over learning in the traditional paper and pencil manner. A total of 49 participants participated in the experiment. Of these, 33 were assigned to the experimental group and the other 16 to a control group.

In the initial phase of the experiment, the participants in both groups learned the basics of ND with handout materials and an instructional video. Next, in the 80-min learning phase, the participants in the experimental group learned ND by solving six example problems with our tutoring system. At the same time, the participants in the control group learned ND by solving the same six problems using handout materials, without the support of our tutoring system; the handout materials contain complete information about how to solve the problems. After the learning phase, the participants in both groups solved two post-test problems printed on the test sheets. One was an easy problem that requires a first-order subproof and the other was a difficult problem that requires a second-order subproof.

We analyzed the post-test scores in order to answer our first research question. In the learning phase, the participants in the experimental group solved an average of 5.21 example problems, but the participants in the control group solved only an average of 2.69 problems. The statistical analysis shows that the difference between the two groups is significant ($t(47) = 5.71$, $p < 0.01$), which means that the tutoring system successfully enhanced the participants' problem-solving performances during the learning phase. A crucial question is whether the performance gain was carried over to the post-test scores, where the tutoring system's assistance was not available.

Figure 3 shows the result of the post-test. The statistical analysis shows that the average score for solving the easy problem was higher in the experimental group than in the control group ($t(47) = 2.59$, $p < 0.05$), but the average score for solving the difficult problem was not ($t(47) = 1.42$, n.s.). This result shows an enhanced performance of those who learned with our tutoring system, but only for solving the easy problem.

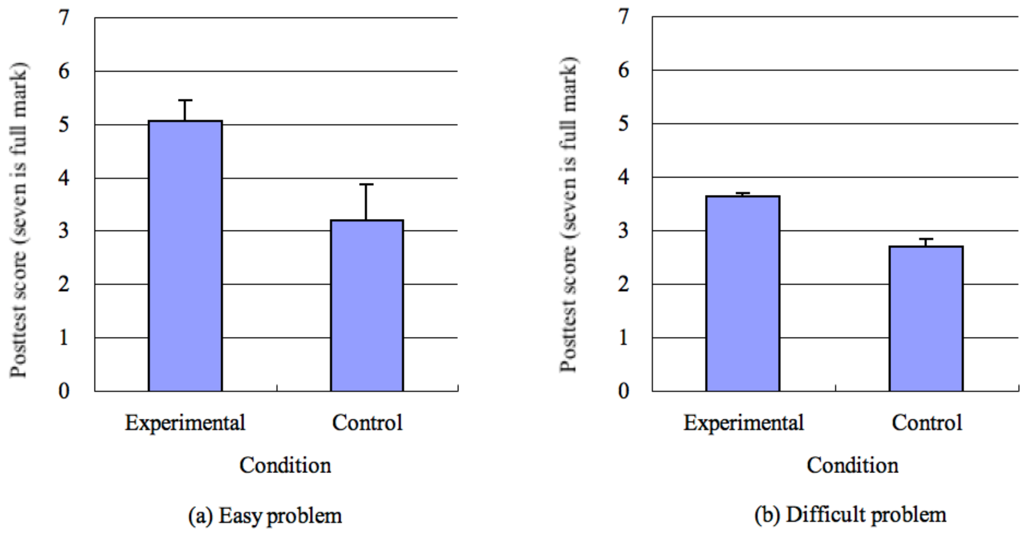


Figure 3: Average score of the post-test for the experimental and control groups.

4. Discussion and conclusions

Our research question concerned the usefulness of our tutoring system. Our results show that the participants in the control group solved only less than half of the problems within the learning phase. This means that they faced serious impasses during this phase. In contrast, the participants in the experimental group solved almost all the problems, which means that the system successfully eliminated the impasses. Although many preceding studies have indicated that learners behave irrationally in seeking help (Wood & Wood, 1999; Aleven & Koedinger, 2000), our overall results implies that the participants adaptively managed the LOSs while using our tutoring system.

The advantage of learning with the tutoring system was carried over only to solving the easy problem in the post-test; the enhanced performance did not reach a statistically significant level for the difficult problem. This may indicate a limitation of the system's direct mode of instruction, such as simply indicating the next step to be performed. An important part of our future research will involve addressing this limitation.

Other tutoring systems exist for teaching ND. One is a system called “Fitch” that provides students with templates for reasoning and automatically performs the reasoning (Barwise, 2003). However, this system does not provide information about which rule should be applied for each phase of the reasoning. As we mentioned above, instructional assistance in rule selection is more important for introductory students than in rule application, and so this is a critical limitation.

Croy et al. (2007) developed an intelligent tutoring system for teaching how to construct propositional proofs. To identify the best assistance for a specific stage of reasoning, they drew on previously acquired log data for students’ reasoning processes and used a Markov Decision Model to infer the next step in reasoning. In this way, they implemented a “Hint Factory” and were able to successfully improve introductory students’ learning in deductive logic courses. When our system provides help, it sometimes narrows the possible rules or strategies to a few candidates rather than a single one, because of the limitations of its inference abilities. Compared to our tutor, Croy et al.’s tutor appears to always provide one best candidate rule for the next step in reasoning. However, our system can handle new problems that have not been solved in advance. Croy’s system must accumulate training data for the Markov Model, and accumulating such data is time consuming, even though the researchers have tried to shorten the preliminary trials (Stamper, 2010). We have begun to expand our system for learning through problem posing, in which students generate their own problems and test the validity of the original problems by having our tutoring system solve those problems.

References

[1] Aleven, V., & Koedinger, K. R. (2000). Limitations of Student Control: Do Student Know when they need help? Proceedings of the 5th International Conference on Intelligent Tutoring Systems, ITS 2000. pp. 292-303.

[2] Barwise, J., & Etchemendy, J. (2003). Language, Proof and Logic. CSLI Publications.

[3] Croy, M., Barnes, T., & Stamper, J. (2007). Towards an Intelligent Tutoring System for propositional proof construction. Proceedings of European Computing and Philosophy Conference, 145-155.

[4] Koedinger, K. R., & Aleven, V. (2007). Exploring the Assistance Dilemma in Experiments with Cognitive Tutors. Educational Psychology Review, 19, 239-264.

[5] Miwa, K., Morita, J., Nakaike, R., & Terai, H. in press. Learning through Intermediate Problems in Creating Cognitive Models. Interactive Learning Environments.

[6] Stamper, J., Barnes, T., & Croy, M. (2010). Enhancing the Automatic Generation of Hints with Expert Seeding. Proceedings of the 10th International Conference on Intelligent Tutoring Systems (ITS 2010), 31-40.

[7] Todayama, K. (2000). Learning Logic through Building it. Nagoya University Publisher.

[8] Wood, H., & Wood, D. (1999). Help seeking, learning and contingent tutoring. Computers and Education, 33, 153-169.

Designing Affective Pedagogical Agents: How learners' and agents' gender and age influence emotion in an online tutoring task

Yugo Hayashi
College of Information Science and Engineering, Ritsumeikan University, Japan
yhayashi@fc.ritsumei.ac.jp

Abstract: In designing pedagogical agents, it is important to understand what factors stimulate the learner's affect to enhance learning motivation. To do so, the present study investigated the influence of affective mood on gender types of the (1) learner and the (2) pedagogical agent through an online tutoring activity. Affective states were ascertained using a questionnaire constructed on the basis of Russell's (1980) two-dimensional affective model data for 16 times during the tutoring activities. The participants were 290 psychology students, who were made to perform a homework activity. The results of two experiments consistently revealed that male students were sensitive to the pleasantness bipolar and female students were sensitive to the activation bipolar.

Keywords: Web-based tutoring, Embodied agents, Affective learning, Gender

Introduction

Past studies on pedagogical agents show that their presence is sometimes sufficient in promoting learners' affective states, such as motivation (Kim, Baylor, & Shen, 2007). One approach is to build tutoring systems that encourage learners, based on their most suitable affective state. The present study investigates the types of agents most suitable for students with personal characteristics and affects during an online learning activity.

1. Related work and research questions

1.1 Affective learning and tutoring systems

Recently, affective learning has become a popular topic in learning sciences and several studies have reported that affective factors are especially important in learning activities (Baylor, Kim, 2005; Hayashi, 2012). For example, Bower and Forgas (2011) revealed that positive moods can increase recall. Psychological studies in communication indicate that social embodied cues such as gaze, gesture, facial expression, and posture are good for facilitating teaching effectiveness (Mehrabian, 1966). Unfortunately, it is quite difficult for a teacher tutoring a large class to provide different types of cues suitable to each learner's preferences and needs. A solution is to use communication technologies such as tutoring systems and pedagogical agents. Then, the following research questions arise: "How do the most effective agents improve student learning?". The present study investigates the kind of factors related to one's affective state. This investigation has implications for designing effective tutoring systems that are capable of presenting information based on the user's affective state. The results of such as study will provide new ideas for designing an online tutoring system that would prove to be effective for teachers, motivators, and collaborators. The present study uses Russell's two-dimensional theory of emotion as a dependant variable to discover affective states relevant to mood during learning activities (Russell, 1980). Pleasure-displeasure (or valence) is a dimension of experience that refers to a hedonic tone. Activation is a dimension of experience that refers to a sense of energy. The vertical axis shows that a person could be somewhere on a continuum ranging from sleep (at the lower end), through drowsiness, relaxation, alertness, hyperactivity, and, finally, frenetic excitement (at the opposite end). The present study uses this model to analyze students' affective states during learning activities.

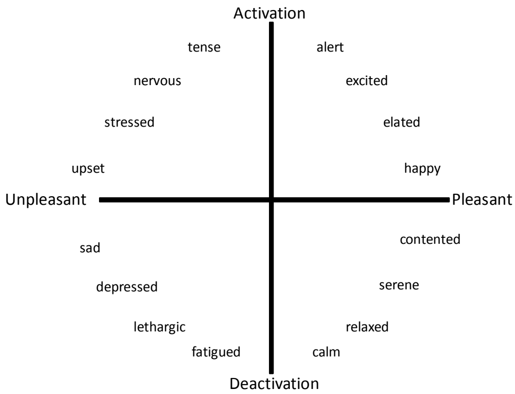


Figure 1. A schematic description of the two-dimensional structure of affect, adapted from Russell (1980)

1.2 Influence of personal status and social cues on affective learning

Studies in human-computer interaction show that people attribute personality traits to non-humans as well, that is, to animals or artificial agents such as robots (see Nass & Moon, 2000 for overviews). It is, therefore, assumed that the social verbal and non-verbal cues that facilitate motivation as discussed above are important factors in tutoring with computer agents. On the other hand, studies have focused on the user's prior knowledge and characteristics during interactions with artificial agents and robots. Research suggests that the effects of gender assigned to humans influence different impressions (Carpenter, Davis, Erwin-Stewart, Lee, Bransford, & Vye, 2009). Previous studies focusing on the use of virtual agents in pedagogical settings also show that these personal elements are important factors that influence interaction (Choi & Clark, 2006). The studies explained above show that personal statuses influence interaction with pedagogical agents. Unfortunately, little is known about the specificity of the influence of the learner's affective states, moreover, such are the activation level indicated in the vertical axis in the two-dimensional affective model (Figure 1). Moreover, little is known about the relationship between individual status and external social cues of the agents. The present study investigates this through an integrative experimental design shown in Figure 2.

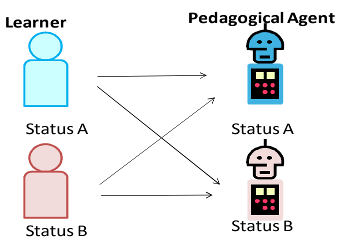


Figure 2. An interactional approach to investigate human-agent interaction

1.3 Aim of this study

The goal of the study is to understand the nature of learners' affective states due to their social status and the effects of the types of embodied agents. This study attempts to answer the following questions: (1) How do personal traits such as gender influence the user's affective mood during online learning activities?, (2) How do representations of agents influence affective mood during online activities?

2. Method

This study constructed a system that guides the learner in a simple web-based tutoring program. Students in a psychology class used the system to review key terms taught in a class, as homework. They were guided by a pedagogical agent who encouraged students by providing metacognitive suggestions and searches about the term on the Web for further understanding. Based on Russell's (1980) emotional models, the present study collected emotional variables as dependent variables. A total of 290 students participated in two evaluation experiments. The participants were all undergraduate students who were taking a psychology course and undertook a web-based tutoring task as part of the course work. The task was to read about a single topic, about key psychological terms, and answer a short quiz based on the literature. The study was conducted through two experiments as a two factorial between subject designs. In both experiments, the two actors, shown in the integrated design model in Figure 2 were investigated. Table 1 shows the experimental design. One hundred and thirty-five undergraduates participated in experiment 1 (male = 66, female = 69, *M* age = 19.72 years) and one hundred and fifty-five undergraduates participated in experiment 2 (male = 72, female = 83, *M* age = 19.49 years).

Table 1. Experiment design

	Male (Student)	Female (Student)
Male (agent)	H(M)/A(M)	H(F)/A(M)
Female (agent)	H(M)/A(F)	H(F)/A(F)

2.1 Materials and settings

In this study, a web-based tutoring system was developed only for the class. The system was constructed using the Web server, database, and rule-based program scripts. It was managed as a member-only system and its main operation was to tutor key terms taught in the class by presenting descriptive content. A total of 30 different key terms (e.g., Gestalt, long-term memory, cognitive dissonance) were extracted from an introductory psychology textbook and its explanations were entered in the system database. Students were randomly assigned to work on one key term. The tutoring sessions comprised 17 short passages, and students proceeded by clicking on to the next page (trial). During the task, there were four short quizzes on the key terms. Students were encouraged to go beyond a mere reading of these passages, to try and search through the web page to further understand the terms. Due to the experimental conditions, these suggestions were made by different types of pedagogical agents. The average time for this activity was approximately 30 minutes. The avatars used in the study were created using a 3D-image/animation-design tool called Poser 8 (www.e-frontier.com). Figure 3 shows an example of the arousal expressions of the avatars that were used in the study. While students participated in the task, they were also required to choose and click on one of the eight emotional icons (emoticons) intended to gauge their mood while undertaking the task. This was the dependant variable for this study: to ascertain affective states related to motivation towards learning activities. These eight emotional icons were depicted from Russell's two emotional dimensions presented in Figure 1 (Russell, 1980). This evaluation was presented at each trial of the passage.

Figure 3. Avatars used in the two experiments: The ones on the left were used in experiment 1 (male and female, respectively) and the ones on the right were used in experiment 2 (male and female respectively)

3. Results

Figure 4 (left) indicates the average of each participant's ratio of evaluation for experiment 1. To investigate the relationships between the gender types of the agent and the participant an analysis was conducted using the ANOVA. A 2 x 2 x 8 ANOVA was conducted on the average scores with the agents' gender (female agent vs. male agent) and participants' gender (female participant vs. male participant) as a between-subject factor, and affective state (pleasure vs. excitement vs. arousal vs. distress vs. displeasure vs. depression vs. sleepiness vs. relaxation) as a within-subject factor. The second-order interaction was not significant ($F(7,917) = 0.217$, $p = .98$). There was significant interaction between the participants' gender factor and the factor of affective state ($F(7,917) = 52.093$, $p<.01$). Consequently, an analysis of the simple main effect was conducted. Focusing on the affective state, the rating score at participant male conditions was significantly higher than at the participant female conditions in the pleasure state ($p<.01$). Results also show that the rating score at participant female conditions was significantly higher than at the participant male conditions in the arousal state ($p<.01$).

Figure 4 (right) indicates the average of each participant's ratio of evaluation for experiment 2. A 2 x 2 x 8 ANOVA was conducted as in experiment 1. The second-order interaction was not significant ($F(7, 1057) = 1.484$, $p = .17$). There was significant interaction between the participants' gender factor and the factor of affective state ($F(7, 1057) = 3.879$, $p < .01$) and the effect between the factor of agents' gender and affective state ($F(7, 1057) = 2.796$, $p =.05$). Given the significant interaction between the factor of participants' gender and the factor of affective state, an analysis of the simple main effect was conducted. Focusing on the affective state, the rating score at participant male conditions was significantly higher than at the participant female conditions in the pleasure state ($p < .01$). Results also showed that the rating score at participant female conditions was significantly higher than at the participant male conditions in the arousal state ($p < .01$). Now, we look at the simple main effect between the factor of the agents' gender and affective state. Focusing on the affective state, the rating score at agent male conditions was significantly higher than at the agent female conditions in the pleasure state ($p < .01$). Results also show that the rating score at agent male conditions was significantly higher than at the agent female conditions in the excitement state ($p < .01$).

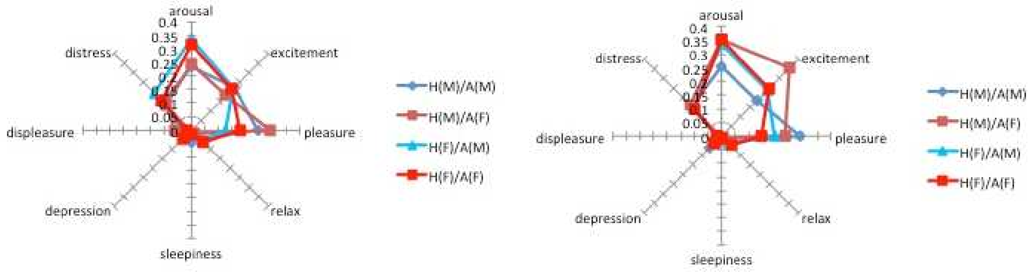


Figure 4. Results of the summary of all data in experiment 1 (left) and 2 (right)

4. Conclusion and Discussions

The present study captured the nature of the affective state based on Russell's (1980) two-dimensional model on an online tutoring activity. The focus of the study was to investigate how affective states may differ due to personal characteristics and factors such as agents that embody character. Results of the large-scale online experiment showed that females are more sensitive to the activation dimension than males and males are more sensitive to the pleasure dimension than females. In experiment 1 and 2, male learners had stronger pleasure ratings than female learners. Results also show that in experiment 1 and 2 female learners had stronger activation ratings than male learners. Consistent results on both experiments indicate that females tend to have more affective states on the vertical axis in Russell's (1980) model. On the other hand, males have more affective states on the horizontal axis. In experiment 2, the interaction between the agents' gender and affective state became significant. This indicates that the affective expressions of the agent appeared only in the child. This indicates that the 'age' of the character strongly influences the affective mood of the learner. The results of the present study may contribute to designing agents that use different types of affective expressions during pedagogy.

References

[1] Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education* 15(1), 95-115.

[2] Bower, G. H., & Forgas, J. P. (2001). Mood and social memory. In: Forgas, J.P. (ed.) *Handbook of Affect and Social Cognition*, pp. 95-120. LEA, NJ.

[3] Carpenter, J., Davis, J, Erwin-Stewart, N., Lee,T, Bransford, J, & Vye, N, (2009) Gender representation and humanoid robots designed for domestic use, *Int. J. Soc. Robot.*, 1, (pp. 261-265).

[4] Choi, S., & Clark, R. E. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning English as a second language. *Journal of Educational Computing Research*, 34(4), 441-466.

[5] Hayashi, Y. (2012), On pedagogical effects of learner-support agents in collaborative interaction. *In S.A. Cerri and B. Clancey (Eds.): Proceeding of the 11th International Conference on Intelligent Tutoring Systems (ITS2011), Lecture Notes in Computer Science, Springer-Verlag*, Vol 7315, pp. 22-32.

[6] Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning* 23(3), 220-234.

[7] Mehrabian, A. (1966). Immediacy: An Indicator of Attitudes in Linguistic Communication. *Journal of Personality*, 34(1), 26-34.

[8] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.

[9] Schlosberg, H. (1941). A scale for the judgment of facial expressions. *Journal of Experimental Psychology*, 29, 497-510.

[10] Graham, J., & Heywood, S. (1975). The Effects of Elimination of Hand Gestures and of Verbal Codability on Speech Performance. *European Journal of Social Psychology*, 5(2), 189-195.

[11] Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56, 81-103.

Combining Language and Speech Features to Predict Students' Emotions in E-Learning Environments

Liang-Chih YU[*], Shou-Fang LIANG, Wei-Hua LIN
Department of Information Management, Yuan Ze University, Taiwan, R.O.C.
*lcyu@saturn.yzu.edu.tw

Abstract: Emotions play an important role in e-learning environments. Text and speech have been recognized as convenient and natural means for expressing emotions, and are increasingly used in human-computer interaction interfaces for e-learning applications, indicating that language and speech could potentially be used to predict learner emotions. In this study, we investigate the use of speech and language features for automatic emotion recognition. A corpus of emotion-laden sentences was collected from student-teacher dialogs in the context of mathematics instruction. The corpus was then annotated to analyze emotion types as they occurred in e-learning applications. The speech and language features were then used to build several classifiers for emotion recognition. Experiments show that the two features combined yielded better results than either feature alone. In addition, among speech features, energy and formant are found to best contribute to successful classification.

Keywords: Emotion recognition, natural language processing, speech processing

Introduction

Students frequently react to satisfactory or dissatisfactory learning performance by expressing positive or negative emotions, which, in turn, may have an impact on subsequent learning outcomes [1][2]. For instance, Rodrigo et al suggested that boredom may have a negative impact on student achievement, while confusion may have both positive and negative effects [2]. This has raised interest in technological solutions for automatic emotion recognition because accurately assessing changes in learner emotional states can allow e-learning systems to provide appropriate suggestions, thus improving learning outcomes.

Text and speech have been recognized as convenient and natural means for expressing emotions, and are increasingly used in human-computer interaction interfaces for e-learning applications such as computer supported collaborative learning (CSCL) [3][4] and intelligent tutoring systems [5][6]. For example, text-based synchronous online chat can be used for group discussion to support collaborative learning [3]. Asynchronous online discussion forums also facilitate knowledge sharing through posting and reading forum articles [4]. Speech has been integrated to help students interact with intelligent tutoring systems [5][6]. This increasing use of text- and speech-based interfaces positions both language and speech as potential features for identifying learner emotions in e-learning applications. Previous research has also demonstrated the effectiveness of using language and speech features for emotion recognition, but mainly in non-e-learning domains such as identifying positive and negative emotions (binary) [7], six basic human emotions [8], and specific emotion types in business [9][10] and medical domains [11][12]. Very little research has investigated the use of language or speech features for emotion recognition in e-learning applications [7]. In addition to language and speech features, mouse movements, facial features and body posture have also been investigated for identifying learner emotions [13][14].

Table 1. Corpus annotation results.

Emotion	Num. of sentence	Proportion	A1-A2 Agreement	A1 Accuracy	A2 Accuracy
Delight	194	26%	93.30%	94.85%	98.45%
Contempt	53	7%	62.26%	77.36%	84.91%
Boredom	81	11%	69.14%	88.89%	80.25%
Frustration	99	13%	59.60%	79.80%	78.79%
Confusion	134	18%	84.33%	92.54%	91.79%
Others	198	26%	87.88%	89.90%	97.98%
Sum/Avg.	759	100%	81.16%	89.33%	91.70%

In this paper, we investigate the use of both speech and language features to identify student emotions. To this end, we first collected a corpus of emotion-laden sentences from student-teacher dialogs in the context of mathematics instruction. The corpus was then annotated to analyze various emotion types as they occurred during use of e-learning applications. Finally, the speech and language features were combined to build several classifiers for emotion classification.

1. Corpus Annotation and Analysis

1.1 Corpus Annotation

The corpus collection process involved communication among three mathematics teachers and 149 students in discussing mathematical problems in the classroom. A total of 759 sentences were collected from student-teacher dialogs to form an emotion text corpus, with emotion types classified as Delight, Contempt, Boredom, Frustration, and Confusion. Sentences in the corpus not explicitly characterized by a specific emotion type were categorized as Other.

To analyze student emotions, the three mathematics teachers annotated the corpus to create a standard of the various emotion types. Each sentence in the corpus was first annotated with one of the six emotion types (including Other) by two teachers (annotators). In case of disagreement between the two annotators, the disputed sentence was judged by the third teacher (adjudicator) for a final decision. Post-adjudication proportions of the various emotion types and the accuracy of the two annotators could then be calculated from the corpus. The annotation results presented in Table 1 show that around 74% of the sentences in the corpus contained an emotion type, while the remaining 26% were out-of-domain sentences (i.e., “Other”). Among the five emotion types, Delight and Confusion were found to predominate.

Table 1 shows that the annotators A1 and A2 agreed on 81.16% of the sentences reviewed. Agreement regarding Contempt and Frustration was relative low, indicating that these two emotion types were more ambiguous. For example, Contempt may be misclassified as Delight, while Frustration may be misclassified as Boredom or Confusion. The accuracy of A1 and A2 (as calculated by their consistency with the adjudicator for samples for which there was disagreement) was 89.33% and 91.70%, respectively. Such human (expert) results can be viewed as the upper bound for automatic emotion classification using machine learning algorithms. The accuracy for Frustration for both annotators was relatively low, again indicating that this emotion type was more ambiguous.

Table 2. Linguistic features and sample sentences for the emotion types.

Emotion	Example sentence	Linguistic feature
Delight	I made a big progress this time. Oh! Great! This question is so easy.	progress, great, easy, simple
Contempt	This question is so stupid and deserves no response. This question is too elementary. Even a kid can do it.	stupid, basic elementary, kid
Boredom	That's so bored. I have addressed such kind of questions many times before. I don't want to waste my time on such a tedious question.	bored, boring, tedious
Frustration	That's too bad. I will be failed. Forget it. That's too hard.	bad, fail, hard, difficult
Confusion	This question is ambiguous. I do not understand the meaning. Why the question can be solved in this way?	ambiguous, why, weird, confuse,

Table 3. Prosodic features for each emotion types.

	Delight	Contempt	Boredom	Frustration	Confusion
Pitch Mean	increased	normal or increased	decreased	decreased	increased
Pitch Max	increased	increased	increased	decreased	increased
Pitch Min	increased	decreased	decreased	decreased	decreased
Energy Mean	increased	normal	increased	decreased	decreased
Energy Max	increased	increased	increased	decreased	decreased
Energy Min	increased	decreased	increased	decreased	decreased
Formant Mean	f1,f5 increased; f2-f4 decreased	f1,f3-f5 increased; f2 decreased	f1,f5 increased; f2-f4 decreased	f1,f3,f5 increased; f2,f4 decreased	f1,f2,f4 decreased; f3,f5 increased
Formant Max	f1-f3 increased; f4,f5 decreased	f1-f5 increased	f1-f4 decreased; f5 increased	f1,f3-f5 increased; f2 decreased	f1-f5 increased
Formant Min	f1,f4-f5 increased; f2,f3 decreased	f1,f3 decreased; f2,f4,f5 increased	f1-f5 increased	f1 decreased; f2-f5 increased	f1,f2 increased; f3-f5 decreased

1.2 Linguistic Features

Table 2 presents several sample sentences for each of the five emotion types. Students may express Delight when they are satisfied with their learning performance or when facing easy questions, but may express Contempt if the questions are too simple. Students may also express Boredom if they feel the questions are pointless or senseless. Conversely, students may express Frustration when they are worried about their performance or when facing difficult, and Confusion when facing ambiguous or incomplete questions. The last column in Table 2 summarizes a number of linguistic features for the various emotion types.

1.3 Speech Features

A total of 379 sentences were randomly selected for recording. The input waveforms were captured at 16kHz, a frame length of 33ms and an average length of utterance 3 seconds.

Table 4. Classification accuracy of different methods with different features (% accuracy).

	Two-class			Five-class		
	NB	C4.5	SVM	NB	C4.5	SVM
Language	83.64	74.41	89.45	64.38	57.26	70.45
Speech	70.45	77.84	76.52	38.26	51.98	55.67
Language + Speech (All)	85.22	79.16	91.29	67.02	59.63	72.03
Language + Pitch	81.79	73.35	88.92	62.80	56.73	69.92
Language + Energy	85.49	**84.17**	89.45	63.59	59.63	69.66
Language + Formant	81.05	74.14	90.50	65.17	55.41	72.30
Language + Pitch + Energy	**86.28**	83.11	89.71	65.17	58.84	70.18
Language + Pitch + Formant	81.27	75.46	90.50	64.12	55.94	72.03
Language+Energy+Formant	86.02	79.68	**91.56**	**67.28**	**60.69**	**72.56**

Recording was conducted in an office environment without obtrusive background noise. To ensure the quality of the recorded corpus, objective tests were performed to validate the correctness of the recorded data which was evaluated by averaging responses from all test subjects. The ground truth of most utterances was decided by a unanimous vote, thus giving the selected utterances significance.

Table 3 summarizes the analysis for each prosodic feature with respect to the various emotion types. According to our observations, the energy related features (i.e., mean, max and min) are useful for differentiating between high and low active states such as Delight and Frustration. The pitch related features are useful for discriminating between both Frustration and Confusion, and Delight and Boredom. In addition, the formant is also an important feature for discriminating among the various emotion types.

2. Experimental Results

The classifiers used in this study include the Support Vector Machine (SVM), C4.5, and the Naïve Bayes (NB) classifier from the Weka Package [15][16]. Each classifier was trained using language features (i.e., individual words), speech features (i.e., pitch, energy and formant as in Table 3), and both. A total of 379 recorded utterances were analyzed with 10-fold cross-validation. Each test utterance was classified as belonging one of the five emotion types from Table 2. A two-class classification was also performed by dividing the five emotion types into positive (Delight and Contempt) and negative emotions (Boredom, Frustration, and Confusion). Performance is measured as a function of *accuracy*, i.e., the number of correctly classified utterances divided by the total number of test utterances.

Table 4 shows the results of different classifiers with different features. For all classifier in both two-class and five-class classification, combining the speech and language features is found to yield higher performance than either individual feature. In addition, different features made different contributions to different classifiers. For NB and C4.5, Energy was the most promising feature because Energy-related feature combinations (i.e., Language+Energy, Language+Pitch+Energy, and Language+Energy+Formant) were more accurate than the other combinations. Conversely, for SVM Formant was found to be the most promising feature. The highest accuracies for two-class and five-class classification were 91.56% and 72.56%, respectively, indicating that there is still much room for improvement in five-class emotion classification.

3. Conclusion and Future Work

Speech and language features are used to identify emotions from a corpus of learner utterances collected within the context of mathematics instruction. The corpus is analyzed to determine emotion types, along with their associated speech and language features. Experimental results show that combining the two features yielded higher performance than using either feature alone and, among the speech features, energy and formant were found to make the greatest contribution to accurate identification. Future work will investigate other significant features to further improve classification performance. An additional possible direction is to realize emotion recognition in text and speech based e-learning applications.

Acknowledgements

This work was supported by the National Science Council, Taiwan, ROC, under Grant No. NSC100-2632-S-155-001 and NSC99-2221-E-155-036-MY3. The authors would like to thank the anonymous reviewers for their constructive comments.

References

[1] D'Mello, S., Graesser, A. & Picard, R. W. (2007). Toward an Affect-Sensitive AutoTutor. *IEEE Intelligent Systems*, 22(4), 53-61.

[2] Rodrigo, M. M. T., Baker, R. S. J. D. & Nabos, J. Q. (2010). The Relationships between Sequences of Affective States and Learner Achievement, *Proceedings of ICCE'10* (pp. 56-60), Putrajaya, Malaysia.

[3] Chiu, C. H. & Hsiao, H. F. (2010). Group Differences in Computer Supported Collaborative Learning: Evidence from Patterns of Taiwanese Students' Online Communication. *Computers & Education*, 54(2), 427-435.

[4] Yeh, Y. C. (2010). Analyzing Online Behaviors, Roles, and Learning Communities via Online Discussions. *Educational Technology & Society*, 13(1), 140-151.

[5] Litman, D. J. & Silliman, S. (2004). ITSPOKE: An Intelligent Tutoring Spoken Dialogue System, *Proceedings of HLT/NAACL'04*, Boston, MA.

[6] Ros´e, C. P., Litman, D., Bhembe, D., Forbes, K., Silliman, S., Srivastava, R. & K. VanLehn, (2003). A Comparison of Tutor and Student Behavior in Speech Versus Text Based Tutoring. *Proceedings of the Workshop on Building Educational Applications Using Natural Language Processing at HLT/NAACL-03*, Edmonton, Canada.

[7] Lee, C. M. & Narayanan, S. S. (2005). Toward Detecting Emotions in Spoken Dialogs. *IEEE Trans. Audio, Speech, and Language Processing*, 13(2), 293-303.

[8] Cook, N. D., Fujisawa, T. X. & Takami, K. (2006). Evaluation of the Affective Valence of Speech Using Pitch Substructure. *IEEE Trans. Audio, Speech, and Language Processing*, 14(1), 142-151.

[9] Bai, X. (2011). Predicting Consumer Sentiments from Online Text. *Decision Support Systems*, 50(4), 732-742.

[10] Xu, K., Liao, S. S., Li, J. & Song, Y. (2011). Mining Comparative Opinions from Customer Reviews for Competitive Intelligence. *Decision Support Systems*, 50(4), 743-754.

[11] Yu, L. C., Chan, C. L., Lin, C. C. & Lin, I. C. (2011). Mining Association Language Patterns Using a Distributional Semantic Model for Negative Life Event Classification. *Journal of Biomedical Informatics*, 44(4), 509-518.

[12] Wu, C. H., Yu, L. C. & Jang, F. L. (2005). Using Semantic Dependencies to Mine Depressive Symptoms from Consultation Records. *IEEE Intelligent Systems*, 20(6), 50-58.

[13] Horiguchi, Y., Kojima K. & Matsui, T. (2009). A Study for Exploration of Relationships between Behaviors and Mental States of Learners for an Automatic Estimation System, *Proceedings of ICCE'09* (pp. 173-175), Hong Kong, China.

[14] Graesser, A. C., Chipman, P., Haynes, B. C. & Olney, A. (2005). AutoTutor: An Intelligent Tutoring System with Mixed-Initiative Dialogue, *IEEE Trans. Education*, 48(4), 612-618.

[15] Witten, I. H. & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques, 2nd Edition, Morgan Kaufmann, San Francisco.

[16] Zhang, P., Zhu, X., Shi, Y., Guo, Li. & Wu, X. (2011). Robust Ensemble Learning for Mining Noisy Data Streams. *Decision Support Systems*, 50(2), 469-479.

Intelligent feedback polarity and timing selection in the Shufti Intelligent Tutoring System

Stuart JOHNSON & Osmar ZAIANE
Alberta Innovates Center for Machine Learning, University of Alberta, Canada
{shjohnso, zaiane}@ualberta.ca

Abstract: It is well known that the training of medical students is a long and arduous process. Students master many areas of knowledge in a relatively short amount of time in order to become experts in their chosen field. The Socratic Method used in the latter stages of medical education, where a physician directly monitors a group of students, is inherently restrictive due to the limited number of cases and length of the students' rotations. Innovative Intelligent Tutoring techniques offer a solution to this problem. This paper outlines the overall structure and design of Shufti, an Intelligent Tutoring System (ITS) focused on mammography and medical imaging. Shufti's aim is to provide medical students with an improved learning environment, exposing them to a broad range of examples supported by customized feedback and hints driven by an adaptive Reinforcement Learning system and Clustering Techniques.

Keywords: Intelligent Tutoring System, Feedback, Hints, Reinforcement Learning, Machine Learning, Data Mining, Breast Cancer, Serious Games

Introduction

Shufti is an Intelligent Tutoring System (ITS) focused on mammography designed to help medical imaging students master the complexities of producing a diagnosis based on relatively poorly defined, low contrast images. Shufti takes the form of a web-based computer educational game where learners accumulate points for correctly diagnosing images. The learners are presented with pairs of mammographic images with overlaid grids and are expected to identify Regions Of Interest (ROI) within those images.

ROI's are regions which would normally necessitate further investigation by a radiologist. Students identify what they believe to be lesions by selecting squares within a grid which has been overlaid on the mammogram (see Figure 1). Once students have completed an exercise they are then given a score, which is derived from their accuracy in identifying lesions minus points for hints they may have requested along the way. Figure 1 depicts an example of feedback given to a student, post-exercise, showing their score relative to other students as well as relative to their previous attempts at the exercise.

Producing a high quality ITS for mammography is a non-trivial task. As demonstrated in Corwley et al.'s Slide Tutor [2], which made use of Natural Language Processing to resolve this issue in the field of pathology. Many attributes of the field do not lend themselves readily to computerized instruction. Amongst these, the two most important are the lack of sufficient time on the part of medical students to explore a large number of cases, and that mammography is an ill-defined domain according to the criteria outlined by Fournier-Viger et al. [3]

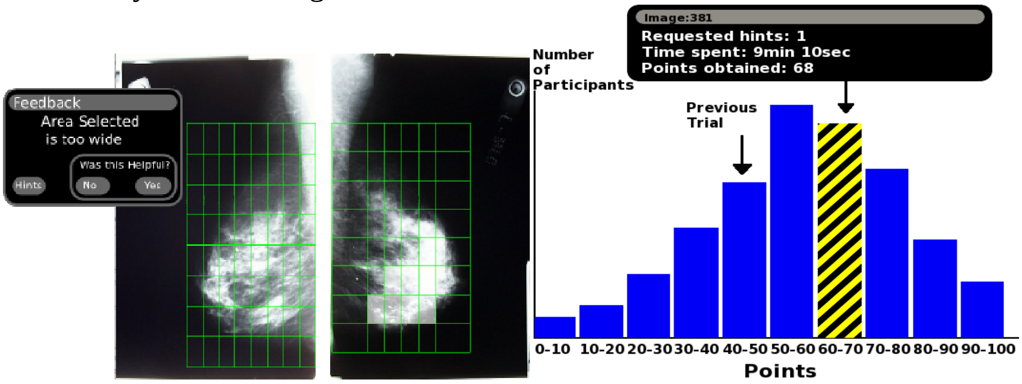


Figure 1: Shufti's feedback during and post-exercise.

To address the issue of learners viewing an insufficiently broad selection of cases, we not only make it easy for them to gain exposure to a wide array of well selected cases but we also incent them to broaden their exposure using a number of concepts from the field of competitive gaming.

The first game design convention used in Shufti is a point system. Points are a means of performance quantification that lack a direct connection with any one metric. In Shufti points are based on accuracy, time spent on an exercise, and the number and type of hints requested by the learner.

The second game design convention used is the concept of a competitive ranking system. Students are ranked based on their total accumulated scores on exercises, emphasizing volume and breadth of cases seen. Competition has been shown to be a good motivator for individuals to spend more time on tasks and can increase performance as demonstrated in Education and Crowdsourcing Schulze et al. [4].

Along with these game design conventions Shufti utilizes ideas taken from human tutoring; hints and feedback. In most literature about ITS, the terms “hint” and “feedback” are used interchangeably. This stems from the concept that all tutor-learner interactions are of the same assistive nature. In Shufti this is not the case; hints and feedback are complementary tools. Hints are much more specific than feedback in the information they impart to the user. An example of a hint would be, the number of ROIs present in the image, whereas an example of feedback, would be a message such as “Good job.”, or “Are you sure you’re done?” Additionally, Shufti provides hints in a prompted manner to the user whereas feedback is provided in an unprompted manner. More specifically, Shufti presents the user with a list of hints and associated score penalties during an exercise. Feedback, on the other hand, is provided automatically with the user having only an indirect influence on whether or not it is issued.

1. Feedback and Hints

One of the core attributes of human one-on-one tutoring is the active role which the tutor plays in the learning experience. An effective human tutor will provide hints, positive feedback and constructive/negative feedback in a strategic fashion so as to aid the learner's progress. Moreover, while some people may appreciate feedback, others may dislike it. Human tutors intuitively understand to whom feedback is beneficial. Mammography lacks the clear domain models, formal theorems, or cognitive models necessary to automatically teach mammogram diagnosis[3], consequently Shufti utilizes a variety of means to effectively simulate attributes of a human tutor.

Users are presented with a set of possible hints, each one being labeled with a description of the type of information the user will receive, and the specific score penalty which will be applied should the user accept the hint.

Shufti assigns penalties to hints to discourage gaming of the system, a phenomena where the user repeatedly requests hints until the answer is fully revealed [1]. Hint penalties may also have the interesting effect that learners will strategically select the minimum number of hints necessary for them to answer an exercise correctly. This strategic hint selection results in a form of user controlled difficulty, as, if the student selects the minimum number of hints necessary for the completion of the exercise, they have in effect adjusted the difficulty to the maximum they can successfully complete. Under normal circumstances mammography lacks an automated means with which to determine exercise difficulty. By giving users control over their own difficulty but incenting users to tackle challenges, Shufti elegantly resolves this issue.

Feedback in Shufti takes the form of both negative and positive messages – an attribute called polarity. Positive feedback is encouraging in nature such as, "Nicely done" and serves as a sign to the learner they are on the correct path to solving the exercise. Negative feedback is corrective in nature serving to steer the learner back onto the correct path with statements such as "You've missed something," or "Look around more."

The timing of a feedback is critical as it needs to be associated with an event or state and should not be disruptive. The selection of feedback polarity and timing is performed using two methods. First a clustering-based method is used which relies on a learner's reaction to feedback in order to determine its relevance. The timing in this method is controlled by one of the timing models described below and displayed in Figure 2. The second method relies on Reinforcement Learning (RL) to control the content, polarity, and timing of feedback delivery.

2. Approaches for Feedback

Exercises in Shufti are categorized by difficulty level, and students move from one level to the next after accumulating sufficient points and diagnosing a defined number of mammograms. Learners are modeled by retaining their current level, the total number of points, the number of images they have attempted, the average number of hints they requested per image, and the accumulated penalties due to requested hints .

Moreover, Shufti records the task state transitions during each exercise. Task State Transition are comprised of the exercise state and actions undertaken by the learners during the exercise. The Task State Transition record includes the current and past states representing the current solution, the last action taken by the learner, the proposed feedback, and the reaction to the feedback by the learner. The state is the number of grid cells selected that differ from the exercise solution (i.e. hamming distance). Actions are operations such as selecting a square, de-selecting a square, some mouse movements, or submitting the exercise for evaluation. The reaction to feedback is whether the learner explicitly found the previous feedback helpful.

Feedback polarity is based upon whether the state of the exercise improved or degraded. The degradation or improvement is determined by comparing whether the hamming distance between the past state and the solution has increased or decreased in contrast with the hamming distance between the current state and the solution.

Feedback is a critical part of the effectiveness of a tutor. Shufti contains methods for determining the content, polarity and timing of feedback. Polarity refers to whether it is a positive, encouraging message or a negative, corrective message. We propose two feedback control approaches: a clustering-based method and a technique based on RL.

2.1 Clustering-based method

The first approach Shufti uses to decide whether or not to give a feedback and the type of feedback to use, considers learners in groups of similar learners – in other words, *clustering*. Clusters of similar learners are based on their levels, points accumulated, their requested hints and the number of attempted exercises.

Feedback timing is determined by one of five models (see Figure 2): *Random* is feedback occurring randomly. *Timed* is feedback delivered after timed intervals. *After Action* feedback is issued in response to a learner's action. *Timed After Action* is triggered by an action however is delayed. *Random After Action* is again similar to After Action except it is randomly delivered (i.e. may or may not be issued).

Delivery of feedback for the timing models is based on the task state transitions of similar learners (i.e. learners in the same cluster as the current learner). The nature of the feedback delivered is chosen by examining the feedback that historically has been most appreciated by other users in the same cluster and the same task state transition.

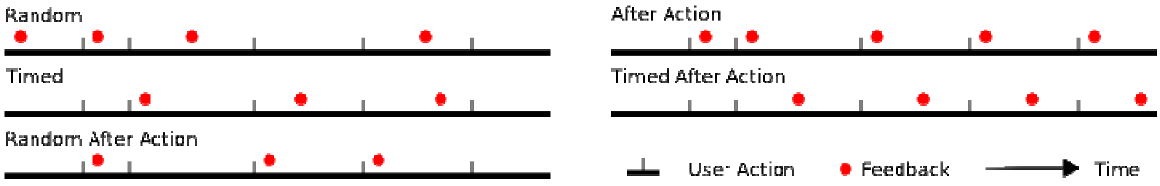


Figure 2: Patterns of feedback timing models.

2.2 Reinforcement Learning based method

Tuning to individual students is one of the ways in which a tutor can offer a superior learning experience. RL offers an automated method with which an ITS can tune its feedback delivery to individual learners and thus approximate a human tutor.

RL is a class of machine learning techniques which resolve problems of mapping situations to actions to maximize or minimize a metric[5]. RL allows Shufti to learn the most effective times to issue feedback, avoiding the use of preset timing models.

An RL system can be thought of as two components; an environment and an agent within which it acts. The environment provides state data and a reward signal to the agent which attempts to maximize the total reward over time. The agent makes use of methods such as Temporal-Difference Learning[5] or Monte Carlo Methods[5] to determine the most long-term rewarding action to take in any given state.

Shufti's environment offers task state transitions as state information to the agent seeking to minimize P in the following formula: $P = \sigma * count(\tau) + \omega * count(f) - \alpha * score$ where P is the total penalty assessed to the agent, σ is the penalty assigned over time, $count(\tau)$ is the total time passed, ω is the feedback penalty, and $count(f)$ is the total number of feedbacks given, α is the reward per score point earned, and $score$ is the total score assigned for the exercise. The longer the learner takes the larger the penalty which encourages the agent to provide feedback. This is balanced by penalties assigned to the agent for giving feedback, which results in strategic feedback issuance. Rate of feedback increases with σ and decreases as ω increases. RL allows Shufti to fine tune for individual learners, thus more closely matching a human tutor, however it lacks the adaptive advantages of learning from many users (clustering).

3. Competition

One of the limitations in traditional imaging analysis training is the amount of cases students are exposed to. Shufti improves on traditional training in two ways; it has an extensive selection of exercises covering a range of scenarios unlikely to be seen during a short rotation as a student in a radiology department, and, it uses competitive techniques learned from gaming to incent students to review as broad a range of scenarios as possible.

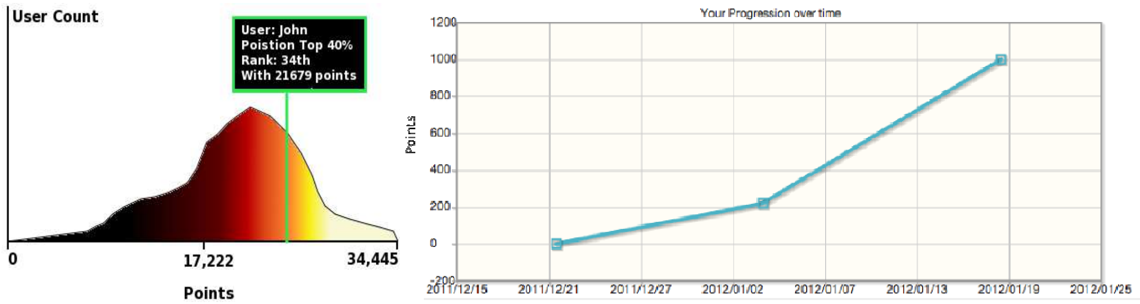


Figure 3: A representation of a user's position

In total, scores are based on problem difficulty, answer accuracy, time spent, and the hints requested. Learners are presented with a variety of means to see how they rank next to their peers, including public leader boards (similar to those used in online games), and performance distribution curves. Figure 3 shows the distribution of participants' scores and the relative position of the user “John”. Overall ranking in Shufti is determined by the sum of all scores they have received, encouraging them to attempt a large number of exercises.

4. Conclusion

Shufti is an innovative solution to many of the issues involved in providing a high quality learning experience to learners in the field of mammography. We present two machine-learning techniques to provide high quality automatic feedback to learners; Clustering and Reinforcement Learning. We also make use of an interesting user controlled hint structure in an effort to not just reduce learners' attempts to game the system but to also exploit the gaming habits of learners in an effort to aid their learning experience. Along with feedback and hints, Shufti also innovates in the field of ITS UI design adopting many features from serious games in an effort to improve the learners' experience.

References

[1] Baker, R. S., Corbett, A. T., Koedinger, K. R., Lester, J. C., Vicari, R. M., & Paraguacu, F. (2004). Detecting Student Misuse of Intelligent Tutoring Systems. In *Intelligent Tutoring Systems* (Vol. 3220, pp. 54-76). Berlin/Heidelberg: Springer Press.

[2] Crowley, R. S., Legowski, E., Medvedeva, O., Tseylin, E. E., & Roh, J. D. (2007). Evaluation of an intelligent tutoring system in pathology: Effects of external representation on performance gains, metacognition, and acceptance. *JAMIA* , *74* (2), 281-296.

[3] Fournier-Viger, P., R., N., & Nguifo, E. M. (2010). Building intelligent tutoring systems for ill-defined domains. *Studies in Computational Intelligence* , 308.

[4] N., K., Schulze, T., & Veit, D. (2011). More than fun and money. Worker motivation in crowdsourcing - a study on mechanical turk. *Seventeenth Americas Conference on Information Systems.*

[5] Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambrige, MA: MIT Press.

Learning Support System to Facilitate Redesigning for Understanding Software Design Patterns

Hiroki OOE[a]*, Tomoko KOJIRI[b], Kazuhisa SETA[c]
[a]*Graduate School of Science and Engineering, Kansai University, Japan*
[b]*Faculty of System Science and Engineering, Kansai University, Japan*
[c]*Graduate School of Science, Osaka Prefecture University, Japan*
*k773152@kansai-u.ac.jp

Abstract: Design patterns are good designs in object-oriented programming and are generated experientially by predecessors. We propose a learning method of understanding design patterns by transforming a program with a design pattern into that without one (alternative solution). We also develop a support system that encourages learners of generating appropriate alternative solution. Experimental results proved that the proposed method was effective for a deep understanding of design patterns.

Keywords: experiential knowledge, design patterns learning support, generating alternative solution, vicarious experience support

1. Introduction

Experiential knowledge is knowledge acquired through experience [1]. Experiential knowledge addressed in this study is design patterns that are a collection of good object-oriented designs. Design patterns are produced through generation of various programs by our predecessors. To obtain design patterns, it is necessary to understand not only their meaning but also the appropriate and inappropriate conditions to apply. For grasping such conditions, to follow predecessors' generation process which is to transform programs with design patterns into those without design patterns (alternative solution) is effective. However, it is difficult for learners to modify a program into a reasonable one.

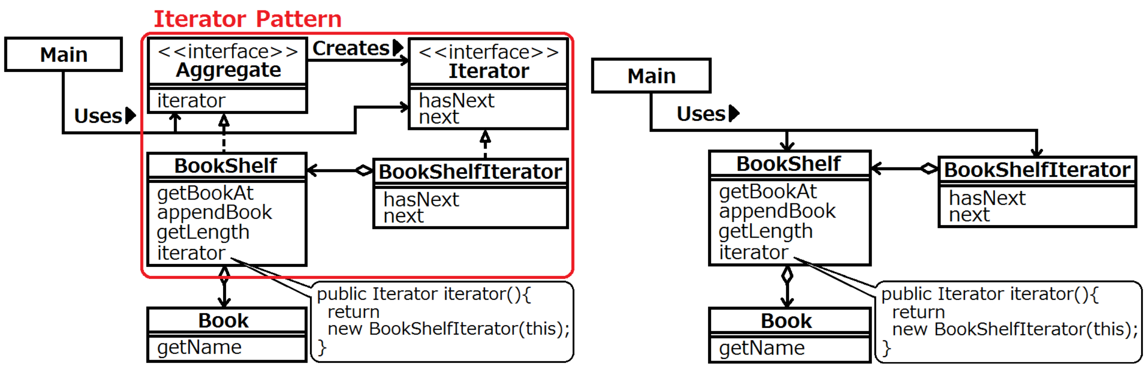
We propose a learning method of reliving predecessors' experience by producing alternative solutions. Additionally, we construct a support system to encourage learners of generating alternative solutions. Based on the method, learners can develop the ability of acquiring experienced knowledge, even if new design patterns are given.

Many studies support the learning of programming [2]. However, they cannot consider the conditions for applying the structure. Stephen introduces target problem to learners and makes them derive the design pattern by themselves as predecessors did [3]. However, in this method, some learners may not be able to derive the design pattern. In this study, in order for all learners to consider predecessors' trial and error easily, the learning method is proposed in which transforming the class diagram of the program using design patterns into alternative solution. We have constructed a learning support system in which the class diagram of the program can be transformed. In addition, the advising mechanism for supporting generation of the alternative solution is introduced.

2. Framework of Design Pattern

Design patterns are produced by experts of programming through their experiences of designing programs repeatedly for similar problems. Design patterns are structures that contain the benefits of object-oriented design, but alternative solutions do not. Therefore, for acquisition of design patterns, to generate alternative solutions that reduce the benefits of object-oriented design is effective to understand the advantages and disadvantages of the design patterns.

Example of alternative solution is shown in Figure 1. The class diagram in Figure 1(a) is the example of using the *Iterator* pattern [5]. The characteristic of this program is that *Main* class does not refer to the method in concrete class, such as *BookShelf* and *BookShelfIterator*, but uses the method in the *Aggregate* interface and *Iterator* interface. In this structure, even if adding or changing another concrete class having the same functionality as concrete class, then a program can be updated without changing the *Main* class. The example of its alternative solution is shown in Figure 1(b). This program does not introduce interface and *Main* class refers to the concrete class directly. In this structure, the change of the concrete method affects to the *Main* class. Therefore, this alternative solution reduces the reusability and flexibility of the original design.



(a) Class diagram of the solution using the Iterator pattern

(b) Class diagram of good alternative solution of (a) without using the Iterator pattern

Figure 1. Example of solution

In this study, learners transform the class diagram to the program structure of the same meaning as alternative solution of the design pattern. To make learners generate good alternative solutions, we have proposed a support system which evaluates the learner's class diagram of alternative solution and generates advice, if necessary. The system presents class diagrams of target design patterns that learners want to learn. Learners change the class diagram and design alternative solution without the design pattern. The system evaluates an inputted class diagram as the good alternative solution. The system gives advice to promote to the derivation of an appropriate alternative solution if the learners' alternative solutions are inappropriate. Learners change the class diagram by referring to advice and finish learning if they generate a good alternative solution.

3. Mechanism for Supporting Alternative Solution Generation

The system holds the data for class diagrams of the problems and class diagrams of good alternative solutions. Good alternative solutions can be expressed by the difference between class diagrams of the problem and class diagrams of alternative solutions. An alternative solution is created by adding/deleting objects/relations from the given problem. Modification of existing objects/relations is equivalent to adding modified objects/relations after deleting the original one, so it can be represented by a combination of addition and deletion. In this study, the system holds alternative solution data as added/deleted object and type of its change, such as addition or deletion. In addition, a good alternative solution is intended to reduce the benefits of object-oriented design. The system must hold benefits of corresponding alternative solution with class diagram data.

Alternative solution data for Figure 1(b) is shown in Table 1. Since it reduces the reusability and flexibility, “reusability and flexibility” are described as the target benefits of object-oriented design. Additionally, it deletes the relation between *Main* class and *Aggregate* interface, so such deletion and its type are noted as ID I. In the same way, all differences are listed.

Table 1. Alternative solution data of Figure 1(b)

Benefits of object-oriented design	ID	Changed object	Type of change
Reusability and flexibility	Ⅰ	Relation（Depend : Uses）: Main class⇒Aggregate interface	Deletion
	Ⅱ	Relation（Depend : Uses）: Main class⇒BookShelf class	Addition
	Ⅲ	Relation（Depend : Uses）: Main class⇒Iterator interface	Deletion
	Ⅳ	Relation（Depend : Uses）: Main class⇒BookShelfIterator class	Addition
	Ⅴ	Class : Aggregate interface	Deletion
	Ⅵ	Class : Iterator interface	Deletion
	Ⅶ	Relation（Depend : Creates）: Aggregate interface⇒Iterator interface	Deletion
	Ⅷ	Relation（Implement）: BookShelf class⇒Aggregate interface	Deletion
	Ⅸ	Relation（Implement）: BookShelfIterator class⇒Iterator interface	Deletion

The alternative solution inputted by the learner is also transformed to this form. The system compares the lists of problem and alternative solution in the system with the learner’s alternative solution and generates advice that helps learners to derive correct alternative solutions if the learner’s diagram is inappropriate.

The advice is derived depending on the types of inappropriateness in the learner’s class diagram. Learners who generated inappropriate solutions do not understand one of the following factors:

- Factor 1. Benefits of object-oriented design that should be reduced from the given class diagram
- Factor 2. The way of changing the class diagram to reduce benefits of object-oriented design

If learners add/delete unnecessary classes/relations, they might not understand factor 2. Such learners can generate correct class diagram if the system points out existence of the unnecessary addition/deletion. However, if the learner does not perform the necessary addition/deletion, they might stumble in both factors 1 and 2. Therefore, the system first tells learners the benefits of the object-oriented design and the part of the inappropriate solution that they should specifically address. The system advises a correct change to let them understand factor 2 if the learner still cannot generate a correct class diagram. The advice templates are shown in Table 2, where <Aspect> corresponds to the benefit of object-oriented design, and <Object> represents class or relation to be added/deleted.

Table 2. Advice template

Type of inappropriateness	Advice template
Lack of necessary addition	First advice: “Focus on <Object> for reducing <Aspect>.” The Second advice: “Add <Object> for reducing <Aspect>.”
Lack of necessary deletion	First advice: “Focus on <Object> for reducing <Aspect>.” Second advice: “Delete <Object> for <Aspect>.”
Extra addition	“Extra Change: delete <Object>.”
Extra deletion	“Extra Change: add <Object>.”

4. Prototype System

We implemented the proposed system using Microsoft Visual C++. The user interface is presented in Figure 2. The class diagram is displayed at the class diagram display unit. At the beginning of learning, the class diagram of the problem is shown, and learner's solution is displayed during the learning. The problem selection unit holds a list of learnable design patterns and learners can select one from it.

The entity edit unit represents a list of entities that exist in the class diagram. Learners can add and delete the entity from text fields. The relation edit unit represents a list of relations. Learners can also add, and delete them. The redraw button is used for reflecting the class diagram that is inputted in the entity edit unit and relation edit unit. The advice generation button is used for evaluating the class diagram created by the learner and for asking advice. When its button is clicked more than once, the former advice is replaced by the new advice. Figure 2 shows an example of the provided advice. This advice shows "deleting the unnecessary class *ooe* whose field is *hiroki* and whose methods are *kansai*, *university*, and *osaka*".

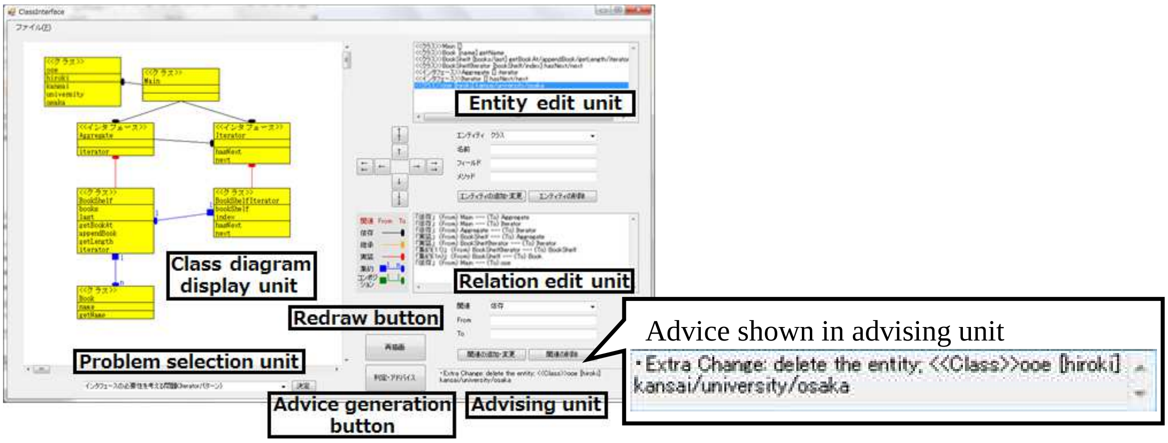


Figure 2. User interface

5. Experiments

We conducted an experiment to verify the effectiveness of the proposed methods. Subjects were six graduate and undergraduate students who had not learned design patterns, or who were unable to apply design patterns to the problem. First, subjects were asked to take a pretest to confirm their understanding of design patterns. The pretest has three questions: question 1 was used to create an alternative solution from the given class diagram with design pattern; question 2 asks the reason for created alternative solution in question 1; and question 3 asks the role of the focused class/relation in the design patterns. After explaining the manipulation method of the system, subjects were asked to create a class diagram of alternative solution using the system. After that, subjects took a posttest to verify the change of understanding of design patterns. The format of the posttest is the same as that of the pretest, but the target design pattern is different. Additionally, we prepared a questionnaire to evaluate the performance of this learning method.

Results of the pretest and the posttest are shown in Table 3, which presents the number of subjects for each situation. In the pretest, most subjects could not answer all the questions correctly. In contrast, all subjects were able to answer question 1 and numerous subjects were able to derive answers for questions 2 and 3 in the posttest. Since many subjects were able to understand the benefits of object-oriented design, our system is effective for learning design patterns.

Table 3. Results of pretest and posttest

Question	Result	Pretest	Posttest
1	Created appropriate alternative solution	1	6
	Created inappropriate alternative solution	4	0
	Could not create at all	1	0
2	Explained correctly	1	3
	Explained incorrectly	2	3
	Could not explain at all	3	0
3	Explained correctly	1	5
	Explained incorrectly	2	1
	Could not explain at all	3	0

Questions of the questionnaire and the results are shown in Table 4. They were asked to select one from 1 (strongly disagree) to 5 (strongly agree). The table shows the number of subjects who select each evaluation value for each question. From questions 1 and 2, it seems clear that subjects can create the solution, but that they cannot understand the meaning of created alternative solution. We should revise the advising mechanism to make learners understand the meaning of the created alternative solution. Results of questions 3 and 4 show that subjects were able to understand the effect of the learning through creating alternative solution. However, only half of them want to use this learning method. The complexity of manipulating our system makes subjects feel this learning method is inconvenient. Therefore, it is necessary to improve the system usability. Additionally, we need further evaluation to prove the effectiveness of our system with more subjects.

Table 4. Questions and results of questionnaire

Question	Evaluation				
	1	2	3	4	5
1. Did you understand the meaning of alternative solution which you created?	0	2	1	1	2
2. Did you think given advices were appropriate?	1	0	4	1	0
3. Do you understand the importance of creating alternative solution?	0	0	0	5	1
4. Do you want to create alternative solution in learning new design patterns?	0	2	1	2	1

6. Conclusion

In this paper, we proposed a design pattern learning support system by creating alternative solutions. In the future, the advising method should be improved to make explain the role of the target classes and relations deeply. Additionally, the system interface must be modified to edit the class diagram more smoothly.

References

[1] Polanyi, M. (2003). The Tacit Dimension, trans. Isao, T., *Thikuma Library* (in Japanese).

[2] Hwang, W., et al. (2008). A Web-based Programming Learning Environment to Support Cognitive Development, *Interacting with Computers*, 20(6), 524-534.

[3] Stephen, W. (2005). Teaching Design Patterns by Stealth, *Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education*, 37(1), 492-494.

[4] Seta, K., et al. (2010). Presentation-based Collaborative Learning Support System to Facilitate Meta-cognitively Aware Learning Communication, *The Journal of Information and Systems in Education*, 9(1), 3-14.

[5] Erich, G., et al. (1994). Design Patterns: Elements of Reusable Object-Oriented Software, *Addison-Wesley Professional*.

A Framework for Identifying Working Memory Capacity from the Log Information of Learning Systems

Ting-Wen Chang[a*], Moushir M. El-Bishouty[a,b], Sabine Graf[a] & Kinshuk[a]
[a]*Athabasca University, Canada*
[b]*City for Scientific Research & Technological Applications, Egypt*
*tingwenchang@athabascau.ca

Abstract: Working memory capacity (WMC) is a cognitive trait that affects students' learning behaviors to perform complex cognitive tasks such as reading comprehension, problem solving, and making decision. Considering students' WMC when providing them with course materials and activities helps in avoiding cognitive overload and therefore positively affects students' learning. However, in order to consider students' WMC in the learning process, an approach is needed to identify students' WMC. To address this problem, we introduce a general framework to automatically identify WMC from students' behavior in a learning system. Our approach is generic and designed to work with different learning systems. It connects to the learning systems' database and extracts students' behavior data to analyze them for indications about their WMC. The proposed approach has been implemented as an extension to a tool for detecting learning styles, enabling this tool to additionally identify students' WMC. By knowing students' WMC, teachers can provide meaningful recommendations to support students with low and high WMC. Furthermore, such information is the basis for designing adaptive systems that can automatically provide students with individualized support based on their WMC.

Keywords: Working Memory Capacity, Student Modeling, Learning Systems

Introduction

Working memory capacity (WMC) is an individualized ability of the human brain to keep active a limited amount of information for a very brief period of time. In recent years, researchers have found that WMC can affect attention control and performance of cognitive tasks [4][7]. Results of these studies have shown that students with low and high levels of WMC have very different performances on different attention levels during performing cognitive tasks. If the students' cognitive load is too high, it will affect them in learning effectively. Knowing the levels of students' WMC can help in many ways to enhance learning and teaching in learning systems. Getting information about students' WMC can be used as input for adaptive systems to provide students with customized learning content and activities to suit their individual WMC.

This paper introduces a student modeling approach and a detection tool to identify students' WMC from the activity log data of learning systems. The student modeling approach profiles students from the activity log information available in the learning systems and identifies students' learning behaviors, including linear navigation, constant reverse navigations, and performing simultaneous tasks. Furthermore, a learning style detection tool, DeLeS [8], has been extended in order to detect not only learning styles but also WMC. By extending DeLeS, all the functionality that is needed to get the data from any learning system is used. Then, different preprocessing and calculation procedures are applied in order to identify both WMC and learning styles. The approach and the tool are developed in a generic way and are therefore applicable for any learning system.

The next section presents related studies about WMC and approaches to identify students' WMC. In section 3, our concept for identifying WMC is introduced, including explanation on the preprocessing steps, the relevant behavior patterns of WMC detection, and the transition from the behavior data to WMC. The architecture for identifying WMC from students' behavior is discussed in section 4. Section 5 concludes the paper.

1. Literature review

1.1 Working Memory Capacity

The human memory system works similar to an information processing system and operates like an advanced computer system [2]. Atkinson and Shiffrin [2] proposed a memory model including three types of memory: sensory memory, short-term memory, and long-term memory. They also pointed out that information is received by sensory memory to arrive in another temporary store called short-term memory (STM) or working memory (WM). STM and WM clearly share a close relationship referring to transient memory. However, they also have different definitions in terms of empirical and conceptual distinctions [11]. The capacity of STM is typically accessed via the immediate serial recall of a list of information. Miller [13] proposed the "magical number seven" in 1956 to give the earliest quantification of the capacity limit associated with short-term memory. The WM is used to hold information actively in the mind and to manipulate that information to perform a cognitive task. WMC refers to the processing of a limited amount of information in transient memory storage for a short time [3]. The works of Woehrle and Magliano [14] have focused on identifying individuals' WMC from different aspects, such as reading comprehension, academic achievement, and attention control. They also suggested in their study that working memory may be a core cognitive ability that underlies and constrains our ability to process information across cognitive domains. WMC is also crucial to many learning activities in online learning because students have to hold information while engaging in an online learning activity.

1.2 Identifying and Considering WMC in learning systems

For enabling rich adaptivity, the student model is an important part of learning systems [8][9]. A student model in an adaptive learning system tracks students' information based on the system's beliefs about students. The process of building and updating a student model is called student modeling. Chen [5] focused on modeling and scaffolding students' cognitive skills related to learning from work-out examples as well as from their exploration activities. Conati and Maclaren [6] analyzed students' browsing data recorded in the log file of a web-based learning system to conclude that students' cognitive style (field dependence and independence) and learning behaviors are related. Cognitive trait model (CTM) is another student modeling technique that profiles students according to the four cognitive traits: working memory capacity, inductive reasoning ability, associative learning skills, and information processing speed [12]. In the CTM, certain learning behaviors, called Manifestation of Traits, are used to infer students' cognitive traits from the students' behaviors in an online course. While the CTM focuses on detecting cognitive abilities in a particular system with a predefined course structure and types of learning objects, our approach aims at identifying cognitive abilities in learning systems in general.

2. Concept of WMC Detection

This section describes our concept of WMC detection from students' continuous behavior in any learning system. In order to analyze students' behavior and detect relevant behavior patterns, some preprocessing of behavior data and course data has to be done. More concretely, the preprocessing includes (1) the identification of learning sessions, (2) filtering out activities that are not dedicated to learning as well as activities where students visit a learning activity only for very short time, and (3) building a Learning Sequence Table called LSEQ table that includes structure of the course in terms of the predefined sequence of learning activities/objects in a course. In learning situations, there are several behavior patterns known in literature that give indications for a person's WMC [9][12]. Three mplementation patterns (IPs) in terms of learning objects, properties, and types of navigational behavior which can then be implemented domain independently. These patterns are linear navigation pattern, constant reverse navigation pattern, and performing simultaneous tasks pattern. These types of navigational behavior are described by a relation function, R(preLO, currLO). This function relates two learning objects (LOs), the source (preLO) and the destination (currLO). The following paragraphs describe the three IPs, including a brief discussion on their effects on WMC, a definition and example of each IPs.

With respect to the IP of linear and non-linear navigation, linear navigation means that students learn the materials linearly and follow the learning sequence of the course defined by teachers. Huai [10] found that students with high WMC tend to focus on linear navigation and students with low WMC tend to use non-linear navigation. If a student can learn materials linearly, her/his working memory is able to deal with the consecutive information easily [10]. For example, when LO B is learned, and the previous LO of LO B (defined in LSEQ table), LO A, has been learned before, linear navigation is found, no matter whether other LOs are visited between LO A and B. If this linear navigation is found, it gives an indication for high WMC. Otherwise, non-linear navigation will be found, which gives an indication for low WMC.

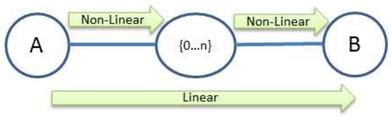


Figure 1. A sample of linear navigational behavior

With respect to the IP of constant reverse navigation, reverse navigation includes revisits of already visited LOs. Constant reverse navigation indicates that a student frequently goes back to an already visited LO. The process of constant reverse navigation is caused by an insufficient WMC to hold on the materials that have just been visited [12]. When the learning materials that a student just read on the previous page should still be fresh and in his/her working memory, the constant need to navigate backwards is a sign of working memory deficiency. The definition of constant reverse navigational behavior is that there are more than two LOs revisited in the same learning session and the navigational relations of these LOs are not defined in the LSEQ table (and therefore not in line with the sequence of LOs in the course structure). Figure 2 shows a sample of constant reverse navigational behavior including the following relations of navigation: *R(A, B), R(B, C), R(C, D), R(D, A), R(A, C)*. In these navigational relations, two relations, *R(D, A)* and *R(A, C)*, are not defined in the LSEQ table and the two destination LOs, A and C, are revisited. Thus, the constant reverse navigational behavior is found, which gives an indication for low WMC.

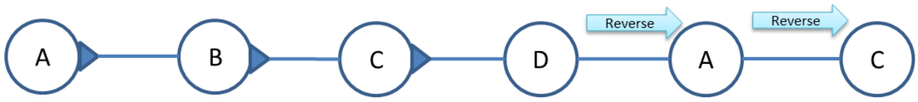


Figure 2. A sample of constant reverse navigational behavior

The IP of performing simultaneous tasks pattern is transferred from the MOT describing the ability of attentional control on performing two tasks simultaneously. The results of previous studies showed that when performing two tasks simultaneously, low WMC participants were less accurate than participants with high WMC [7][14]. For identifying this pattern, the overlaps navigational behavior is investigated which indicates that a student tries to perform two tasks simultaneously. As shown in Figure 3, if a student visits at least one other LO in between *LO A* and its evaluation, *EA*, overlaps navigational behavior is found. In such a situation, the student learns *LO A* first and then learns other LOs before taking the evaluation of *LO A*. Therefore, she/he will need to remember the concept of *LO A* in her/his working memory while learning other LOs. If the student then passes the evaluation of *LO A*, the simultaneous tasks pattern is found, which gives an indication for high WMC; if she/he fails, the non-simultaneous tasks pattern will be identified, which gives the indication for low WMC.

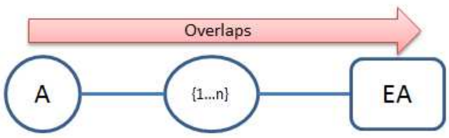


Figure 3. A sample of overlaps navigational behavior

3. An architecture for identifying WMC from learning systems

The proposed architecture for identifying WMC from any learning system is based on the architecture of DeLeS, a tool developed to automatically detect learning styles in any learning system. The proposed architecture (and DeLeS' architecture) consists of two components: the data extraction component and the calculation component. In order to identify WMC, the data extraction component of DeLeS is extended and extracts the learning sessions of each student, the learning activities of each learning session, and the LSEQ table of each course. These extracted data are then passed on to the calculation component, which is responsible for calculating the students' WMC based on the three IPs and the corresponding navigational behaviors of students. If a navigational behavior is detected in a relation between two LOs, this relation is considered as an activated relation for the particular behavior. In each learning session, a value is calculated for each pattern based on the number of activated and non-activated relations in this session. This value shows how strongly the student's behavior represents the respective pattern. Subsequently, each value is transferred to its indication for WMC (e.g., a high value for linear navigation provides an indication for high WMC) and then the indications for all patterns are summed up and divided by the number of patterns. The result of this calculation represents the indication for WMC of this session. Each learning session also contains a weight, which determines the influence of each session on the overall value of WMC and is calculated based on the number of activated relations in a session for all patterns. In order to calculate the student's WMC, the WMC indication of each session is multiplied by the weight of the respective session. Subsequently, the results for all sessions are summed up and divided by the number of sessions.

4. Conclusions

As identified in past studies [1][14], different levels of WMC have potential to affect students' learning performances. This paper introduced a framework for identifying students' WMC from their activity log data in learning systems. The proposed framework is based on a student modeling approach in order to identify different levels of WMC, and extends the DeLeS tool to analyze activity log data from a variety of learning systems [8]. Therefore, the framework is not restricted to a particular learning system. The DeLeS tool is used to preprocess the log data. Then, these data are used to find three navigational behaviors: linear navigation, constant reverse navigation, and overlaps navigation, and calculate students' WMC from these behaviors. Future research will deal with investigating the use of additional behavior patterns to be added to our student modeling approach. Furthermore, we will use the information about students' WMCs to provide teachers with recommendations for designing and improving learning contents and course presentation.

Acknowledgements

The authors acknowledge the support of NSERC, iCORE, Xerox, and the research related gift funding by Mr. A. Markin.

References

[1] Antonenko, P. D., & Niederhauser, D. S. (2010). The influence of leads on cognitive load and learning in a hypertext environment. *Computer in Human Behavior, 26*(2), 140-150.

[2] Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposal system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The Psychology of Learning and Motivation* (pp. 89-195). New York: Academic Press.

[3] Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

[4] Broadway, J. M., & Engle, R. W. (2011). Lapsed attention to elapsed time? Individual differences in working-memory capacity and temporal reproduction. *Acta Psychologica, 137*(1), 115-126.

[5] Chen, L.-H. (2010). Web-based learning programs: Use by learners with various cognitive styles. *Computers & Education, 54*(4), 1028-1035.

[6] Conati C., & Maclaren H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User Adapted Interaction, 19*(3), 267–303.

[7] Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Current Anthropology, 51*(1), Working Memory: Beyond Language and Symbolism (June 2010), 17-26.

[8] Graf. S., Kinshuk, & Liu, T.-C. (2009). Supporting teaching in identifying student's learning styles in learning management system: An automatic student modelling approach. *Educational Technology & Society, 12*(4), 3-14.

[9] Garf, S., Liu, T., Kinshuk, Chen, N.-S., Yang, S. J.-H. (2009). Learning styles and cognitive traits – Their relationship and its benefits in web-based educational systems. *Computers in Human Behavior, 25*, 1289-1289.

[10] Huai, H. (2000). Cognitive style and memory capacity: effects of concept mapping as a learning method. *Doctoral Dissertation*, Oct 2000, Twente University, The Netherlands.

[11] Hutton, U. M., & Towse, J.N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory, 9*(4-6), 383-394.

[12] Lin, T., Kinshuk, & Patel A. (2003). Cognitive trait model – a supplement to performance based student models. In *Proc. of International Conference on Computers in Education* (pp. 629–632). Hong Kong.

[13] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

[14] Woehrle, J. L., & Magliano, J. P. (2012). Time flies faster if a person has a high working-memory capacity. *Acta Psychologica, 139*(2), 314-319.

Quick View of Descriptive Answers of Quiz by Auto-Extracted Keywords

Haruhiko TAKASE[a*], Hiroharu KAWANAKA[a] & Shinji TSURUOKA[b]
[a]*Graduate school of Engineering, Mie University, Japan*
[b]*Graduate school of Regional Innovation Studies, Mie University, Japan*
*takase@elec.mie-u.ac.jp

Abstract: In this paper, we develop a system that helps teachers to make an effective improvement based on answers of a quiz in descriptive answer style. Grasping students' understandings is necessary for effective improvements for teachers. The system enables teachers to grasp understandings even before students finish their answers. First, it provides keywords that are automatically extracted from answers. Then, it shows related phrases or related answers according to teachers' selection of keyword. As a result, teachers can grasp interested phrases quickly. It makes improvement of classes effective.

Keywords: Quiz, e-Learning system, text mining, keyword extraction

Introduction

Grasping students' misunderstandings are important for teachers to make their lecture effective. But, it is hard for teachers in large classes. Most easy way to grasp students' misunderstandings is to make a quiz. Some teachers use clickers, which allow students to answer by using electronic devices [1]. With clicker, students can answer for true/false questions, multiple choice questions, or numeric questions. Teachers easily grasp responses of all students. Since clicker accepts only selective answers or numerical values, teachers need to prepare quizzes carefully to grasp students' misunderstandings correctly. Quizzes in descriptive answer style would be preferred for such purpose. Students need to answer such kind of quizzes with their own words.

Though the effectiveness of quiz in a descriptive answer style, it is hard for teachers to grasp all answers in a short time. As a result, some teachers avoid quizzes in descriptive answer style, though its effectiveness. E-learning systems, such as Moodle [2] or Blackboard Learning System [3], can summarize answers in short time by using information technologies. Many researchers have developed systems to analyze/visualize answers, especially descriptive ones. Since sets of answers have difference characteristics with general document collections, some methods that are based on general text mining techniques have a room for improvement.

In this article, we propose a supporting system for teachers to quiz in a large class. The system enables teachers to grasp descriptive answers in short time. It means that they can improve their classes just after quizzes. Teachers can grasp answers by phrases that include a keyword, which is suggested by the system automatically. The system does not require any model answers, teachers can quiz without special preparations. To develop such system, we discuss an interface of grasping answers and a method to suggest keywords.

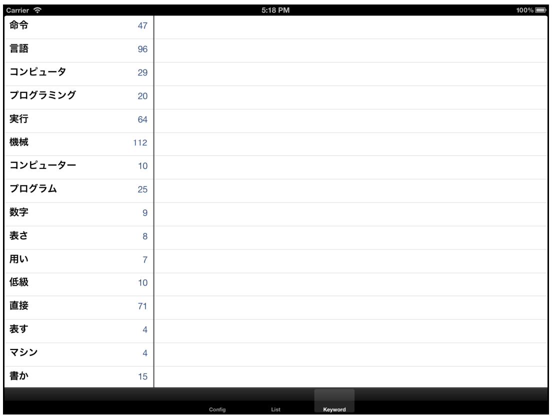
1. Requirements to analyze descriptive answers of a quiz

In this section, we discuss various text-mining techniques for descriptive answers, and state our problem. There are many techniques to support teachers with many documents submitted by students. Ishioka et al. developed JESS (Japanese Essay Scoring System) [4]. It scores Japanese essays by three features: rhetoric, organization, and content. It provides a score and a diagnosis for each essay. Villalon et al. developed Concept Map to visualize conceptual understandings [5]. It visualizes concepts and their relation as a map from students' compositions. Though these techniques are useful for their purpose, they may not work well for our purpose. They do not provide information related to misunderstandings directly.

To develop effective supporting system for quizzes, there are three requirements as follows. The first requirement is that the system provides useful information related misunderstandings of students. It is just our purpose. The second requirement is that the system can accepts incomplete answers to analyze. It is related to the quick improvements as mentioned above. The third requirement is that the system does not require any additional preparations for a quiz. It is related to the advantage of quizzes in descriptive answer style. Teachers can quiz with less preparations than other answer style, they only require a question: no model answers, or no choices for students. The requirement implies to keep this advantage. In addition, we assume all answers are in Japanese. Most of conventional methods do not satisfy the second requirement. They perform deep analysis, which need complete answers.

2. E-Learning system for quizzes and its improvement

We propose a new interface for teachers as shown in fig. 1. It consists of three views: keyword view, phrase view, and answer view. At first, the system shows a keyword view. The view provides a list of words in order of their importance. Here, "importance" represents how effective the word is to grasp misunderstandings. We discuss it in the next chapter. A teacher selects a keyword on the keyword view. Then, the teacher gets related phrases on the phrase view. Finally, the teacher can read whole answers that contain selected phrases on the answer view. We explain this flow with fig. 1. It shows the result of a quiz "*Explain the term*: *machine language*" after 7 minutes elapsed (not finished). 80 students are answering the quiz. Fig. 1 (a) is the keyword view. The view provides keywords with their frequency in the order of importance. They are *instruction* (47 times), *language* (96 time), *computer* (29 times), and so on. After a teacher selects keyword (2nd word: *language*), the system provides the phrase view on the right side of keyword view as fig. 1 (b). On the phrase view, the system provides phrases that are heading/tailing with the selected keyword with their frequency. In this case, there are 5 phrases tailing with *language* and 7 phrases heading with *language*. In this case, the system extracts a train of 6 morphemes as a phrase, and shows only phrases that are appeared multiple times. When the teacher selects a phrase, the system show answers that contain the selected phrase in answer view as fig. 1 (c). In this case, the system shows three answers that contain the phrase *be a language*. With the proposed interface, teachers can grasp misunderstandings without confused by a flood of characters. They follow a their flow to marking answers naturally. It would useful for grasping answers.



(a) Keyword view

(b) Phrase view

(c) Answer view

Figure 4: Teacher's view of proposed system

3. Keyword extraction technique

Since the new interface provides only keywords at first, the extraction of keywords decides the effect of it. In this section, we propose a keyword extraction method for our new interface. Because of the second requirement, we cannot use complicated techniques to estimate importance of each word. Fortunately, MeCab [6], a famous Japanese language morphological analyzer, works even for incomplete answers. By using MeCab, we get morphemes divided from a given text and a word class of each morpheme. We try to estimate importance of each word (a subset of morphemes) from this information. Since the purpose is to extract keywords, we focus only on nouns, verbs, and adjectives in the following discussion.

In general, words that appear in various documents are common words. On the other hand, words that appear only in particular documents may be misspelled words. It is not similar for answers. Words that appear in many answers would be essential words to answer the quiz. Fig. 2 shows this relationship. Here *corpus* means the set of general documents, and *answers* means the ones for particular quiz.

Based on this idea, we estimate the importance of a word by using the function as shown in fig. 3. The figure is a contour graph of the importance corresponding to two input values, the frequency in corpus and the frequency in answers. We use a radial basis function as the function. Here, each frequency is regularized into the range [0,1]. By using this function, the system can automatically estimate importance of each word only by frequency of each word. In addition, the system accept incomplete answers for analysis.

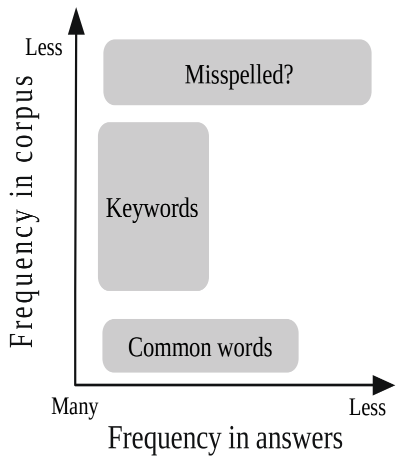


Figure 2: Relationship between type of words and their frequency

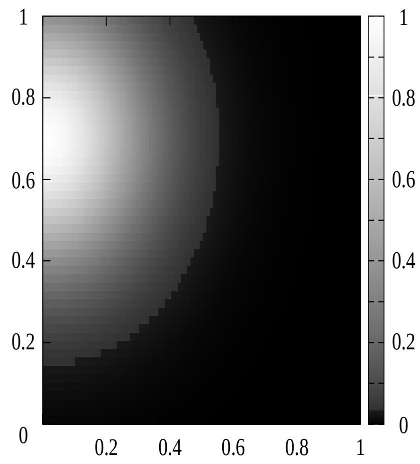


Figure 3: Evaluation function for importance of a word

4. Demonstration

In this section, we demonstrate auto-extraction of keywords. To estimate importance of each word, we use Google's n-gram data [7] as a corpus. It shows frequency of each morpheme in all web pages that are crawled by Google at June 2007. There are 2,565,424 morphemes from 20,036,793,177 sentences.

We apply our system to answers of following quizzes in the course "Introduction to Computer Engineering I and Exercise" for the 1st grade students in our department. Since all quizzes are done in Japanese, we show translated answers in this article.

Quiz 1: *Why is high level programming language needed? Answer with three keywords: Machine language, Program, and Binary code.*
This quiz is an example that keywords are indicated. A sample answer is that "*It is hard for programmers to develop programs in machine language, which is in binary code, directly. They develop programs in high level programming languages that adopt human friendly elements to use.*" There are 80 answers for this quiz.

Quiz 2: *Explain the term "Compiler" in the broad sense.*
This quiz is an example that does not have any requirements for expression. A sample answer is that "*A program that translates source codes in a high level programming language into object code, which is based on a machine language.*" Keywords would be *translate, source code*, and *object code*, and so on. There are 83 answers.

Table 1 shows the result of estimation, which shows top 10 words in importance. Table 1 (a) shows that keywords indicated by the question are ranked in top 10 words. Table 1 (b) contains expected keywords. In addition, keywords in these tables would be useful for answering each question. These results show a validity of our proposal.

Table 1: Estimated Importance

(a) Quiz 1

Word		Importance
Language	(言語)	0.775
Computer	(コンピュータ)	0.714
Programming	(プログラミング)	0.699
Machine	(機械)	0.577
Convert	(変換)	0.572
High level	(高級)	0.572
Execute	(実行)	0.547
Program	(プログラム)	0.502
Use	(用い)	0.445
Description	(表記)	0.397

(b) Quiz 2

Word		Importance
Convert	(変換)	0.827
Programming	(プログラミング)	0.824
Translate	(翻訳)	0.759
Object	(オブジェクト)	0.738
Compile	(コンパイル)	0.723
Language	(言語)	0.720
Compiler	(コンパイラ)	0.678
In the broad sense	(広義)	0.639
Computer	(コンピュータ)	0.631
Source	(ソース)	0.557

5. Conclusion

In this article, we aim to develop a supporting system for quizzes, which are regard as a method to aware students' misunderstandings. The system supports for teachers to grasp misunderstandings quickly. It satisfies three requirements: (1) provides effective information, (2) accept incomplete answers, and (3) do not need special preparations. The proposed system provides keywords, phrases, whole answers in a step-by-step manner as necessary. Teachers would find students' misunderstandings quickly, since they can get necessary information in each phase. In addition, we discuss the method to extract keywords automatically. The proposed method estimates importance of each word only by its frequency in answers and a corpus. As a result, teachers do not need to prepare for analysis, and the system extracts keywords even during a quiz.

Acknowledgements

This research was supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 23501106, 2011.

References

[1] Sally A. Gaucii, Arianne M. Dantas, David A. Williams, & Robert E. Kemm (2007). Promoting student-centered active learning in lectures with a personal response system. *Advances in Physiology Education*, 33, 60–71.

[2] Moodle.org: open-source community-based tools for learning (retrieved May, 2012), http://moodle.org/.

[3] Blackboard International (retrieved May, 2012), http://www.blackboard.com/.

[4] Tsunenori Ishioka, & Masayuki Kameda (2006). Automated Japanese Essay Scoring System based on Articles Written by Experts. *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, 233–240.

[5] Jorge Villalon, & Rafael A. Calvo (2011), Concept Maps as Cognitive Visualizations of Writing Assignments. *Educational Technology & Society*, 14(3), 16–27.

[6] Taku Kudo, Kaoru Yamamoto, & Yuji Matsumoto (2004). Applying Conditional Random Fields to Japanese Morphological Analysis. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, 230–237.

[7] Taku Kudo, & Hideto Kazawa (2007). Web Japanese N-gram Version 1. *Gengo Shigen Kyokai*.

Designing Lessons with an Ontological Modeling Approach in Practice

Yusuke HAYASHI[a*] & Riichiro MIZOGUCHI[b]
[a]*Faculty of Engineering, Hiroshima University, Japan*
[b]*Research Center for Service Science, JAIST, Japan*
*hay@icts.nagoya-u.ac.jp

Abstract: In this paper, we discuss the effectiveness of an ontological modeling approach in practice. Although a lesson plan is a document that describes the plan of a lesson, the design rationale behind it tends to be implicit. The authors have developed an ontology called OMNIBUS and a theory-aware authoring system called SMARTIES. This paper describes the result of the experimental use of them in the real task of schoolteachers.

Keywords: instructional design, ontological engineering, and authoring system

Introduction

Teachers develop their teaching skill in their practice through self-reflection and discussion with other teachers or experts [1][2]. "Lesson study" is a systematic activity to foster such development in a group setting [5]. Lesson study aims to improve design of a lesson (lesson design) through discussion among teachers before the teacher carries out the lesson and to bring reflection through evaluating the lesson after the lesson. In designing a lesson a teacher makes a document called "lesson plan". This is the description of a lesson design and a document for sharing it among teachers. If a teacher can make a lesson plan faithfully reflecting the lesson design in his/her mind, teachers can share the design rationale of the lesson. However, as most of the lesson plans describe mainly concrete activities of teachers and learners, the design rationale underlying lesson plans is often unclear.

This study aims to help teachers make high-quality lesson design and reflect it on lesson plans through an ontological engineering approach [6][7]. The task of making lesson plan consists of the following two subtasks: considering the content and expressing it in a format. This study considers that a difficulty in making lesson plans faithfully reflecting lesson design in a teacher's mind is caused by doing the two subtasks at the same time in design process. Therefore, the approach of this study is to separate these two tasks clearly. This study sets an objective to achieve the goal. It is to enable teachers to faithfully describe lesson design in their head without the constraints of lesson plans. This helps them to check the validity of lesson design and improve it and then to reflect the lesson design to a lesson plan sufficiently.

This paper discusses the effectiveness of OMNIBUS ontology and SMARTIES authoring system [3] in practical lesson design activities with the result of practical experiences that the authors have conducted with an official research group of schoolteachers of Tokyo prefecture in Japan, named "ToChuSha". The group consists of only practicing schoolteachers of all ages that are from novices to experts. The authors conducted thier practical study when they were preparing for presentation of lesson plans and demonstrations of lessons according to them at an annual domestic conference on educational research of social studies in junior high school in Japan. This paper discusses the results of the practical study from the viewpoint of changes of lesson plans by introducing OMNIBUS and SMARTIES in designing a lesson.

The structure of the rest of this paper is as follows. The next section gives an overview of OMNIBUS and SMARTIES and defines the role of them in this study. The third section explains how to introduce them in practical lesson design activity. The fourth section discusses findings from this practical study. The final section concludes this paper.

1. Lesson Design Supported by OMNIBUS and SMARTIES

Figure 1 shows a screenshot of SMARTIES that displays a model of a process of learning and instruction based on OMNIBUS. In SMARTIES, the process is represented in the form of a tree-structured graph of learning goals. This structure represents the sequence of learning and instruction from left to right in a unit of learning such as a lecture or a learning session in a learning content. The root represents the goal of the unit of learning, and the bottom sequence represents concrete interaction between the instructor and the learners. A node is called "I_L event" (instructional and learning event) and a vertical link between them is called "WAY". A tree-structured graph composed of I_L events and WAYs is called "I_L scenario model". By the combination of these two concepts the hierarchical structure represents the design rationale of the sequence. That is to say the intention of each I_L event is represented by the upper one linked with a WAY and all the concrete interaction represented by the bottom I_L events are rationally linked to the goal of the unit of learning.

The essential of learning and instructional process model based on OMNIBUS is a distinction between learning goals and ways to achieve them. This distinction enables to manage a diversity of learning and instructional methods. There can be many methods to achieve a learning goal, and there is a method that can achieve some different learning goals. This approach can organize relationship between a variety of learning goals and methods to achieve them.

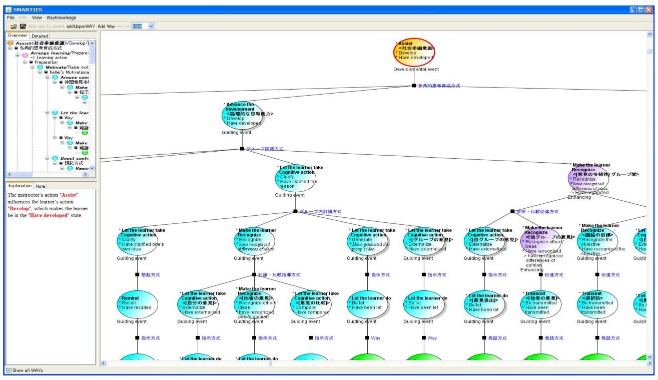


Figure 1 A screenshot of SMARTIES

2. Practice of Lesson Design with OMNIBUS and SMARTIES

The authors made some field trials to use OMNIBUS and SMARTIES in designing lessons in ToChuSha. The goal of these trials is to confirm the following hypotheses formed in this study;

1. Making I_L scenario models enables teachers to make lesson design clearer.
2. I_L scenario models help teachers to improve lesson design by considering alternative learning and instructional methods.

Previous study [4] illustrates the potential to support the former. This time, the authors made some field trials in order to find further support for the hypothesis and analyzed the results quantitatively. On the other hand, this study confirms the latter by analyzing alternative WAYs made in the I_L scenario models in these field trials.

In the field trials SMARTIES mainly played a role of a tool to describe design rationale of lessons made by teachers of ToChuSha. The major goal of the activity of ToChuSha is to make use of the results achieved up to now by them. Therefore, the priority is, rather than to make use of learning and instructional theories, to improve instructional methods they have used after clarifying the design intention of lessons. The authors repeat the following procedure in the filed trials with teachers in ToChuSha.

1. **A teacher makes a lesson plan:** the teacher designs a lesson and then describes it as a lesson plan.
2. **The author makes an I_L scenario model from the lesson plan:** the author presumes the lesson design from the lesson plan and makes an I_L scenario model.
3. **The teacher confirms the model:** the teacher checks whether the model reflects the design that the teacher has considered when he has made the lesson plan.
4. **The teacher discusses the lesson design with the author:** the teacher and the author check the validity of the design and try to improve it if desired.
5. **The teacher updates the lesson plan:** the teacher updates the lesson plan according to the I_L scenario model.
6. **The teacher discusses the lesson plan with the other teachers in ToChuSha:** the teacher suggests the plan reflecting the result of the discussion to the other teachers in ToChuSha and asks for feedback. And then go back to the second step.

In this procedure, from the second to the fifth steps are differences from the usual procedure that teachers in ToChuSha. That is, from the second to the fifth steps are the additional steps to investigate the effectiveness of OMNIBUS and SMARTIES.

The authors conducted this procedure on six lessons made by teachers in ToChuSha. Two of the lessons are for presentation at a domestic annual conference of teachers of social studies, and the others are for lesson studies in their schools or the school board. The number of times of this procedure the author could carry out differs from one lesson to another because it was necessary to follow their schedule. The highest number is five times for a lesson while there are only one or two chances for the rest.

3. Findings from Modeling Lesson Design in Practice

We officially summarized findings from the field trials with ToChuSha as follows:

A) **Clarification of the design rationale of lessons:** the design rationale that has not been described or described implicitly in the lesson plan but planed in the teacher’s mind is described more explicitly in the I_L scenario model.

B) **Improvement of lesson design:** lesson designs are improved through discussions between the teacher and the author based on both of the I_L scenario models and past achievements of ToChuSha.

This section explains these findings with some data or examples. Note that the main topic of this section is not the quality of the resultant lesson plan or the originality of learning and instructional methods included in it. What we will discuss are the activities by teachers for careful consideration for improvement of lesson design.

Nevertheless some subject matter expert evaluated the resultant lesson plan. Firstly, ToChuSha authorized it. Members of ToChuSha accepted the lesson plan supported by OMNIBUS and SMARTIES, and then published it. Secondly, the teacher that has made the lesson plan demonstrated a lesson according to the plan at an annual domestic conference on educational research of social studies in Japanese junior high schools. At the conference, there was a reviewer for the lesson demonstrated. He highly appreciated it as well-designed one with a clearly defined position in the curriculum. Consequently, although the quality of the resultant lesson design did not undergo quantitative evaluation, the quality is ensured to a certain extent because some subject matter experts properly assessed it.

In the field trials, not the teachers but the author made I_L scenario models as stated in the previous section. The teachers checked whether the author translated the original lesson plans into the models faithfully. Then, the teachers and the authors made discussion for improving the lesson design. Through this process, the teachers and the authors clarified lesson design in the teachers' mind and then improved it.

A lesson plan describes the goal of a lesson, the aim of instruction, and concrete activities of learners and teachers. It is considered that it is a description of the result of the teacher's consideration of lesson design. The consideration includes, for example, the consistency between the goal of the lesson and concrete activities of teachers and learners, alternatives of learning and instructional methods can be adopted and so on. The authors tried to expose such information that tend to be implicit in lesson plans and made I_L scenario model according to it through interviews from teachers.

Table 1 shows improvement process of the lesson plan in terms of number of items in a lesson plan and concordance between the items and the I_L scenario model made from it. This table indicates that, in essence, both of the number of items in the lesson plan and the concordance rate are increasing step by step. This can be considered that the teacher updated the lesson plan in a reflection of improvement of the lesson design described as the I_L scenario model. In fact, the teacher commented that he could update the description of the lesson plan by reconfirming the lesson design with the scenario model. Thus, this suggests that the increase of the number of the I_L events means the progression of externalization and improvement of lesson design in his mind. In addition to that, this also suggests that the increase of the number of items in the lesson plan means the reflection of changes of lesson design on the lesson plan. That is to say, repeating update of models and the lesson plan helped him to clarify and externalize the design rationale of the lesson. Furthermore, the repetition also helped him reflect the change of lesson design on the lesson plan. Consequently, this can be a case supporting both hypotheses of this study as previously mentioned.

Note that the concordance rate once decreases in the second cycle. In the first cycle, the lesson plan was a rough note, and the teacher had a difficulty to organize his idea of the lesson. Therefore, the scenario model at this time mainly had I_L events representing only the goal of the lesson and concrete interactions between teachers and learners. This caused high concordance rate yet the design rationale was not clear. On the contrary, the concordance rate in the second cycle is lower than the first one. This is also the result of improvement because there is the increase in the number of I_L events. This means the teacher has enriched the lesson design in his mind. However, the concordance rate is low because the lesson plan has not reflected enough on the lesson design yet he has expressed it when making the I_L scenario model. Thus, it suggests that, in the cycle, the teacher could update lesson plan with improvement of lesson design through making an I_L event scenario model. Finally, he made the lesson plan reflecting results of improvement of lesson design.

Table 1 Improvement process of a lesson plan and an I_L scenario model

Cycle	1	2	3	4	5
# of items in the lesson plan	17	21	22	25	31
# of I_L events in the model	73	82	94	91	91
# of concordance of the items and the I_L events	56	57	77	78	88
The concordance rate (%)	76.7	69.5	81.9	85.7	96.7

4. Conclusion

We have discussed practical experiences of the field trials the authors carried out with teachers in ToChuSha. In the field trials, OMNIBUS worked as the basis for describing design rationale of lessons and SMARTIES worked as a tool for describing them as I_L scenario model. In this study, teachers firstly made lesson plans based on their idea and then discuss for improving it with the authors. In fact, the teacher could modify lesson design or make new ideas for a lesson plan in discussion using I_L scenario model after they made the lesson plan by themselves. The quality is ensured to a certain extent because some subject matter experts properly assessed it. This can be considered as the contributions of OMNIBUS and SMARTIES in this study. As the result, this study obtained case examples supporting the hypotheses mentioned in Section 2. Of course, there is still room for argument about the comparison of the proposed approach with the others and the learning effect of lessons designed with this approach.

Some doubt remains about this result. Interpretations of lesson plan by the authors might have some influence on the result because the authors made I_L scenario model from the lesson plan. However, as stated in Section 2, the teachers confirmed the model and then discussed with the authors to improve lesson design. It should be noted that the teachers and the authors discussed continuously to improve lesson design until they finish making lesson plan. This is because they can record design rationale of each lesson plan consistently as I_L scenario model. The teachers gave comments that the record is helpful to look back on thinking when they had described the lesson plan.

The future work is to improve SMARTIES with which teachers can easily make and improve lesson design by themselves. In this study, not teachers but the authors made I_L scenario models in view of our previous study [4]. It is necessary to make OMINBUS and SMARTIES user-friendly to allow teachers can use SMARTIES by themselves.

Acknowledgements

The authors are deeply grateful to ToChuSha, the community of social studies teachers in Tokyo. In particular the authors would like to thank Mr. Takuma Suzuki at the Toyotama junior high school, whose cooperation in this study was immensely valuable throughout the course of this study. This research was partly supported by *Grant-in-Aid for Scientific Research (A) 22240011 and Grant-in-Aid for Young Scientists (B) 22700811*.

References

[1] Donald A. Schon: *The reflective practitioner: How Professionals Think in Action.* New York, Basic Books, 1983.

[2] Karen Hammerness, et al.: How Teachers Learn and Develop, *Preparing Teachers for a changing World: What teachers Should Learn and Be Able to Do*, San Francisco, Jossey-Bass, pp. 358-389. 2005.

[3] Hayashi, Y., Bourdeau, J. and Mizoguchi, R.: "Using Ontological Engineering to Organize Learning/Instructional Theories and Build a Theory-Aware Authoring System", *International Journal of Artificial Intelligence in Education,* Vol. 19, No. 2, pp. 211-252, 2009.

[4] Hayashi, Y., Kasai, T., and Mizoguchi, R.: "Ontological Modeling for Reflective Instructional Design: A Case Study on Modeling a Lesson Plan", *Proc. of ICCE2010*, pp. 25-32, 2010.

[5] Stigler, J.W. and Hiebert, J., *Teaching gap*, The Free Press, NY, 1999.

[6] Mizoguchi, R. and Bourdeau, J.: Using Ontological Engineering to Overcome Common AI-ED Problems, *International Journal of Artificial Intelligence in Education*, 11(2), pp. 107-121, 2000.

[7] Devedzic, V.: *Semantic Web & Education,* Springer ScienceBusiness Media, 2006.

Experimental Study toward Estimation of a Learner Mental State from Processes of Solving Multiple Choice Problems Based on Eye Movements

Kazuaki Kojima*, Keiich Muramatsu & Tatsunori Matsui
Faculty of Human Sciences, Waseda University, Japan
*koj@aoni.waseda.jp

Abstract: Recently, the research area of intelligent educational systems has increasingly approached Educational Data Mining, which addressed a variety of learner aspects such as mental states. Although data of eye movements is promising in realizing mental states of learners, we have to empirically understand features of eye movements in advance due to difficulty in handling them. Toward the final goal of automatic estimation of learner mental states based on eyes, this study experimentally described changes of eye movements by confidence of correct answers. In our experiment, participants were asked to answer multiple-choice problems and respond to questionnaires about the problems, with the results indicating that transitions of eyes in initial processes where all choices were scanned differed depending on the levels of confidence.

Keywords: Educational data mining, mental state, confidence, eye movements, multiple-choice problems

Introduction

Recently, the research area of intelligent educational systems has increasingly approached Educational Data Mining (EDM), attempts to develop methods for exploring data from educational settings and adapt those methods to understand learners and their settings [1]. Such EDM studies have addressed a variety of learner aspects. Besides knowledge structures, those aspects include affective states such as confidence or confusion. Development of such EDM methods that automatically and directly estimate learner mental states would allow teachers to know situations of the learners in distant education, or systems to sophisticate their supportive interventions.

One important data generally used in EDM is information of eyes. Data from eyes is considered to be promising in realizing mental processes that can instantly change in a short activity (e.g., solving of a single problem). Accordingly, it is adopted in exploring human cognitive processes or mental states. In fact, some ITS studies have adopted techniques for measurement of eye movements (e.g., [2]). Furthermore, methods to automatically estimate cognitive processes or mental states from eyes have been implemented [3, 5]. Due to unstability in obtaining data from eyes and absence of a general method to analyze it, however, eye data is considerably difficult to handle in any ways. Hence, the automatic estimation basically needs specific models of eye transitions according to structures of tasks. To automatically estimate mental states in problem solving through observation of eyes, we have to understand features of eyes in advance based on data empirically collected.

Toward the final goal of automatic estimation of learner mental states, this study experimentally described changes of problem solving processes by confidence of correct answers as a mental state based on eye movements. To provide data description not depending on task domains, we adopted multiple-choice problems that only needed recall of knowledge. This study, therefore, aimed to produce qualitative descriptions useful in direct and bottom-up analysis of learner behaviors.

1. Method

In this experiment, participants were asked to answer multiple-choice problems and respond to questionnaires about the problems. Eye movements of the participants were recorded while they answered the problems. Confidence in each problem was evaluated based on the questionnaires. We then explored relationships between confidence and features of the eye movements.

1.1 Tasks and Procedures

This experiment used 30 four-choice problems that questioned encyclopedic knowledge such as history or geography. To observe behaviors in reading of texts and searching of answers, the problems that require no mental processing of integration and reasoning were created by the authors.

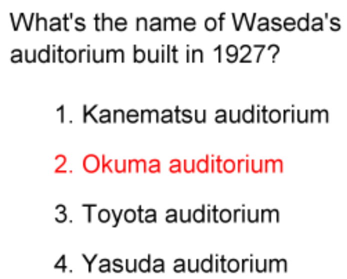


Figure 1. Example of problems presented by the program

Each of the problems was presented on a full screen of a PC monitor by a program implemented by the first author. Figure 1 shows examples of the problems presented. Participants responded each of the problems in the following procedures.

1. Reading a text: A button labeled "*proceed next*" appeared in a position where a text of each problem would be presented. When a participant clicked the button with a mouse, a text of a problem was shown. S/he was instructed to click a button labeled "*I have read the text*" beneath the text as soon as s/he had read it.
2. Responding a pre-questionnaire: The text was once hidden and a dialog window showing the questionnaire that asked to what extent the answer was familiar (*pre-familiarities*) appeared. The participants responded to it by selecting one of "I know its answer, I can answer without choices (*recall*)", "I can remember its answer from choices (*recognition*)", "I don't know, but I may be able to guess from choices (*guesstimate*)", and "I have no idea (*no-idea*)".
3. Selecting an answer: The text and four choices were presented after response of the questionnaire. Each of the choices turned red when the mouse cursor entered. The red choice was selected as the answer when it was clicked.
4. Responding post-questionnaires: The problem was then removed and questionnaires about the problem were presented. The first questionnaire asked whether or not the

participants selected the choice that they had judged to be the answer. The second again asked the familiarity for the answer (*post-familiarities*). In the post- questionnaires, choices in the questionnaire here were altered to “I know the answer, I could answer without the choices (*recall*)”, “I remembered the answer from the choice (*recognition*)”, “I don’t know, but guessed from the choices (*guesstimate*)”, or “I had no idea (*no-idea*)”. The third asked *evaluations* for each choice. Every choice was evaluated with “it’s definitely the answer/ not the answer”, “it’s probably the answer/ not the answer”, or “I cannot judge”.

Each participant in the experiment was seated in front of a desk where a PC monitor (the resolution was 1280 x 1024 pixels) was set up. S/he was asked to answer problems with a mouse as quickly and correctly as possible, and to respond to the questionnaires after answering of each problem. To train for the experimental tasks, s/he answered one problem and its questionnaires prior to answering the 30 problems.

While each participant engaged in selecting answers, eye movements were recorded with EMR-AT VOXER produced by nac Image Technology Inc. The sampling rate of the record was 60 frames/sec. Data in each frame included values of x and y coordinates on the screen.

1.2 Analysis

Data of the 30 problems of each participant was categorized into groups according to the pre- and post-questionnaires. The number of alternatives in finally selecting an answer of each problem was estimated with responses of the evaluations of the third one of the post-questionnaires: answering from *one* (three choices were evaluated not to be the answer), *two*, *three* or *all* (nothing was evaluated not to be the answer). The groups were formed by combining the alternatives, pre-familiarities and post-familiarities.

Data of eye movements were used to analyze transitions of positions where participants watched. Fixations of eye movements were computed from the data. In this study, a fixation was defined as a sequence of six or more serial frames (0.1 sec or more) whose range was smaller than a circle of a 22-pixel radius (the angle was 0.8 degree or smaller) according to the experimental setting. Targets that the participants watched and their transitions were estimated with vertical positions of fixations. Prior to the analysis, we removed frames of blinks from the data, and then corrected and smoothed it.

In answering a single problem, first of all, each participant must have read its text and all four choices in turns. We defined this initial process where the all choices were viewed as *initial scanning*, and extracted transition patterns of fixations in it.

2. Results

Ten undergraduate students participated in the experiment. Four of them were excluded from the analysis because they inappropriately responded to the questionnaires due to misunderstanding of instructions.

2.1 Groups of Problems

Data of 180 problems was obtained from the six participants. In the aspect of the alternatives, answers were selected from one alternative choice in 58.1% of the problems, from two in 16.8%, from three in 3.9% and from all in 21.2%. Hereafter, data[1] of

[1] The label “one/recall-recall” indicates a group of problems whose alternative was one, whose pre-familiarity was recall, and whose post-familiarity was recall.

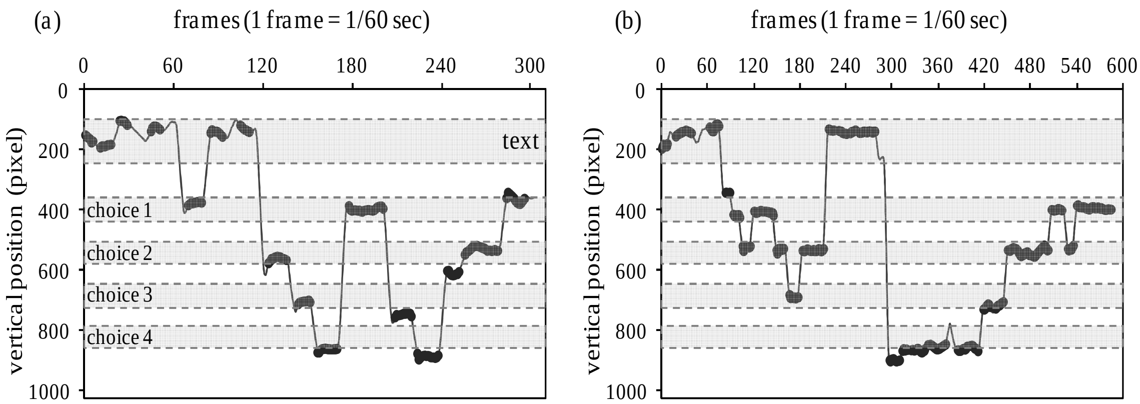
one/recall-recall, one/recognition-recognition, one/guesstimate-guesstimate, all/guesstimate-noidea and all/noidea-noidea was analyzed because data frequencies of the other groups were few.

It was assumed that the degrees of confidence of correct answers got higher in the order of the groups described above. Actually, the proportions of correct answers basically got higher in the order (97.3% in one/recall-recall, 100% in one/recognition-recognition, 46.7% in one/guesstimate-guesstimate, 27.3% in all/guesstimate-noidea and 20% in all/noidea-noidea).

2.2 Transactions of fixations

Figure 3 presents examples of eye movements obtained from the participants. Transitions of targets that the participants watched were reproduced with vertical positions of fixations. In the case (a) of the figure, the participant first read a problem text for about 1 second. He again saw the text after viewing choice 1, and then viewed choices 2, 3 and 4 in turns. He finally selected choice 1 as the answer after examining the four choices. Transitions in initial scanning in this case were [t (text) > 1 > t > 2 > 3 > 4]. Transition patterns in initial scanning were categorized into the following three.

1. *Simple sequential scanning (simple)*: Four choices were simply scanned in turns in initial scanning.
2. *Sequential scanning including returning (returning)*: Like as the case mentioned above, returning to a text was included during scanning of four choices in turns.
3. *The others*: Transitions of choices were not orderly but varied. An example of this is (b) in Figure 3 (The transitions were [t > 1 > 2 > 1 > 2 > 3 > 2 > t > 4]).



(Lines indicate a vertical position of eyes in each frame, and bolded lines that of fixation)
Figure 3. Examples of eye movements

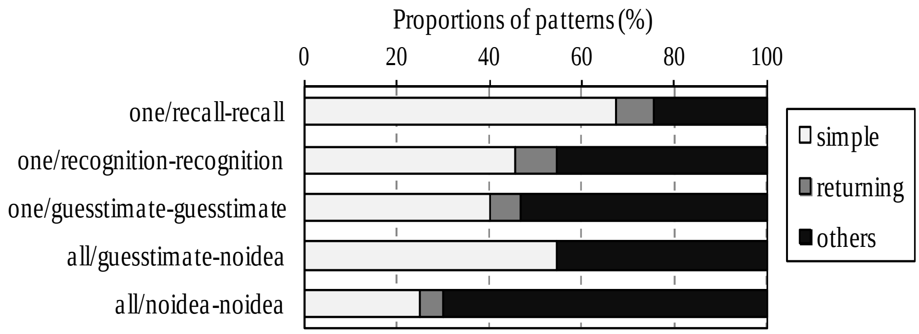


Figure 4. Proportions of transition patterns

Figure 4 indicates the proportions of each transition pattern in each group. We examined differences of the numbers of the patterns among the groups by the chi-square test, however, the differences were not significant ($\chi^2(8)=12.93$, *n.s.*). We examined differences between one/recall-recall and all/noidea-noidea, with the results indicating that the differences were significant ($\chi^2(2)=11.36$, $p<.01$). The residual analysis indicated that the number of simple was high and that of others were low ($p<.01$) in recall/one, and the number of simple was low and that of others was high ($p<.01$) in no-idea/three.

3. Discussion and Future Work

It was assumed that confidence of correct answers was the highest in working on the problems in one/recall-recall. In this case, most transition patterns of eyes in initial scanning were simple. This indicates that the participants scanned choices in turns when confidence was high. On the other hand, transitions of eyes varied or were disordered in all/noidea-noidea, where confidence was the lowest. Thus, analysis of initial scanning must be useful in directly estimating confidence in an earlier stage of problem solving.

Our experiment revealed that the transitions were differed depending on confidence. This fact must indicate that purposes of viewing choices were changed along with confidence. When confidence was high in answering to a problem, viewing choices was considered to be an action to search an answer a participant had already had, whereas it was an action to understand a problem or consider its answer when confidence was low. In many cases of low confidence, initial scanning included transitions to return to texts from choices. Eye transactions were disordered because of comparison between a text and a choice, or between choices. In text reading, it has been documented that backtracking of eye movements occurs when the reader does not understand the text [4].The backtracking in text reading is similar to the disordered eye transactions observed in the current study. Therefore, it is considered that initial scanning served as problem understanding when confidence was low. One important future work is further collection of empirical data to study more diverse cases, such as problems answered in middle confidence. Of course, it is also important task to implement computational models to estimate confidence based on transaction patterns of things that learners watch. Based on our finding, such a model may be implemented with simple techniques because it has only to distinguish a normative and orderly process in initial scanning.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 22300294.

References

[1] Baker, R. S. J. D. and Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1, 3-17.

[2] Conati, C. and Merten, C. (2007). Eye-Tracking for User Modeling in Exploratory Learning Environments: An Empirical Evaluation. *Knowledge-Based Systems*, 20, 557-574.

[3] Nakayama, M. & Takahasi, Y. (2008). Estimation of Certainty for Multiple Choice Tasks Using Features of Eye-movements, *Proceedings of Symposium on Eye Tracking Research & Applications* (pp. 73-76).

[4] Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124, 372-422.

[5] Salvucci, D. D. (2001). An Integrated Model of Eye Movements and Visual Encoding, *Journal of Cognitive Systems Research*, 1(4), 201-220.

Using CbKST for Learning Path Recommendation in Game-based Learning

Simone Kopeinik[a*], Alexander Nussbaumer[a], Michael Bedek[a] & Dietrich Albert[ab]
[a]*Knolwedge Management Institute, Graz University of Technology, Austria*
[b]*Department of Psychology, University of Graz, Austria*
*simone.kopeinik@tugraz.at

Abstract: This paper presents a novel approach how learning paths consisting of game units can be created and adapted to learners based on their behavior during the game play. Non-invasive assessment procedures interpret the behavior and calculate information about the competences of the learners. A user model holds probabilistic information on the competence profile. Based on this competence profile game units/stories are recommended fitting to the actual competence state of the learner. This approach is part of the EC-funded TARGET project which provides the technical infrastructure regarding the 3D virtual game environment. The innovative part of this paper is the adaptive learning strategy and how it can be included in a game-based environment. The user perspective is demonstrated on a concrete scenario where the learner has to solve a task in the game-based environment.

Keywords: Digital Educational Game, Adaptive Learning Paths, Competence-based Knowledge Space Theory, Simplified Updating Rule

Introduction

An important research area in Technology-enhanced learning (TEL) focuses on adaptivity and personalization. Several approaches have been elaborated that demonstrate how a system and its content can be adapted to the learner´s knowledge level. To allow individually tailored educational software solutions, it is necessary to keep track of an individual learner's knowledge state at a specific moment in time [6]. In Adaptive systems relevant information is typically described in user models, domain models, and adaptivity models [3].

One research area in TEL is Game-Based Learning (GBL) and Digital Educational Games (DEG). They provide powerful opportunities for the leaner regarding motivation and flow experience. It has also shown that these factors in game-based settings have positive influence on learning effectiveness and learning outcomes [9]. The European research project ELEKTRA (http://www.elektra-project.org/) firstly explored and presented the micro-adaptivity approach. This methodology allows assessing a learner non-invasively and continuously without interrupting the learner´s potential game flow experience. Assessment data is retrieved from the user's behavior while being engaged in the game [8]. This approach was revisited and implemented in subsequent projects. In 80Days (http://www.eightydays.eu/) for instance, information was derived from specific actions indicated by the manipulation of objects [9].

This paper presents an approach how learning paths consisting of game units can be created and further adapted to learners based on their behavior during the game play. Non-invasive assessment procedures interpret the learner´s observable behavior and infer information about the competence level, which is stored in the user model. Based on the user model, stories are recommended fitting to the actual competence state. This learning cycle is done until a learner achieves a desired competence state. This approach is part of the EC-funded TARGET project [10] which provides the technical infrastructure regarding the 3D game environment. The innovative part presented in this paper is the adaptive learning strategy (LS) and how it is embedded in a game-based environment.

1. Conceptual Framework

In a DEG like TARGET, learning happens during the game play. Therefore, the game is structured into game units that teach competences and thus act as learning objects (LOs). Traditionally, LOs are designed as multimedia documents containing texts, images, animations, and other 2D elements. In our case these units are designed as immersive 3D environments where learners can move around and interact. Since there are also tasks to solve in a defined and contextualized situation, they are seen as stories. For the completion of a story, a certain level of proficiency is required but might also enhance throughout the confrontation with comprised challenges. Hence, stories or LOs do not only teach subjects, but also test the knowledge of the learners by observing the learners' performance when interacting in a story (see Figure 1).

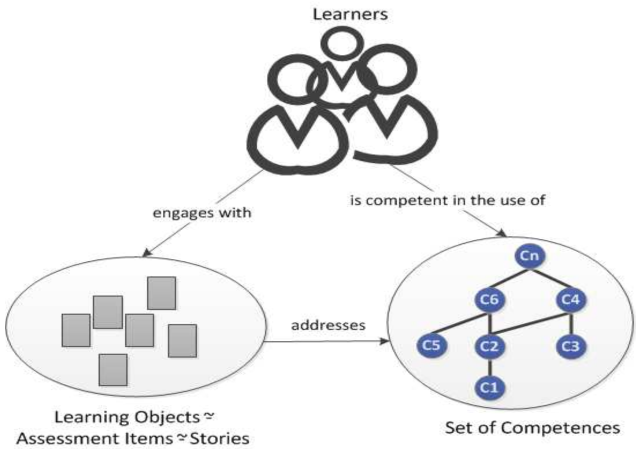


Figure 1: Relation between stories, competences and learners

In order to formally structure the stories with respect to knowledge and competences, the conceptual framework is based on the Competence-based Knowledge Space Theory (CbKST) [5][7]. This framework allows for representing knowledge of knowledge domains and learners and provides algorithms to test the knowledge of learners in terms of competences. The learning path is adapted according to these competences. The basic idea is to define a domain model by defining competences and to build a structure on them (a competence can be a prerequisite for another one). These competences are assigned to learners (a learner can demonstrate a competence), to learning objects (a learning object teaches competences), and to assessment items (an item can test whether a learner can demonstrate a particular competence). Each story is assigned with the competences that are required for solving the tasks of that story and hence for story completion. Competences and their relationships to stories form the domain model. Competences may or may not concern with one another but if they do, this is displayed in the domain model using pre-requisite relationships. In addition to the domain model there is also a user model that describes a learner's individual progress and state in terms of obtained competences. Figure 1 outlines the relationships between stories, competences, and learners.

Figure 2 illustrates the learning cycle: In the initialization phase the learner is invited to set learning objectives in terms of competences (Target Competence Profile, TCP) and to provide pre-knowledge in terms of competences through self-assessment (User Competence Profile, UCP). Based on the competence profile a story is selected that addresses the competences that the learner should learn next. The learner plays this story by interacting with the game and by trying to solve the task in the given situation. The interactions are observed and used to identify if the learner shows the respective competences. The result of this non-invasive assessment goes into and updates the user model. When the learner has finished a story, the system recommends the next story based on the user profile taking into account the results of stories a learner previously engaged with.

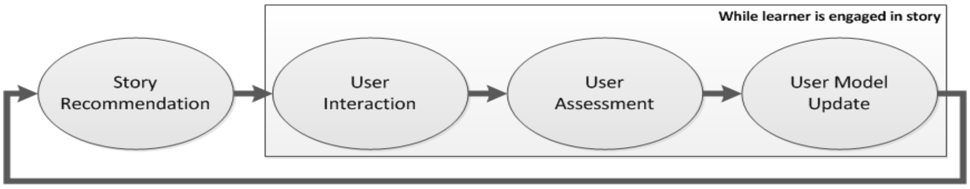


Figure 2: Learning Cycle

2. Adaptation Model

Adaptation in our case means to adapt the story path to the learner's competence state. Stories are recommended to the learner according to currently shown competences.

Assessment items are usually provided to the learner before or after she has consumed a set of LOs. However, as it often happens in serious games, both teaching and testing are within a single learning unit or game scenario. This offers the great opportunity to assess the learner’s ability while she or he is being engaged with the game. It allows assessing the learner without destroying a potential flow experience [4]. Therefore, a non-invasive or implicit assessment procedure was introduced. The approach is based on the non-invasive assessment procedure that is already implemented in the TARGET project. Basically, it grounds on the interpretation of the learner´s actions and interactions within the virtual environment [2]. These observations result in values (ranging between 0 and 1) for the set of competences assigned to the current story. For example, if a learner is playing story A and the competences x, y, and z are assigned to this story, then the result could be [0.1, 0.7, 0.8] meaning that the learner performed well in respect to competences y and z and poorly as to competence x. We call these values competence performance values.

In Figure 3 one can find a more detailed view on the structure of the learning strategy’s (LS) logic. Starting at the top of the illustration, the Domain Model encompasses all identified competences of the domain and their (pre-requisite) relations to each other. This model does not change during the learning cycle. On the other hand, the user model located in the center of the illustration keeps track of the competences a learner demonstrates. It is initialized with the values of the user competence profile, the target competence profile and relevant parts of the competence domain. All competences of the TCP and those that are pre-requisite relations to them are relevant for the user model.

Each competence within scope has assigned a probabilistic value that indicates the probability of the learner being competent in the use of it. The Assessment part receives competence performance values in a continuous range from 0 to 1 for single competences. Incoming values are applied through an algorithm called Simplified Updating Rule [1]. As the algorithm can only handle binary updates, solely values smaller than 0.35 for negative assumptions and values higher than 0.65 for positive assumption are taken into account. After the classification of the input, the algorithm can be applied on the affected competence and its related competences in the user model. For example, if a competence x is a prerequisite for competence y then we can assume that a learner that shows competence y also shows competence x. If the assessment procedure delivers a probability value for competence x, we can also make the assumption that the same learner demonstrates competence y to a certain extent and thus we can increase also the probability value for competence y. According to this consideration all related probability values are modified each time the assessment procedure delivers data for the competences assigned to a story.

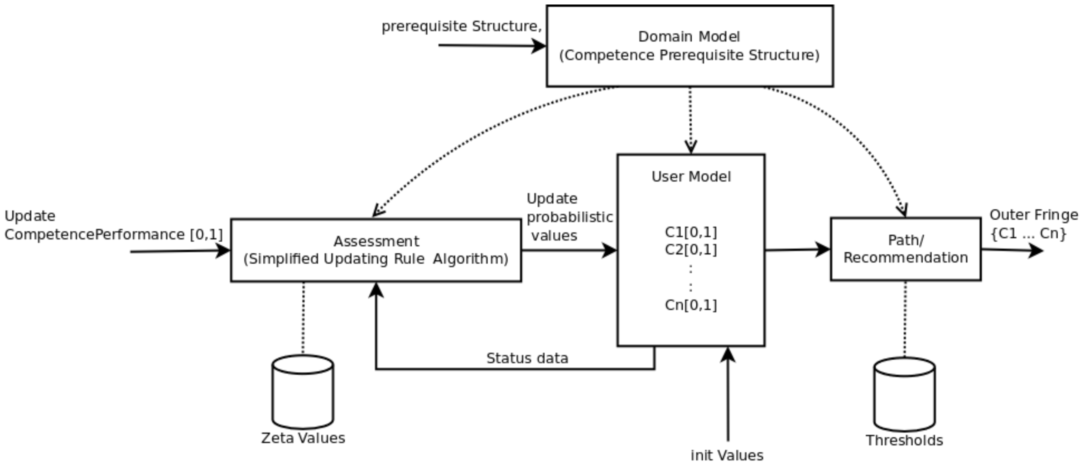


Figure 3: The Core Logic of the Learning Strategy Component

The recommendation strategy is based on the competence profile (of the user model). This strategy is done in a two-step process: First the competences that the learners should obtain next are determined. Then, an according story is selected. To that end, the Path/Recommendation module accesses the user model and selects a small set of competences whose probabilistic values differ from a defined threshold the least. If the value of a competence is very high, it is likely that the learner already demonstrates this competence. So this competence will not be selected. If the value is very low, it is likely that the learner is not competent in this area yet. This competence is also not selected, because it is assumed to be too difficult for the learner at this stage of the learning process. Therefore, a competence should be selected that has a probability value of about 0.5, because such a competence is expected to be of medium difficulty for the learner. In the second step a story is selected that addresses the picked competences. Then the learner continues with this new story and a new assessment is happening. This cycle is conducted until all competence values are above a certain threshold value.

3. The Learner's Perspective

The learner's active part in the personalization process takes place during the initialization phase, when the learning plan is created. In the TARGET project, a tool called Competence Analyzer is provided as an input device to assign selected competences to the User Competence Profile (UCP) or the Target Competence Profile (TCP). When the learner finishes, the UCP should include all competences the learner demonstrates at this point in time. The TCP should include all competences the learner would like to achieve during the execution of the resulting learning plan. Based on these profiles the first story recommendation can be provided and presented to the learner. Within the virtual environment the learner is represented as an avatar and has to interact (non-verbally and verbally) with so called non-playable characters (NPCs) to master story-dependent tasks. Stories are tailored to contribute to a learner's competence development. The story description encompassing tasks, characters, and background information is presented to the learner at the very beginning. After reading the initial story manual the learner enters a scene of the Game scenario. The player learns and is being assessed till the end of the game is reached, which happens either when the story tasks have been mastered successfully or the playing time has expired. In any case, the learner gets the chance to reflect on a diagram that presents her or his performance of the story competences throughout the last game play. A next story is offered to the learner.

4. Conclusion and Outlook

Focus of this paper lies on the CbKST based modeling of user competences to support the adaptive guidance through competence based learning and assessment in a DEG. A brief insight into the implementation and the application of the algorithm was provided. Evaluations of the overall TARGET platform have started. Initial feedback from users indicates that recommended stories are experienced slightly above the medium difficulty level. Further studies will provide information about the appropriateness of selected competences addressed in these stories. Subjects of adjustment could be the thresholds related to competence probabilistic that lead to the competence selection and also the number of competences addressed by one story, in order to improve the adaptation.

Acknowledgements

This paper is part of the EC-Project TARGET funded by the 7th Framework Program of the European Commission. The authors are solely responsible for the content of this paper. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein.

References

[1] Augustin, T., Hockemeyer, C, Kickmeier–Rust, M. D., Podbregar, P., & Albert, D. (submitted). The simplified updating rule. *Journal of Computational Science*.
[2] Bedek, M.A., Petersen, S. A., Heikura, T.(2011). From Behavioral Indicators to Contextualized Competence Assessment. *In: Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies, pp. 274 – 276.*
[3] Brusilovsky, P., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. (P. Brusilovsky, A. Kobsa, & W. Nejdl, Eds.)*LNCS: The Adaptive Web, 4321, 3 - 53. Springer. doi:10.1007/978-3-540-72079-9_1*
[4] Csikszentmihalyi, M.: Flow: The psychology of optimal experience. *Harper & Collins, New York (1990)*
[5] Falmagne, J.-C., & Doignon, J.-P. (Eds.). (2011). Learning Spaces. *Springer. doi:10.1007/978-3-642-01039-2*
[6] Fischer G. (2001). User Modeling in Human-Computer Interaction. *User Modeling and User-Adapted Interaction, Kluwer Academic Publishers, 2001, 11, 65-86.*
[7] Heller, J., Steiner, C., Hockemeyer, C., & Albert, D. (2006). Competence-Based Knowledge Structures for Personalised Learning. *International Journal on E-Learning, 5(1), 75-88.*
[8] Kickmeier-Rust, M. D., Hockemeyer, C., Albert, D., & Augustin, T. (2008). Micro adaptive, non-invassive assessment in educational games. In M. Eisenberg, Kinshuk, M. Chang, & R. McGreal (Eds.), *Proceedings of the second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning (pp. 135-137), November 17-19, 2008, Banff, Canada.*
[9] Steiner, C. M., Kickmeier-Rust, M. D., Mattheiss, E., Göbel, S, Albert, D. (2012). Balancing on a High Wire: Adaptivity, a Key Factor of Future Learning Games. *In: An Alien's Guide to Multi-Adaptive Educational Computer Games. Informing Science Press, Santa Rosa, USA*
[10] TARGET project, http://www.reachyourtarget.org/, retrieved at 2012-05-28

A Framework for Analyzing Course Contents in Learning Management Systems with Respect to Learning Styles

Moushir M. EL-BISHOUTY*[a,b], Ting-Wen CHANG[a], KINSHUK[a] & Sabine GRAF[a]
[a]*Athabasca University, Canada*
[b]*City for Scientific Research & Technological Applications, Egypt*
[*]moushir.elbishouty@athabascau.ca

Abstract: The harmony between the learning styles that a course supports and the actual learning styles of learners helps to magnify the efficiency of the learning process. The aim of this research is to present a framework for analyzing existing course contents in learning management systems and providing the teacher with information regarding how well the course supports different learning styles of students based on the Felder and Silverman's learning style model. This allows the teachers to be aware of the suitability of their courses for students with different learning styles and to improve their courses to support more learners. This paper introduces the design and the implementation of the framework.

Keywords: Course analyzer, learning management systems, learning styles

Introduction

In online learning, teachers build courses according to their teaching methods. Teaching methods vary. Some instructors lecture, others demonstrate or discuss; some focus on principles and others on applications; some emphasize memory and others understanding [1]. On the other hand, learners have different backgrounds, motivation and preferences in their own learning processes and web-based systems that ignore these differences have difficulty in meeting learners' needs effectively [2]. Therefore, when designing instructional material, it is important to accommodate elements that reflect individual differences in learning. One of these elements is learning styles. Understanding a student's particular learning style and how to best meet the needs of that learning style is essential to perform better. Clay and Orwig [3] defined learning style as a unique collection of individual skills and preferences that affects how a person perceives, gathers and processes information. Learning styles affect how a person learns, including also the aspects of how a person acts in a learning group, participates in learning activities, relates to others, and solves problems. Basically, a person's learning style is the method that best allows the person to gather and to understand knowledge in a specific manner. Once a learner's particular learning style is identified, it will be possible to identify ways in which the learning process can be improved [4]. There are many models about learning styles in literature such as Kolb [5], Dunn & Dunn [6], Honey & Mumford [7], and Myers-Briggs [8]. This research paper utilizes the Felder and Silverman's Learning Style Model (FSLSM) [1] because of its applicability to e-learning and compatibility to the principles of interactive learning systems design [9]. In this model, Felder and Silverman proposed four dimensions of learning styles (active/reflective, sensing/intuitive, visual/verbal, and sequential/global) and teaching styles (active/passive, concrete/abstract, visual/verbal, and sequential/global), where each teaching style corresponds to (matches with) a learning style.

Many researchers have conducted research to detect the learners' learning styles and provide recommendations and adaptations for online courses based on learning styles. For example, Paredes & Rodríguez [10] presented a framework that collects explicit information about the students by means of the Index of Learning Styles (ILS) questionnaire developed by Felder and Soloman [11], adapts the course structure and sequencing to the student's profile and uses the implicit information gathered by the system during the course in order to dynamically modify the course structure and sequencing. Graf & Kinshuk [12] introduced a concept for enhancing learning management systems (LMSs) with adaptivity based on learning styles. They used the open source LMS Moodle as a prototype and developed an add-on that enables Moodle to automatically provide adaptive courses that fit to the learning styles of students. Mejía et al. [13] proposed an approach of an adaptation process that allows adjusting different types of resources to the user's preferences by means of the identification of the user's learning style in LMSs.

Our research is different from the previous research in that we focus on analyzing existing online courses and making teachers aware of how well those courses fit with diverse learning styles. Currently, LMSs contain tons of existing courses but very little attention is paid to how well these courses actually support learners. This research presents a framework to analyze existing course contents in learning management systems and provide the teachers with a visualization tool, which makes the teachers aware of the suitability of their courses for diverse learning styles. The aim is to help the teachers in improving their courses to support students with different learning styles. In the next section, the mechanism for analyzing course contents is presented; the framework architecture is illustrated in section 2. In section 3, the implementation of the framework is introduced, followed by the conclusions and the future plans of the research.

1. Course Analyzing Mechanism

Making teachers aware of how well their courses fit with diverse learning styles can help them in improving their courses to support students with different learning styles. Therefore, we propose a mechanism for analyzing existing courses in LMSs in order to infer which learning styles they currently support. The mechanism currently considers eleven types of learning objects (LOs), as listed below; however, from technical point of view, new types of LOs can easily be included in this mechanism, if required.

- *Commentaries:* provide learners with a brief overview of the section.
- *Content Objects:* are used to present the learning material.
- *Reflection Quizzes*: include one or more open-ended questions about the content.
- *Self-Assessment Tests*: include several close-ended questions about the content.
- *Discussion Forum Activities:* provide learners with the possibility to ask questions and discuss topics with their peers and the instructor.
- *Additional Reading Materials:* provide learners with additional sources for reading about the content.
- *Animations:* demonstrate the concepts of the course in an animated multimedia format.
- *Exercises:* provide learners with an area where they can practice the learned knowledge.
- *Examples:* illustrate the theoretical concepts in a concrete way.
- *Real-Life Applications:* demonstrate how the learned material can be related to and applied in real-life situations.
- *Conclusions*: summarize the learned content.

In order to use the proposed mechanism to analyze an existing course, this course only needs to fulfill two requirements. First, for the general structure of a course, it is assumed that a course consists of several units and a unit can (but does not have to) consist of several sections. Second, it is assumed that each section (or unit if no section exists) contains at least one content object which presents the learning material of this section.

A section normally starts with a commentary. Subsequently, there is an *area before content* (ABC) that may include a few LOs that aim at motivating the learners and making the section interesting for them. After this area, the content is presented. In the next area, namely *area after content* (AAC), different types of LOs may be presented. The conclusions of the section can exist either right after the last content object or at the end of the section. This mechanism recognizes how well a section of an existing course fits to each of the eight poles of FSLSM by calculating the average of three factors: the availability, the frequency and the sequence of the learning objects in that section, as illustrated below. Consequently, the results can be summarized for each unit and for the whole course.

Certain LO types can support diverse learning styles. The availability of types of LOs is considered as a factor to infer the learning styles that a section of the course fits well. It measures the existence of LO types in the section that can support each learning style in respect to all types of LOs that can support that learning style. The availability factor (Ava_{ls}) of a certain learning style (*ls*) is calculated using formula 1. On the other hand, the frequency factor ($Freq_{ls}$) measures the existence of LOs in the section that support each learning style with respect to the total number of LOs in that section, and is calculated as described in formula 2. The obtained values for both, the availability factor and the frequency factor, range from 0 to 1, where 1 indicates a strong suitability for the learning style and 0 means no support.

$$Ava_{ls} = \frac{(\#\ of\ existing\ LO\ types\ that\ support\ ls\ )}{(\#\ of\ LO\ types\ that\ support\ ls)} \qquad (1)$$

$$Freq_{ls} = \frac{(\#\ of\ existing\ LOs\ that\ support\ ls\ )}{(\#\ of\ existing\ LOs\ )} \qquad (2)$$

Not only the types but also the order and the position of the LOs affect the suitability of a course regarding different learning styles. The sequence factor measures the suitability of the sequence of LOs for different learning styles. It is calculated for each LO according to its type, location (ABC or AAC) and order. It is determined according to how much this object type in that place fits with each of the eight learning styles of FSLSM. The sequence factor ($\mathrm{Seq}_{\mathrm{ls}}$) is calculated using formula 3. In this formula, fls (LO) =1, if the LO is suitable for that learning style at that location, and fls (LO) =0 otherwise. The weight w represents how well the position of a learning object in AAC/ABC fits to the learning style, and n is the number of LOs in the section. This formula represented the weighted mean of fls(LO). Its value ranges from 0 to 1, where 1 indicates a strong suitability for the learning style and 0 means no support.

$$Seq_{ls} = \frac{\sum_{i=1}^{n} f_{ls}(LO_i) \times w_i}{\sum_{i=1}^{n} w_i} \quad ,\ 0 < w \leq 1 \qquad (3)$$

2. Framework Architecture

The proposed framework is based on DeLeS tool [14]. DeLeS stands for “Detecting Learning Styles”. It is developed to detect students’ learning styles, in an automatic way, by using data from students’ behavior in a course in a LMS. DeLeS is a standalone tool; it is developed to be used for any LMS. The tool consists of two components, the data extraction component and the calculation component. The data extraction component is responsible for extracting the relevant data from the LMS database in order to calculate learning styles with respect to the four dimensions of FSLSM. Since one of the main aims of the tool is to be applicable for LMSs in general rather than only for one specific system, heterogeneity of database schemata in different LMSs is considered. The calculation component is responsible for calculating students' learning styles.

The proposed framework, for identifying how well a course fits to students' learning styles, extends the data extraction component to retrieve the existing course structure from the LMS database, and adds a new component, the course analyzer, that implements the mechanism described in the previous section. As shown in Figure 1, the data extraction component connects to the LMS database, keeps the tool generic for any LMS schema, and feeds the course analyzer with the raw data to apply the proposed mechanism. In addition, an internal database is implemented in order to facilitate a reliable way of storing and exchanging the data among the different components. Besides, the processing and visualization component retrieves the available information and supplies the teachers with a visualization tool through an interactive graphical user interface (GUI) in a way that supports the teachers to be aware of the suitability of their courses regarding different students' learning styles.

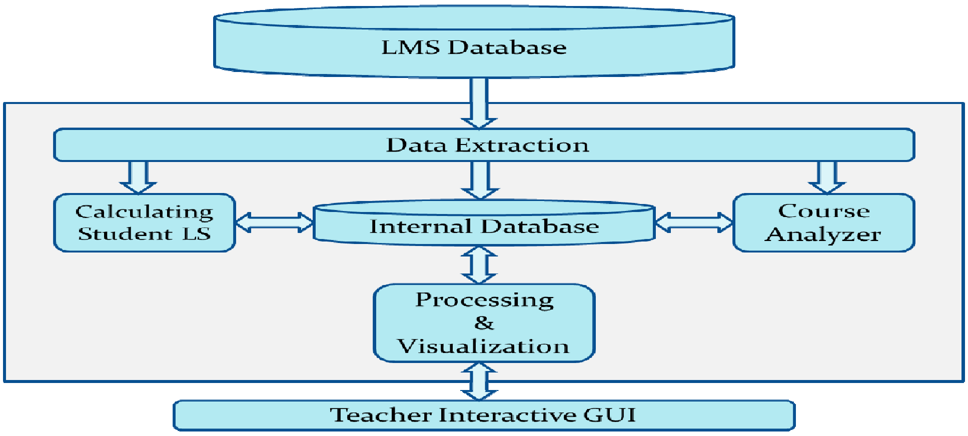


Figure 1. Framework Architecture

3. Framework Implementation

The framework is implemented as a client-server application. It is mainly developed using MySQL relational database management system and PHP scripting language. It is a stand-alone application that runs on the server side, connects to a LMS database (e.g., Moodle), extracts data, analyzes them, and provides the user (teacher) with an interactive graphical user interface that visualizes information about the suitability of the courses for students' learning styles. It allows the teacher to select one course from a list of existing courses in the LMS. Once a course is selected, a chart consisting of a set of bars is displayed. This chart shows how well the course and its sections fit with diverse learning styles.

4. Conclusions and Future Plans of the Research

This paper introduces the design and implementation of a framework for analyzing existing course contents in learning management systems and providing the teachers with information regarding the suitability of their courses for students' learning styles based on Felder and Silverman's learning styles model. The framework utilizes the proposed mechanism for recognizing how well a course supports different learning styles by calculating three factors: the availability, the frequency and the sequence of learning objects in that course. It provides teachers with an interactive graphical user interface, which can be used to analyze each section of the course and to make teachers aware of the quality of their courses as well as how to improve the courses to support learners with different learning styles.

The future plans of the research includes extending the proposed mechanism to additionally provide teachers with recommendations on how to best extend their courses to support more students with different learning styles, and to fit the course to the current cohort of learners. Also, experiments with teachers are planned to evaluate the efficiency and user-friendliness of the framework.

Acknowledgements

The authors acknowledge the support of NSERC, iCORE, Xerox, and the research-related gift funding by Mr. A. Markin.

References

[1] Felder, R. and Silverman, L. (1988). Learning and Teaching Styles. *Journal of Engineering Education* 94(1), 674–681.
[2] Atman, N., Inceoğlu, M.M. and Aslan, B.G. (2009). Learning Styles Diagnosis Based On Learner Behaviors in Web Based Learning. In: Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y., Gavrilova, M.L. (eds.), *Proceedings of the International Conference on Computational Science and Its Applications* – ICCSA 2009, 5593/2009, 900–909. Springer, Heidelberg.
[3] Clay, J. and Orwig, C.J (1999). Your learning style and language learning. *Lingual links library, Summer Institute of Linguistic*, Inc (SIL), International version 3.5.
[4] Onyejegbu, L. N. and Asor, V. E. (2011). An Efficient MODEL for Detecting LEARNING STYLE Preferences in a Personalized E-Learning Management System. *Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Software Engineering (JSSE)*, May Edition.
[5] Kolb, A.Y. and Kolb, D.A. (2005). The Kolb Learning Style Inventory. Version 3.1, *Technical Specification*, Hay Group, Boston.
[6] Dunn, R., Dunn, K. and Price, G.E. (1996). Learning Style Inventory. Lawrence, KS. Price Systems.
[7] Honey, P. and Mumford, A. (2006). The Learning Styles Helper's Guide. Peter Honey Publications Ltd., Maidenhead.
[8] Myers, I.B. and McCaulley, M.H. (1998). Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator. Consulting Psychologists Press, Palo Alto.
[9] Kuljis, J. and Liu, F. (2005). A Comparison of Learning Style Theories on the Suitability for Elearning. In: Hamza, M.H. (ed.) *Proceedings of the IASTED Conference on Web Technologies, Applications, and Services*, 191–197. ACTA Press.
[10] Paredes, P. and Rodríguez, P. (2004). A mixed approach to modelling learning styles in adaptive educational hypermedia. *Advanced Technology for Learning*, 1(4), 210–215.
[11] Felder, R.M. and Soloman, B.A. (2012). Index of Learning Styles Questionnaire. North Carolina State University, online at: http://www.engr.ncsu.edu/learningstyles/ilsweb.html.
[12] Graf S. and Kinshuk (2007). Providing Adaptive Courses in Learning Management Systems with Respect to Learning Styles, in: Richards G., ed., *Proceedings of the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (e-Learn), 2576-2583, AACE Press, Chesapeake, VA.
[13] Mejía, C., Baldiris, S., Gómez, S. and Fabregat, R. (2008). Adaptation process to deliver content based on user learning styles. In Gómez Chova R., Martí Belenguer R. & Candel Torres I. (Eds.), *International Conference of Education, Research and Innovation* (ICERI 2008), 5091-5100, IATED, Madrid.
[14] Graf, S., Kinshuk, and Liu, T.-C. (2009). Supporting teachers in identifying students' learning styles in learning management systems: An automatic student modelling approach. *Educational Technology & Society*, 12 (4), 3–14.

Negotiated Learner Models for Today

Susan BULL[a*], Ravi VATRAPU[b]
[a] *University of Birmingham, UK*
[b] *Copenhagen Business School, Denmark*
*s.bull@bham.ac.uk

Abstract: Negotiated learner models can help improve the accuracy of a learner model as well as promote metacognition. This paper defines directions for negotiating learner models based on multiple data sources, as the range of technologies and interaction types increases.

Keywords: Open learner models, negotiated learner models, technology-rich classrooms

Introduction

Learner Models (LM) offer personalisation, and are a core part of intelligent systems [1],[2]. Open LMs (OLM) give access to the LM in a human-understandable way, e.g. skill meters [3],[4], concept maps [5],[6] and, recently, treemaps [7],[8]. The OLM therefore not only allows personalisation, but also offers itself directly to users to prompt metacognition (e.g. reflection, planning) [9]. Control over the LM may rest with learner or system, or there may be mixed control [10]. We here focus on LMs that can be negotiated. This aims for an agreed LM by allowing the same argumentation and justification moves to both negotiation parties. Likely advantages are increased LM accuracy (taking into account the user's view of their skills), while promoting reflection during negotiation [11],[12]; and can form the basis of learner assessment [13]. These are strong educational reasons for negotiated LMs, but classroom needs have developed rapidly: there are now many activities and technologies in use, e.g. e-portfolios, blogs, wikis, social networks [14]. There is a call to integrate OLMs with e-learning approaches, such as e-portfolios [15], or more broadly in e-learning with the OLM at the centre [16]. Indeed, OLMs built on a range of sources are being investigated [17],[18]. Our aim is to unite current e-learning practices with benefits of negotiated LMs.

1. Negotiated Learner Models

LMs are usually negotiated by a student and tutoring system. However, other stakeholders can be involved, and the notion of system can include a range of technologies. We consider (i) fully negotiated LMs; (ii) partially negotiated LMs; (iii) other types of LM discussion.

(i) Mr Collins aims to increase LM accuracy by user-system discussion of the LM, while also promoting learner reflection through discussion [11]. The LM has separate belief measures: system inferences about user knowledge, and the user confidence in their skills (input with responses to questions). Mr Collins uses menu-based discussion to allow users to challenge and respond to the system at any time, and it initiates discussion if there are discrepancies between its inferences and the user's stated confidence in their knowledge. This follows the notion of interaction symmetry (system and student have identical negotiation moves) [19]. These include initiating, maintaining, ending discussion; and allow each party to request explanations, challenge beliefs, justify viewpoints, amend beliefs, accept compromise, maintain beliefs (if student and system do not agree, both inconsistent beliefs are retained). Adult users challenged Mr Collins if they disagreed with their LM, and suggested changes. STyLE-OLM [12] uses a dialogue game based model in negotiation, with the following dialogue moves (adapted from [20]): inform, inquire, challenge, disagree, justify, agree, suggest, skip. Initial findings gave additional support for the potential to promote reflection in university users. Based on the negotiation options of Mr Collins [11], CALMsystem uses a chatbot in negotiation [21]. Evaluation with 10-11 year olds showed significant improvements in self-assessment and reduction of LM discrepancies.

(ii) Close to negotiated LMs is xOLM [22]. Based on Toulmin's argumentation model [23], xOLM uses: data (actual belief); claims (summary belief - level I, II); warrants (evidence for beliefs); backings (qualitative/quantitative attributes supporting warrants). However, xOLM relies on users to initiate discussion. For example, students can challenge claims, warrants, backings; and receive justifications from xOLM. They can agree, disagree or move on (without resolution). New evidence is added to the LM, which can then be explored by the user. xOLM allows user challenge to succeed if there is unresolved disagreement [22]. In contrast, EI-OSM defers the decision to the (human) teacher if student-teacher interaction cannot resolve discrepancies using the system's evidence-based argument approach [24], also based on Toulmin. In addition to data, claims, warrants and backings, EI-OSM uses rebuttal and rebuttal data. Teachers had mixed reactions to considering assessment claims from students without evidence, but they believed these could be a useful starting point for formative dialogue [24].

Also relevant here are persuadable OLMs. The main difference between these and fully negotiated LMs are that (as with [22],[24]) models that can be persuaded do not offer each partner the same moves, or matching roles in diagnosis. A system has to agree before changes can be effected in the LM, and this occurs if a challenge comes from the student. For example, EER-Tutor has a component to allow users to challenge LM concepts [4]. A student can initiate a dialogue at any time. The system offers a question, and the LM is updated accordingly. Flexi-OLM also allows users to challenge the LM [5]. It gives evidence in the form of responses that led to its inferences, to provide a reflection resource. Students can try to persuade after seeing evidence, and similar to EER-Tutor, can demonstrate their knowledge by answering more questions.

(iii) While not negotiated, OLMlets [3] was used with Facebook for university users to discuss their LMs [25], indicating willingness to critically consider understanding in an open-ended way. This is crucial for model negotiation between humans. Another case where the LM is not negotiated is children giving self-assessments if they disagree, quantitative and text, for the teacher. This can become a focus for teacher-child discussion [26].

Research on student-system collaborative assessment found university participants acting as (human) teacher-student pairs would challenge an assessment and resolve disagreements in a manner resembling negotiated LMs [27]. This also supports the idea of student-teacher LM negotiation. Later work on negotiation between assessee (student) and assessor (system/teacher) [13] raised issues such as: assessment criteria; reasons for criteria; extent that student can challenge criteria; evidence to collect in interaction; sources of material to consult; negotiation ground rules; how to choose/communicate ground rules; extent student can influence negotiation outcome; learning during negotiated assessment [13].

2. Challenges for Negotiated Learner Modelling in Today's Classrooms

A challenge is to help teachers make effective use of information about students for classroom orchestration or offline consideration. OLMs are taking up the challenge of collecting data from multiple sources and presenting it in a meaningful way [17],[18], where the resulting models can allow technologies, learners or teachers to adapt learning/teaching, with technology or face to face. Our aim is to enhance such approaches by incorporating the benefits of negotiated LMs. As identified, main themes are: identical negotiation moves (fully negotiated); evidence for LM data; objects/artifacts of discussion/for consultation; ability to challenge LM; learning during negotiation; control over negotiated LM. By definition, negotiation implies identical argument or dialogue moves and rights. However, as seen, there are also reasons to use partially negotiated LMs. We therefore include all here.

For negotiation to have meaning, there must be *evidence* to support arguments. This could be based on more complex reasoning, e.g. using Toulmin's argument structure (data, claims, warrants, backings, rebuttal, rebuttal data) [22],[24]. However, while a detailed, more formal approach, this would be less flexible in today's classrooms. Alternatively, evidence could come from work produced by a learner, system explanation of its inferences and their sources, a student claim to have completed an activity. The latter may not be acceptable to teachers in formal assessment, but it may be a beneficial focus for student-teacher discussion [24].

Much evidence will point to *objects or artifacts*, e.g. essay (with simple statistical information), teacher appraisal, quiz outcome, spreadsheet calculation showing a skill, avatar activity log, wiki, blog or discussion entries (maybe with peer appraisal). It will be necessary to provide evidence at the appropriate level of granularity. For example, if a learner *challenges* inferences from detail of an online science experiment (e.g. choice of experimentation method), the system or teacher will need to present evidence accordingly. The artifact might be an activity trace combined with inferences drawn from the trace. If the student challenges the model at a broader level (general competence in scientific experimentation), the system might initially present an overview of sources of data. Challenge is also intended to promote reflection and encourage metacognition [10]. Therefore a user may *learn* while negotiating. The LM must update accordingly, raising the issue of how updates are represented. If negotiation around data from a specific application results in learning related to that application, a narrowly focused representation could be entered. If it could also apply in other contexts, it becomes a question of how broadly to apply new learning. It is also crucial to consider which party has *control* over negotiation outcome. This may be system [4]; student [22]; separate equally valid representations for each party [11]; or teacher (in student-teacher negotiation) [24]. In the following section we offer an example.

3. Combining Negotiated Learner Modelling with Current E-Learning Approaches

The Next-TELL OLM may use various data sources: self, peer and teacher input, alongside automated data from various activities, applications, and software (e.g. Moodle quizzes, Google docs, spreadsheets, social networks, OpenSim, e-portfolios) that may provide different granularity and levels of access [18],[28]. It is an “independent OLM” [10]; there is no system teaching - the OLM gives responsibility for learning to the learner, or help for the teacher. Using independent OLMs with e-portfolios is suggested as a useful way to combine two learner-centred information sources to inform adaptive training systems [15]. This could be relevant to Next-TELL in the future, but our focus is on the reverse: multiple sources of information to the LM, in line with the aim of harnessing current practices in technology-rich classrooms, where various sources can contribute to the LM [16],[17],[18]. Negotiating the LM may be even more important in such contexts, as students can lose track of activities contributing to their LM; may not appreciate the relative weighting of activities in the LM (recency/type of data); or may not realise that so many sources contribute. The opportunity to negotiate the LM in discussion with a teacher aims to help them recognise this. This may result in agreement with the representations, or provide information to help them form an argument. OLM visualisations and related activity data and evidence (e.g. e-portfolio contents) form the objects or artifacts of discussion. In negotiation, the teacher will need to consider assessment criteria, evidence, materials, extent of student influence over the LM, take account of learning during negotiation, and ground rules for negotiation (identified by [13]). A student challenge to the LM may occur, for instance, if they believe certain activities were not taken into account in the LM. They may use the negotiation tool (Figure 1 left), the outcome of which is sent to the teacher; or work face to face, with the teacher inputting the result of negotiation if changes to the LM are needed. If a teacher receives an argument from a student – e.g. artifact-focussed discussion is supported by a linear threaded discussion associated with a particular node (activity or competency) – they can connect to a URL given as evidence. In Figure 1 (right) this is a student claiming their ability to use mathematical information in communication (in English) by a spreadsheet calculating expenditure for a holiday. Figure 1 also shows how the teacher can add LM data (clicking on stars relating to competencies) and provide feedback in fields for strengths and suggestions for how to proceed.

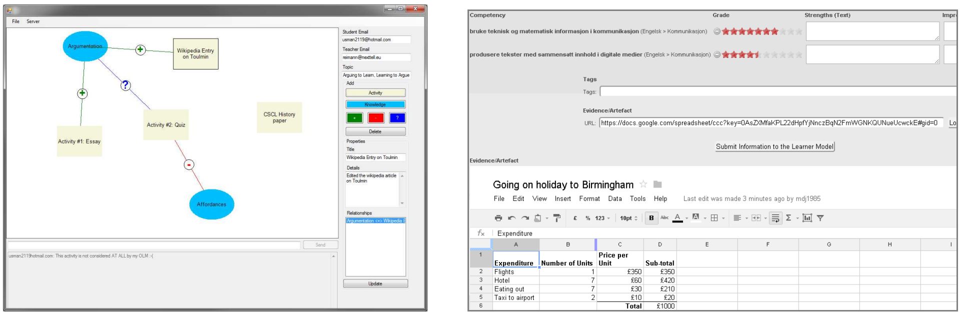


Figure 1: Next-TELL learner model negotiation example

Our example is from the Norwegian national competence goals and curriculum plan for English (see [28]), using the “communication” competence of “use of technical and mathematical information in communication”. After receiving a challenge from the negotiation tool, a teacher may acknowledge the evidence (spreadsheet) in the feedback fields, but also, for example, explain: weighting of the evidence is low compared to more extensive activities (e.g. marked essay, interaction in a virtual world); the data has since been superceded; or, when aggregated with other data, this entry has relatively little influence - if, for example, a student was challenging a skill at a broader level such as communication rather than use of technical & mathematical information in communication. Alternatively, discussion may be face-to-face. The negotiation tool may still be used to help a student understand the argument/evidence relationships: for example, teachers may explain, change or add new evidence nodes (for, against, unknown). This allows users to perform the kind of discussion required for negotiated learner modelling, including the following key issues identified above: evidence, challenge and artifacts. Instead, negotiation may occur around the LM visualisations and evidence, without using the the negotiation tool. However discussion occurs, some degree of learning might take place during this process [13]. This will also need to be reflected in the LM. The current solution is for the teacher to further update the model should such learning be identified (as described with reference to Figure 1).

To address concerns that teachers may be reticent to accept claims without evidence [24], control of the LM is not with the student. The student can enter self-assessments (as above), but these do not override data unless agreed by the teacher. This teacher control is similar to the power of the system in persuadable LMs ([4],[5]), but the teacher may also initiate negotiation if they consider this beneficial (e.g. to encourage reflection). Thus, there is also some similarity to the symmetrical approach of fully negotiated LMs ([11],[12]).

Acknowledgement

This project is supported by the European Community (EC) under the Information Society Technology priority of the 7th Framework Programme for R&D under contract no 258114 NEXT-TELL. This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

References

[1] Holt, P., Dubs, S., Jones, M. & Greer, J. (1994). The State of Student Modelling, in J.E. Greer & G.I. McCalla (eds), *Student Models: The Key to Individualized Educational Systems*, Springer, New York.
[2] Woolf, B.P. (2010). Student Modeling, in R. Nkambou, J. Bourdeau & R. Mizoguchi (eds). *Advances in Intelligent Tutoring Systems*, Springer-Verlag, Berlin Heidelberg, 267-279.
[3] Bull, S. Jackson, T. & Lancaster, M. (2010). Students' Interest in their Misconceptions in First Year Electrical Circuits and Mathematics Courses, *Int. J. of Electrical Engineering Education* 47(3), 307-318.
[4] Thomson, D. & Mitrovic, A. (2010). Preliminary Evaluation of a Negotiable Student Model in a Constraint-Based ITS, *Research and Practice in Technology Enhanced Learning* 5(1), 19-33.
[6] Mabbott, A. & Bull, S. (2006). Student Preferences for Editing, Persuading and Negotiating the Open Learner Model, in M. Ikeda et al. (eds), *Intelligent Tutoring Systems*, Springer, Berlin, 481-490.
[7] Hsiao, I-H. Bakalov, F., Brusilovsky, P. & Koenig-Ries, B. (2011). Open Social Student Modeling: Visualizing Student Models with Parallel Introspective Views. UMAP 2012, Springer, Berlin, 171-182.
[8] Kump, B., Seifert, C., Beham, G., Lindstaedt, S.N. & Ley, T. (2012). Seeing What the System Thinks You Know - Visualizing Evidence in an Open Learner Model, *Proceedings of LAK 2012*, ACM.
[9] Bull, S. & Kay, J. (in press). Open Learner Models as Drivers for Metacognitive Processes, to appear in R. Azevedo & V. Aleven (eds), *International Handbook of Metacognition and Learning Technologies*.
[10] Bull, S. & Kay, J. (2010). Open Learner Models, in R. Nkambou, J. Bordeau & R. Miziguchi (eds), *Advances in Intelligent Tutoring Systems*, Springer, 318-338.
[11] Bull, S. & Pain, H. (1995). 'Did I Say What I Think I Said, And Do You Agree With Me?': Inspecting and Questioning the Student Model, in J. Greer (ed), *AIED95*, AACE, Charlottesville VA, 501-508.
[12] Dimitrova, V. (2003). STyLE-OLM: Interactive Open learner Modelling, *IJAIED* 13(1), 35-78.
[13] Brna, P., Self, J., Bull, S. & Pain, H. (1999). Negotiated Collaborative Assessment through Collaborative Student Modelling, *Workshop: Open, Interactive & Other Overt Approaches to Learner Modelling*, AIED 1999, 35-42.
[14] Richardson, W. (2009). *Blogs, Wikis, Podcasts, and Other Powerful Web Tools for Classrooms*, Corwin Press, California.
[15] Raybourn, E. & Regan, D. (2011). Exploring E-portfolios and Independent Open Learner Models, *Interservice/Industry Training, Simulation and Education Conference*, Orlando, Florida. Paper 11361.
[16] Morales, R., Van Labeke, N., Brna, P. & Chan, M.E. (2009). Open Learner Modelling as the Keystone of the Next generation of Adaptive Learning Environments. In C. Mourlas & P. Germanakos (eds), *Intelligent User Interfaces*, Information Science Reference, ICI Global, London, 288-312.
[17] Mazzola, L. & Mazza, R. (2010). GVIS: A Facility for Adaptively Mashing Up and Representing Open Learner Models, *EC-TEL 2010*, Springer Verlag, Berlin Heidelberg, 554-559.
[18] Reimann, P., Bull, S., Halb, W. & Johnson, M. (2011). Design of a Computer-Assisted Assessment System for Classroom Formative Assessment, *CAF11*, IEEE.
[19] Baker, M.J. (1990). *Negotiated Tutoring*, Unpublished PhD thesis, Open University, Milton Keynes UK.
[20] Pilkingon, R. (1999). *Analysing Educational Discourse: The DISCOUNT Scheme*, Technical Report 99-2, Computer Based Learning Unit, University of Leeds, UK.
[21] Kerly, A. & Bull, S. (2008). Children's Interactions with Inspectable and Negotiated Learner Models, in B.P. Woolf, E. Aimeur, R. Nkambou & S. Lajoie (eds), *ITS 2008*, Springer, Berlin, 132-141.
[22] Van Labeke, N., Brna, P. & Morales, R. (2007). Opening up the Interpretation Process in an Open Learner Model, *IJAIED* 17(3), 305-338.
[23] Toulmin, S. (1959). *The Uses of Arguments*, Cambridge University Press. (Cited in [22]).
[24] Zapata-Rivera, D., Hansen, E., Shute, V.J., Underwood, J.S. & Bauer, M. (2007). Evidence-Based Approach to Interacting with Open Student Models, *IJAIED* 17(3), 273-303.
[25] Alotaibi, M. & Bull, S. (2012). Using Facebook to Support Student Collaboration and Discussion of their Open Learner Models, *Workshop: Web 2.0 Tools, Methodolog, & Services for Enhancing ITS*, ITS 2012.
[26] Zapata-Rivera, D. & Greer, J. (2004). Interacting with Inspectable Bayesian Student Models, *IJAIED* 14.
[27] Pain, H., Bull, S. & Brna, P. (1996). A Student Model 'For Its Own Sake', *Euro-AIED*, Lisbon, 191-198.
[28] Bull, S., Wasson, B., Johnson, M.D., Petters, D. & Hansen, C. (2012). Helping Teachers Effectively Support Group Learning, *Workshop: Intelligent Support for Learning in Groups*, ITS 2012.

Definition Response Scoring with Probabilistic Ordinal Regression

Kevyn COLLINS-THOMPSON[a*], Gwen FRISHKOFF[b] & Scott CROSSLEY[b]
[a] *Microsoft Research, Redmond, U.S.A.* [b]*Georgia State University, Atlanta, U.S.A.*
*Contact author: kevynct@microsoft.com

Abstract: Word knowledge is often partial, rather than all-or-none. In this paper, we describe a method for estimating partial word knowledge on a trial-by-trial basis. Users generate a free-form synonym for a newly learned word. We then apply a probabilistic regression model that combines features based on Latent Semantic Analysis (LSA) with features derived from a large-scale, multi-relation word graph model to estimate the similarity of the user response to the actual meaning. This method allows us to predict multiple levels of accuracy, i.e., responses that precisely capture a word's meaning versus those that are partially correct or incorrect. We train and evaluate our approach using a new gold-standard corpus of expert responses, and find consistently superior performance compared to a state-of-the-art multi-class logistic regression baseline. These findings are a promising step toward a new kind of adaptive tutoring system that provides fine-grained, continuous feedback as learners acquire richer, more complete knowledge of words.

Keywords: Intelligent tutoring systems, definition scoring, ordinal regression.

1. Introduction

Because knowledge of a word's meaning is often acquired gradually, by exposure to the word over time and in different contexts, learners may have partial or incomplete knowledge of many words (Frishkoff, Collins-Thompson, Perfetti & Callan, 2008; Frishkoff, Perfetti, & Collins-Thompson, 2011). Further, they may benefit from instruction that is tuned to support different kinds of interactions with words that are partially known versus ones that are unknown (Frishkoff, Perfetti, & Collins-Thompson, 2010). In previous work we have described a method called MESA (Markov Estimation of Semantic Association), for estimating degrees of word knowledge by applying a random walk model to compute the distance between a user-generated synonym for a newly learned word and the actual (target) meaning. We used MESA to examine average trajectories across words within different instructional categories, e.g., when encounters with a new word are massed *vs*. spaced or when contexts that provide more *vs*. fewer clues to a word's meaning. These average measures have provided insights into how word knowledge develops through time and as a function of different learning and instructional variables, an important step towards an interactive and dynamic approach to vocabulary training.

In this study, we describe an extension of MESA for estimating degrees of word knowledge on a trial-by-trial basis. As in previous work, each learner is presented with a sentence that contains a rare target word, such as *aleatoric*, and is asked to provide a synonym for this word. The inputs to our model predictor are a *response word* from the student, along with a *target word* that the student is aiming to learn. The prediction output is a number on a four-point ordinal scale that captures how closely the student's response word matches the meaning of the target word. Our definition response scoring approach is based on probabilistic ordinal regression. It exploits rich semantic features on multiple types of word relations, provides probabilistic scores and confidence estimates, and can achieve satisfactory performance on relatively small sets of human-labeled examples for training. We provide an initial evaluation of this approach by evaluating its prediction accuracy and the ability of the input features to discriminate between four levels of response accuracy.

2. Method

2.1 Dataset

Our target word list was a set of 60 English adjectives, verbs, and nouns selected by trained psychology experts. These words ranged from 'rare' to 'very rare' according to their frequency in the Kučera-Francis corpus (1979) with the requirement that very rare words appeared no more than 1 time out of 1 million tokens. We also provided, for each target word, a list of 1 to 3 reference words, which were higher-frequency synonyms that summarized the meaning of the target. For example, the reference words for the rare word 'limpid' were 'clear' and 'transparent'. The scoring algorithm uses these reference words as secondary targets when the target word itself is extremely rare and is found in few resources.

To create the labeled examples, each target word was paired with a short definition and two instances of the word in context: a high-constraint context, which provided rich cues to meaning (e.g., "I could not see a thing in the *X* room until I found the light switch."), and a low-constraint context that provided few if any cues to meaning (e.g., "Sharon did not expect to find that it would be this *X*."). For each target word, coders were asked to provide four kinds of responses:

(1) **Best Fit**: A word that matches the target and can be used in both contexts;
(2) **Strongly Related**: A word related to the target that can be used in both contexts;
(3) **Weakly Related**: A word that is weakly related to the target definition and can only be used in the low-constraint contexts;
(4) **Unrelated**: A word unrelated to the target, which cannot be used in either context.

We also evaluated scoring of *antonyms* of the 60 target words. Antonyms are challenging for many word similarity algorithms to score correctly because their semantic qualities are easily confused with those of synonyms. To label the responses, we employed three expert coders, resulting in three response files, each with 240 responses (60 targets x 4 response words, corresponding to the 4 ordinal levels as above)[1].

2.2 MESA Features

Prior work has used multiple resources to compute semantic distance, such as co-occurrence information, WordNet 2.0 (Harabagiu, Miller & Moldovan, 1999) dictionaries, and other resources (Mihalcea, Corley, & Strapparava, 2006). Each of these resources covers only a fraction of the potential relations between word pairs. By combining multiple resources using probabilistic chains of inference, it may be possible to bridge key gaps in a semantic network model. The MESA model adopts this approach (Collins-Thompson & Callan, 2007), and assigns a likelihood score to each target word on each learning trial. The target word's likelihood is derived from the stationary distribution of the Markov chain, which is approximated using a random walk. Details on the multiple word relations used by MESA are described in (Collins-Thompson & Callan, 2007).

[1] Researchers interested in using this dataset (non-commercial research purposes only) should contact the authors.

We used MESA synonymy, association, and morphology relations since these are the most effective combination for scoring synonyms. We added the ability to dynamically create new edges between terms that are not in the current word graph, and all current graph nodes. This is required us to handle (spell-corrected) free-form response words, such as those that would be likely to occur within an intelligent tutor. The edges are given uniform probability here, but more semantically focused schemes are possible and might give further prediction gains. In practice, we use a small number of walk steps (five) on a sparse representation of the word graph to perform the random walk. The random-walk based features we derived were the minimum, maximum, and average MESA log-likelihood scores of a response word, computed over all reference words for the target.

2.3 Features based on LSA

To add features that exploit word co-occurrence as a source of semantic information, we computed the LSA similarity score between each coder's response and the target concept, averaged over all reference words for the target, using the LSA Pairwise Comparison term-term comparison (http://lsa.colorado.edu/) with the default parameter settings. The resulting score between 0 and 1 was our LSA feature for that (target, response) pair.

2.4 Ordinal regression baseline

We used Gaussian Process Ordinal Regression, a state-of-the-art method recently introduced by Chu and Ghahramani (2005). GPOR outperforms previous approaches such as SVM-based ordinal regression or metric regression (Chu and Ghahramani, 2005). The GPOR method provides probabilistic prediction with confidence estimates for prediction and incorporates feature weighting as part of its model learning. GPOR also explicitly models the ordinal nature of the ratings.

2.5 Multi-class logistic regression baseline

We compare the effectiveness of GPOR for definition response scoring with a multi-class classification baseline using regularized logistic regression (Andrew and Gao, 2007). For each of the four response levels, a one-vs-all log-linear model was learned using the same training data as used for GPOR, also using 3-fold cross-validation. The L1 and L2 regularization weights were both set to default values of 1.0.

Prediction method	Precision (Micro-averaged)	Precision (Macro-averaged)
Random	0.250	0.250
Multi-class logistic regression	0.473	0.376
GPOR (all features)	**0.500**	**0.461**

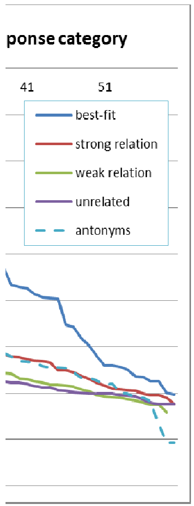
Table 2. GPOR achieves higher precision than multi-class logistic regression baseline

3. Evaluation

We evaluated prediction effectiveness of GPOR using MESA and LSA features with the standard quadratic weighted kappa measure (Fleiss, 1971) and micro/macro-averaged precision (correct vs. incorrect label prediction). We used GPOR settings of a Gaussian kernel with noise variance set to S = 0.40, and default settings for other parameters. We pooled the labeled datasets from the three raters, and used 3-fold cross-validation to produce train/test splits. Examples for the case where two raters' data are used to predict the ratings for the third rater are shown in Table 1. The results for all rater combinations (where R1+R2 → R3 means raters 1 & 2 are used for training and rater 3 for testing) are given, showing the full confusion matrix: each row corresponds to a different 'true' label, and columns show the predicted labels in each column. Weighted kappa varied from a minimum of 0.416 to a maximum of 0.519 (with kappa being on a [0,1] scale). Precision for individual labels was best for labels 1 and 4 in a range of 70 to 80%. Predicting the intermediate labels 2 and 3 was more difficult and had lower precision of 12 to 35%. Space does not permit showing a full learning curve analysis, but we found that as the amount of training data was varied: a) GPOR had consistently higher precision than LR across all training set sizes and b) this difference increased for small training sets (less than 20% of the original size).

3.1 Baseline comparison

The micro- and macro-average precision comparisons between GPOR and baselines based on random labeling and multi-class logistic regression are shown in Table 2. The random baseline results in a precision of 0.25 per category, since we have the same number of training examples for all four categories, and randomly picking a category is correct 25% of the time. GPOR prediction attains superior prediction accuracy over logistic regression for both micro- and macro-averaging (with each test slice having 240 instances).



'true' response
te between the
orted by score.
hood under the

3.2 Feature comparison

To compare LSA and MESA (random-walk) score properties across response categories, we plotted their scores for responses categorized by their 'true' label (here, from rater 3). The results are shown in Figure 1. Both methods discriminate among the best-fit and unrelated categories: LSA was less effective at discriminating strongly- from weakly-related words. MESA scores were less effective at distinguishing weakly- from unrelated words. Antonyms were scored by both methods as intermediate between 'synonym' and 'strongly related', which seems appropriate for this task.

4. Conclusion

To summarize, we have shown that a supervised approach based on Gaussian Process ordinal regression can exploit relatively small amounts of expert-labeled training data for competitive performance on a difficult prediction problem: scoring definition responses on an ordinal scale. Our approach uses classification features that combine MESA and LSA scores to capture complementary aspects of word relationships. With these features GP regression achieves consistently higher precision than a multi-class logistic regression baseline, over a range of training set sizes. Further performance gains are likely with additional feature sets or by refining the prediction model.

Our long-range goal is to develop a robust method for online scoring that can be embedded in an intelligent tutoring system (ITS). By tracking changes in word-specific knowledge on a single trial basis, the ITS will be able to provide feedback to students in real time, enabling them to adjust their focus and strategies on subsequent trials. Further, the ITS will be able to adapt the presentation of stimuli based on student performance, combined with cognitive models of robust word learning. In this context, we view our results as a promising step towards an adaptive tutoring system that provides fine-grained, continuous feedback as learners gain richer and more complete knowledge of words.

References

G. Andrew and J. Gao. (2007) Scalable Training of L_1-Regularized Log-Linear Models. *Proc. of ICML-2007.*

W. Chu and Z. Ghahramani. (2005) Gaussian Processes for Ordinal Regression. *J. Machine Learning Research.* 6 (July 2005). pp. 1019–1041.

K. Collins-Thompson and J. Callan. (2007) Automatic and Human Scoring of Word Definition Responses. HLT-NAACL 2007. pp. 476–483.

S. Deerwester, S. Dumais, T. Landauer, G. Furnas and R. Harshman. (1990) Indexing by latent semantic analysis. Journal of American Society for Information Science. 41:391–407, 1990.

J.L. Fleiss. (1971) Measuring Nominal Scale Agreement Among Many Raters. Psychological Bulletin, Vol. 76, No. 5. pp. 378–382.

W.N. Francis and H. Kučera. (1979) Standard Corpus of Present-Day American English. Dept. of Linguistics, Brown Univ., Providence, 1979.

G. Frishkoff, K. Collins-Thompson, C. Perfetti, J. Callan. (2008) Measuring incremental changes in word knowledge: Experimental validation and implications for learning and assessment. *Behavior Research Methods*, Vol. 40, No. 4. pp. 907–925.

G. Frishkoff, C. Perfetti, K. Collins-Thompson. (2010) Lexical quality in the brain: ERP evidence for robust word learning from context. *Developmental Neuropsychology*, Vol. 35, No. 4. pp. 376–403.

G. Frishkoff, C. Perfetti, K. Collins-Thompson. (2011) Predicting Robust Vocabulary Growth from Measures of Incremental Learning. *Scientific Studies of Reading*, Vol. 15, No. 1. pp. 71–91.

S. Harabagiu, G. Miller, D. Moldovan. (1999) WordNet2 - a morphologically and semantically enhanced resource. In Proceedings of SIGLEX-99, pp. 1–8.

R. Mihalcea, C. Corley, and C. Strapparava. (2006) Corpus-based and Knowledge-based Measures of Text Semantic Similarity. AAAI 2006.

Analyzing lecturer's speech and slide by estimated word difficulty

Yoshinobu Kito[a*], Kiyoaki NAKANISHI[a], Nobuyuki KOBAYASHI[b], Hiromitsu SHIINA[c] & Fumio KITAGAWA[c]
[a]*Graduate School of Informatics, Okayama University of Science, Japan*
[b]*Faculty of Human Sciences, Sanyo Gakuen University, Japan*
[c]*Faculty of Informatics, Okayama University of Science, Japan*
*sigel4280@live.com

Abstract: As research goal of providing the difficulty level of the Japanese used in lectures from the viewpoint of foreign exchange students and lecturers who are not native speakers of Japanese, we propose a system for providing Japanese word difficulty used in speech and slides of lecturers in lectures. The grades of the old version of Japanese Language Proficiency Test are used to determine the Japanese word difficulty. In the case that the Japanese word difficulty is already determined by materials, that Japanese word difficulty is used, and in the case of words with an unknown word difficulty, estimates reached by use of a support vector machine are used to arrive at a word difficulty. This system provides three kind of information of lecture by word difficulty.

Keywords: Lecture difficulty, Word difficulty, Support Vector Machine

Introduction

For foreign exchange students whose native language is not Japanese, it is not easy to attend lectures intended for Japanese people and they will have trouble comprehending such lectures unless they achieve a certain level of proficiency in Japanese. And yet, lecturers themselves do not understand the difficulty of speech and slides from the perspective of foreign exchange students. We are conceivable that it would be useful for both sides to be able to know the level of usage of Japanese words used in lectures. Students would be able to understand the difficulty of lectures and learn words that are necessary, while lecturers would be able to understand problems and ways to improve slides that they created and understand the nature of the spoken language used during the lecture with the insight gained from knowing the Japanese word difficulty used, and also update their slides and improve their lectures. In this research, we are developing a system that provides lecture difficulty on Japanese word difficulty using VOD lectures [1] which use an internet environment in which the speech and materials used in a lecture are saved as files.

1. System Summary

In this system, the data that is processed is subtitle data that is obtained from conversion of speech during the lecture into subtitles, as well as data that is extracted from the text portions of PPT files. The grade from the JLPT [2] which corresponds to each noun that is obtained from each of these sources is taken as the word difficulty and the Japanese word difficulty of a VOD lecture can thus be rated.

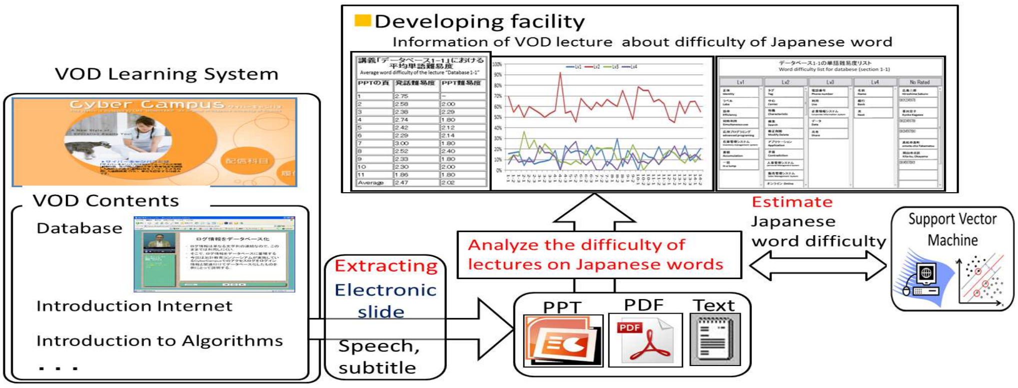


Figure 1 System Layout

2. Estimation of Japanese Word Difficulty

We use support vector machine (SVM)[3] to estimate Japanese word difficulty. It is a classification algorithm and the feature to assign each grade of JLPT [2] to classes.

2.1 Building Learning Parameters of SVM from Dictionary Data

The hypothesis surrounding the estimation of the grade of Japanese word difficulty using dictionary data is that with regards to the relation between entry words and their semantic descriptions, the grade of Japanese word difficulty used in semantic descriptions is equal to or easier than the entry word itself. Based on this hypothesis, we assumed a correlation between the grade of difficulty distribution according to the grade of the words expressed in the semantic description and the grade of difficulty for the entry word and generated the learning parameters. As evaluation experiment in this study, we used Tokuhiro's data [4] as the initial data, and Meikyo Japanese Dictionary [5] as entry words of dictionary data. In the case of the word “keiei (management)” in Figure 2, two Lv1 words, one Lv2 word, one Lv3 word and zero Lv4 words appear in the semantic description, and the word is trained using a combination of the learning parameter (1/4,3/4,0/4,0/4), and a Lv2 teaching signal.

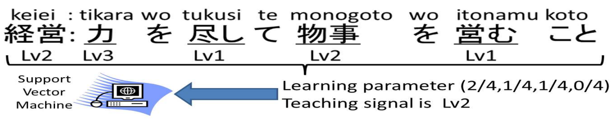


Figure 2 Building Learning Parameter

2.2 Difficulty Level Estimation of Compound Words using Web Search Data by SVM

There are many Japanese compound words that are not entried in dictionaries. Because of this we use descriptions obtained from web searches to estimate word difficulties. The method used for estimation is to enter the compound word into a web search, and then use the distribution of grade of word difficulty among the words near the compound word in the search results. In Figure 3, the learning parameter (1/2, 1/2, 0/2, 0/2) is created from the description of web search results, then SVM estimates the grade of difficulty is Lv2.

Figure 3 Estimation of Difficulty for Compound Words

3. Analysis of Lecture difficulty on Japanese Words

In this research, we are developing analyze system of lecture difficulty of Japanese word. It provides information related to Japanese word difficulty from speech and PPTs used in a lecture and the variation for each lecture. In particular, we provide the average word difficulty level for each lecture, word difficulty ratio distribution and word difficulty list.

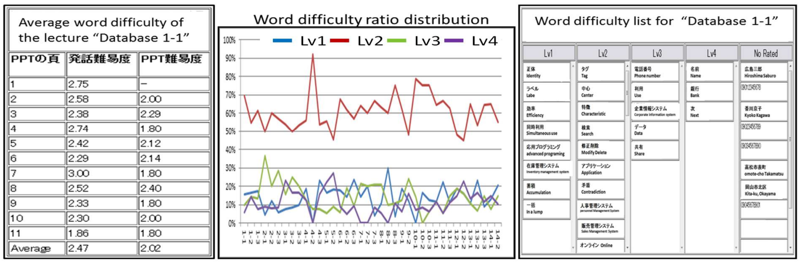


Figure 4 Information of Lecture Difficulty

The survey on the comprehensiveness of the word difficulty list in Table 1 shows the number and percentage of words in each grade that were judged to be Lv1 words. And it shows that almost all the difficult words are included up until Lv2, and also that the ratio of words judged to be more difficult by the Chinese foreign exchange student was low.

Table 1 Survey on Comprehensiveness of Word Difficulty Level List

Estimated grade of Word	Average for Japanese students	Chinese student
Lv1 (8)	3.333 (0.416)	2 (0.25)
Lv2 (36)	7.666 (0.213)	2 (0.053)
Lv3 (5)	0.333 (0.067)	0 (0.0)
Lv4(3)	0.000 (0.0)	0 (0.0)

4. Future works

This research proposed various types of information about word difficulty in order to help lecturers and lecture takers to understand the situation of lectures. The word difficulty of speech and PPT show a trend whereby the word difficulty used in PPT has a higher difficulty than those used in speech. Also, since lectures proceed by following a PPT, we think that a lecturer’s speech is somewhat dependent on the contents of the PPT. We believe that this means that it can be useful for lecturers to be careful of the distribution of word difficulty levels in PPT and when displaying of words with high difficulty. As future works, we will consider developing of learning system that allows for preparatory study of the Japanese needed to understand a lecture. We also have plans to create a general Japanese difficulty level index that takes into account not only word difficulty level, but that also indexes the complexity of a sentence according to construction.

References

[1] Kitagawa, F., Onishi S. (2007). An Experiment on selective Course with Face-to-Face or/and E-learning, and Students Behavior (in Japanese), *Japan Society of Educational Information*, *22*(3), 57-66.
[2] Japanese Language Proficiency Test, http://www.jlpt.jp
[3] Vapnik,V. (1998). Statistical Learning Theory, *Springer*.
[4] Tokuhiro, Y. (2008), Kanji2100 Listed according to Frequency and Familiarity, *Sanseido*.(in Japanese)
[5] Kitahara,Y. (2010), Meikyo Japanese Dictionary Second Edition, *Taishukanshoten*. (in Japanese)

Hybrid Question Generation Approach for Critical Review Writing Support

Ming Liu[a*], Rafael A. Calvo[a] and Vasile Rus[b]
[a]School of Electrical and Information Engineering, University of Sydney, Australia
[b]Department of Computer Science, University of Memphis, USA
[*]Ming.Liu@sydney.edu.au

Abstract: Research towards automated feedback can build on the work in other areas. In this paper we explore question generation techniques. Most research in question generation has focused on generating content specific questions that help students comprehend a set of documents that they must read. However, this approach is not so useful in writing activities, as students would generally understand the document that they themselves wrote. The aim of our project is to build a system which automatically generates feedback questions for academic writing support, particularly for critical review support. This paper presents our question generation system which relies on both syntax-based and template-based approaches, and uses Wikipedia as background knowledge.

Keywords: Question Generation, Academic Writing Support, Natural Language Processing

Introduction

Critical review requires students to read relevant sources in enough detail so that they can present a fair and reasonable evaluation of the selected sources. The ability to critically review relevant literature is a key academic skill. Afolabi [1] identified the most common problems that students have when doing literature review which includes not being sufficiently critical, lacking synthesis and not discriminating between relevant and irrelevant materials.

Questions can be a very useful form of feedback to develop critical reviewing. An example of a generic feedback question is *Have you critically analyzed the literature you use? Instead of just listing and summarizing items, have you assessed them, discussing strengths and weaknesses?* However, such question is too general and not likely to be effective for helping a student write on a specific topic.

More specific feedback questions are needed to develop students' critical review skills and help them with their writing. Table 1 shows an example of two system-generated questions from an engineering research student's literature review document. These questions were triggered by the citation sentence, copied from a literature review document, shown at the top of the table. Q1 asks the student-writer to critically evaluate the citation sentence by asking for supporting evidence and other people's views. Q2 asks the writer to critically compare the technique Principal Component Analysis (PCA) with another similar technique called factor analysis.

Table 1: Example of system-generated feedback questions

Citation 1: Cannon challenge this view mentioning that physiological changes were not sufficient to discriminate emotions. *Q1: Why did Cannon challenge this view mentioning that physiological changes were not sufficient to discriminate emotions? Does any other scholar agree or disagree with Cannon?*

key phrase: Principal Component Analysis (PCA) Q2: *What do you think of the differences between PCA and factor analysis in relation to your project?*

1. The Automatic Question Generation Framework

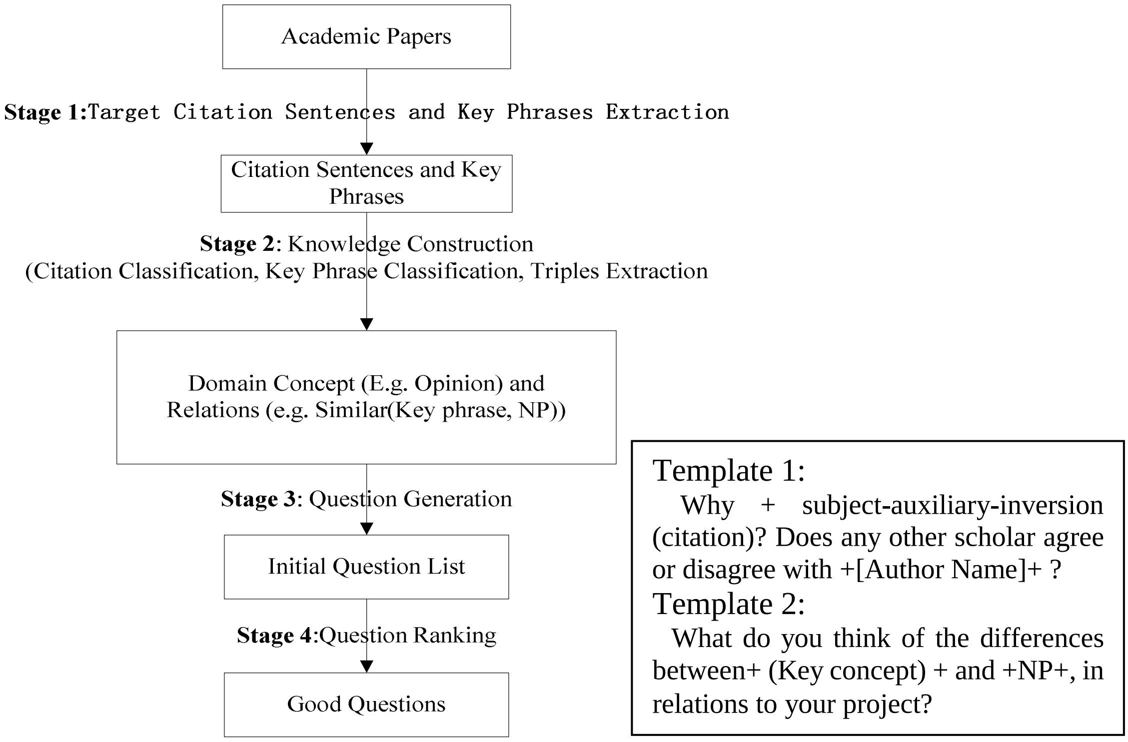


Figure 1: Four Stages Question Generation Processing

Our question generation approach is based on two important topic independent features of an article: citations and key phrases. The question generation process has been divided into four main stages as shown in Figure 1. In stage 1, the citations and key phrases are identified using regular expressions and the Lingo algorithm [2], respectively. In stage 2, according to the function of a citation, each citation is classified in one major citation category in the scientific register. We have defined the following 6 citation categories: Opinion, Result, Study Aim, System, Method and Application. Meanwhile, each key phrase is linked to a Wikipedia article and classified in a scientific instrument category based on the first section of the Wikipedia article. We defined 4 useful scientific instrument categories: Research Field, Technology, System and Term. If the key phrase is classified as one of the instruments, information extraction techniques are used to extract triples from the linked Wikipedia page. Each triple contains a key concept, another noun phrase (NP) or sentence (S), and their relation. We have defined 7 common relations including Has-limitation, Has-Strength, Include-Technology, Apply-to, Similar-to, Different-to and Kind-of. In stage 3, questions are generated based on these concepts and their relations. As we mentioned before, the goal is to generate specific questions using the syntax-based approach to transform the declarative citation sentence into questions. We also need deeper questions and use the template-based approach to fill the extracted citation or triples into pedagogical question templates. In stage 4, like Heilman and Smith's approach [3], a statistical question ranker is used to rank the quality of the generated questions in terms of readability.

For example, we assume that the citation and key phrase shown in Table 1 are extracted from stage 1. In stage 2, the citation 1 is classified as an Opinion concept. The Principle Component Analysis (Key Phrase) is linked to the Wikipedia article called PCA and classified as a Technology Concept. From the Wikipedia article, we extracted a triple which contains a key phrase, a noun phrase (NP)/sentence (S) and its relation, which is denoted RelationName(Key Phrase, NP/S). In stage 3, Q1 is generated using the template 1 shown in Figure 1. In this case, we need to perform Subject-Auxiliary Inversion commonly used in the syntax-based approach to generate “Why did Cannon challenge this view mentioning that physiological changes were not sufficient to discriminate emotions?” We then use the template-based approach to fill the extracted author name in the question template (“Does any other scholar agree or disagree with Cannon?”).

2. Discussion and Future Work

This paper presents a hybrid approach for generating specific deep questions which are used for critical review writing support. Citations and key phrases are the target elements for the question generation because they are important features of an article. Asking critical questions about citations can be effective for critical literature review writing because citations directly relate to the literature and such questions have high specificity. In this case, the questions generation process does not necessarily require any background knowledge. However, questions generated from key concepts (key phrases) used background knowledge extracted from Wikipedia. Wikipedia can be seen as huge lexical semantic resource that includes knowledge about domain specific terms. Therefore, it would be useful to ask questions about understanding these key concepts based on the knowledge extracted from Wikipedia.

Previous works [4-6] only separately proposed and evaluated the question generation based on citations or key phrases. Our future work will evaluate the new question generation framework and investigate the differences between questions generated from both.

Acknowledgements

This project was partially supported by a University of Sydney TIES grant, an Australian Research Council Discovery Project grant (DP0986873) and Google Research Award for measuring the impact of feedback on the writing process.

References

[1] Afolabi, M. (1992) "The review of related literature in research," *International journal of information and library research,* vol. 4, pp. 59-66.

[2] Osinski, S., Stefanowski, J., and Weiss, D. (2004),"Lingo: Search results clustering algorithm based on singular value decomposition," in *The International Conference on Intelligent Information Systems*, Zakopane, Poland, pp. 359--368.

[3] Heilman, M. and Smith, N. A.(2010) "Good Question! Statistical Ranking for Question Generation," presented at the Annual Conference of North American Chapter of the Association for Computational Linguistics - Human Language Technologies, Los Angeles.

[4] Liu, M., Calvo, R. A., Aditomo, A., and Pizzato, L. A. (to appear 2012) "Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support," *IEEE Transactions on Learning Technologies,*

[5] Liu, M., Calvo, R. A., and Rus, V. (2012) "G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support," *Dialogue and Discourse: Special Issue on Question Generation,* vol. 3, pp. 101-124.

[6] Liu, M., Calvo, R. A., and Rus, V.(2010) "Automatic Question Generation for Literature Review Writing Support," presented at the 10th International Conference on Intelligent Tutorial Systems, Carnegie Mellon University, USA.

Should We Use Examples in Intelligent Tutors?

Amir Shareghi Najar[*], Antonija Mitrovic
University of Canterbury, New Zealand
*Amir.shareghinajar@pg.canterbury.ac.nz

Abstract: Although examples are frequently used by human tutors, they are not common in Intelligent Tutoring Systems (ITS). Previous research studies over the last three decades compared learning from examples to unsupported problem solving. Only recently there have been studies comparing learning from examples to problem solving in ITSs. This paper reviews those studies. We discuss unsolved issues such as when and how examples should be provided in intelligent tutoring systems, and some options to improve learning from examples.

Keywords: Learning from examples, problem solving, self-explanation

Introduction

In this paper, we review related work on using examples compared to both unsupported and tutored problem solving. We also review studies comparing different combination of problems with examples, and come up with our idea on how we can reinforce examples and problems with an appropriate Self-Explanation (SE) prompts. An SE prompt is a question that encourages the student to explain the knowledge encapsulated in a problem or an example. We look at previous studies from two different angles: studies comparing examples with unsupported problem solving, and studies comparing examples with supported problem solving in ITSs. In unsupported problem solving, learners do not get any feedback, while in tutored problem solving students receive feedback on their solution steps and final answers. First we start with those studies comparing learning from examples with unsupported problem solving.

There has been no agreement on how much assistance should be provided to students. However, it has been shown that maximum assistance (e.g. examples) is more efficient than minimal assistance (e.g. unsupported problem-solving) for novices [1]. Recently researchers focused on different example-based learning strategies. van Gog, Kester, & Paas [2], investigate the difference between worked examples only (WE), worked examples/problem pairs (WE-PS), problem/worked examples pairs (PS-WE) and problem solving only (PS) on novices. They show that the participants in WE and WE-PS had a higher performance in the post-test than PS and PS-WE. van Gog [3] conducted a study using Modelling Examples (ME) (i.e. the problem solution is demonstrated to learners by a model , who can be an expert or not) in two conditions PS-ME-PS-ME and ME-PS-ME-PS in the Frog Leap game. Result showed no difference in learning performance since the students learnt most after studying the second worked example.

Most of the prior studies showed the example effect in well-defined problems. Well-defined tasks are those for which there is an algorithm for solving problems [4] (e.g. mathematics, physics). Nevertheless , it has been shown that the worked-examples effect can be obtained in ill-defined problems like well-defined problems [5]. Kalyuga [6] show that instructional support which is enhanced with SE and self-visualization technique, may improve students' abilities to transfer their skill and knowledge. Hilbert & Renkl [7] investigated the best structure of examples to teach concept mapping. They found that students learn more when the examples are presented with SE than without it.

Now we review those studies compared worked-examples with ITSs. Schwonke et al. [8] compared a cognitive tutor (Geometry Tutor) with a new version which was enriched with faded worked examples. They conducted two experiments. In the first experiment, they found an improvement in learning time from using examples. In the second experiment, they used the think-aloud protocol in order to study relevant cognitive processes. According to the result, the efficiency advantage of worked examples was replicated.

McLaren and colleagues [9] discuss their three studies on example-based strategy using the stoichiometry tutor. In all the studies, the problem and example conditions were compared while students were given SE prompts after examples. The result showed no significant difference learning gains. McLaren and Isotani [10] show that the students benefit most by learning with worked examples only, at least with respect to learning time They found no significant difference in learning gain between the students who worked with examples only, problems only, or a mixture of examples and problems. However, the examples were followed by SE prompts while the problems were not. Salden and colleagues [11] compared fixed worked-out examples with adaptive ones. Fixed faded examples are the same for all students, but the adaptive faded examples are adapted with respect to the student's prior knowledge. The lab results indicate that adaptive examples led to a better learning gain compared to the other conditions. In contrast, the classroom results depict no significant difference in immediate post-test, but in the delayed post-test students who used adaptive examples learned more.

Most of the studies show that using worked examples in ITSs results in reduced learning time. Although there are some studies showing the higher learning gain or the faded examples, most studies have found no differences in the amount learnt. In addition, all the prior studies on using examples in ITSs were in Geometry, Chemistry and Algebra domains. All these tutors teach well-defined tasks. Therefore, there is a need for more research in order to explore the usage of examples in ill-defined tasks.

How should we design examples for ITSs

This section covers the design of examples when used in conjunction with problem solving in ITSs. In the following subsections we discuss a number of issues when we use examples with tutored problem solving.

When to give examples? It has already been shown that novices benefit more from the example strategy than learners who have enough prior knowledge to start practicing using the problem-solving strategy [12]. Traditionally, systems with tutored problem-solving strategy have indirectly followed this idea, and students with a high expertise level can solve a problem without using any hint level while students with a low level of expertise have to transfer the problem to an example gradually, until they can solve the problem. Although the research contributions are not conclusive to decide whether to use the combination of examples and problems or not, it is important to know whether novices benefit more from instant solutions (i.e. examples), or gradually solved problems (tutored problem solving).

How to design examples? Mayer [13] proposed seven principles for designing multimedia messages. Modeling examples and worked examples emphasize differently on these principles; therefore, using each of these two types of examples has its own advantage and drawbacks. Perhaps an adaptive example (worked-example/modeling example) may improve learning more than a fixed example, but when and how should ITSs switch between a worked example and a modeling example? We also can adapt examples based on their difficulty level. For instance, in a faded example, the solution steps can gradually be faded until an example (low difficulty) transforms to a problem (high difficulty). Prior studies like Schwonke and colleagues [8] show that students learn more effectively by using adaptive examples compared to tutored problem solving.

How to scaffold examples? Self-explanation, as an effective scaffolding strategy, is a potential option to scaffold examples, so a good future research question is to find an appropriate SE design to reinforce learning from examples. We think that SE prompts designed for problem solving are not appropriate for examples. Therefore, it is a good idea to categorise SE prompts into two new types named: example-adapted and problem-adapted prompts.

In conclusion, while many human tutors use different combinations of examples and problems, the previous studies that compared tutored problem solving with examples are not conclusive to replace this new teaching strategy with traditional problem-solving strategies in ITSs. Shareghi Najar & Mitrovic [14] suggest a new approach to use examples with tutored problems. In this model, for each example or problem, a corresponding SE prompt is provided that reinforces examples for procedural and problems for conceptual knowledge acquisition. Future research on using examples in ITSs will draw on three perspectives: when to give examples, how to design examples, and how to scaffold examples as the prior studies are not enough to show the best approach.

References

[1] Kirschner, P. A., Sweller, J., and Clark, R. E. 2006. Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. Educational Psychologist. 41:75–86

[2] van Gog, T., Kester, L., and Paas, F. 2011. Effects of worked examples, example-problem, and problem-example pairs on novices' learning. Contemporary Educational Psychology. 36:212–218

[3] van Gog, T. 2011. Effects of identical example–problem and problem–example pairs on learning. Computers & Education. 57:1775–1779

[4] Mitrovic, A., and Weerasinghe, A. 2009. Revisiting Ill-Definedness and the Consequences for ITSs. In The 14th International Conference on Artificial Intelligence in Education, ed. Vania Dimitrova; Riichiro Mizoguchi. Amsterdam, The Netherlands: IOS Press, p. 375–382.

[5] Rourke, A., and Sweller, J. 2009. The worked-example effect using ill-defined problems: Learning to recognise designers' styles. Learning and Instruction. 19:185–199

[6] Kalyuga, S. 2009. Instructional designs for the development of transferable knowledge and skills: A cognitive load perspective. Computers in Human Behavior. 25:332–338

[7] Hilbert, T. S., and Renkl, A. 2009. Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. Computers in Human Behavior. 25:267–274

[8] Schwonke, R., Renkl, A., Krieg, C., Wittwer, J., Aleven, V., and Salden, R. 2009. The worked-example effect: Not an artefact of lousy control conditions. Computers in Human Behavior. 25:258–266

[9] McLaren, B. M., Lim, S.-J., and Koedinger, K. R. 2008. When and How Often Should Worked Examples be Given to Students? New Results and a Summary of the Current State of Research. In Conference of the Cognitive Science Society, Austin, TX: Cognitive Science Society, p. 2176–2181.

[10] McLaren, B., and Isotani, S. 2011. When Is It Best to Learn with All Worked Examples? In Artificial Intelligence in Education. Springer Berlin / Heidelberg, p. 222–229

[11] Salden, R., Aleven, V., Renkl, A., and Schwonke, R. 2009. Worked Examples and Tutored Problem Solving: Redundant or Synergistic Forms of Support? Topics in Cognitive Science. 1:203–213

[12] Sweller, J., and Cooper, G. A. 1985. The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra. Cognition and Instruction. 2:59–89

[13] Mayer, R. E. 2002. Multimedia Learning. 2nd ed. New York, NY, USA: Cambridge University Press.

[14] Shareghi Najar, A., and Mitrovic, A. 2012. Using examples in intelligent tutoring systems. In The Intelligent Tutoring Systems (ITS) conference. p. 579-581

Developing Program Grading Suggestion Mechanism by Reusing Teacher Grading Records

Chih-Yueh Chou, You-Chia Cheng
Department of Computer Engineering and Science, Yuan Ze University, Taiwan
cychou@saturn.yzu.edu.tw

Abstract: In this paper, we describe a work-in-progress research to develop a virtual grading assistant system for assisting teachers in grading programming assignments based on the output of student programs. The system provides teachers with an interface to grade student programs through marking token patterns of the program output and offering grades. The system records the teacher marking output token patterns and grades as teacher grading cases. The system will match the outputs of other student programs with the previous record of teacher grading cases to suggest grading to reduce the grading load. The paper presents the system prototype and proposes two research issues, which we will investigate.

Keywords: Virtual teaching assistant, computer assisted programming grading, token pattern matching

Introduction

Assessing and grading student programs are time-consuming and labor-intensive for teachers and teaching assistants, thus some computer systems are developed for assisting teachers in assessing student programs through assessing the program output and program code [1,2]. Most systems evaluate the function correctness of student programs by compiling and executing the programs with test inputs and comparing the output of student programs with that of the model program [3-5]. Generally, these systems evaluate the program through strict comparison of program output. It may arouse two limitations. First, a program with correct function but with minor different output form or order may be evaluated as wrong [6]. For instance, an assignment asks students to write a program to calculate and output the factorial result of an input integer number. If the input of a test case is 5 and the program output of the model program is “The factorial result of 5 is 120”, a student program with the output “5! = 120” may be evaluated as wrong because the result of strict output comparison of these two outputs is different. Teachers usually setup rigid output format specifications to limit the student program output. Second, these systems evaluate programs either pass (correct) or failed (wrong), but teachers may grade programs within a range of grades from full mark to zero according to the correct ratio of the program output. Researchers proposed a program output comparison approach through token patterns to improve the flexibility of computer assessment [7].

Teachers can understand the meaning of program output and recognize the key token pattern, such as “120” in the factorial program when the test case input is 5, in program output, but computer assisted assessing systems are unable to understand program output and just compare the program output of student programs and that of the model program. Applying natural language processing techniques may solve this problem, but the system development is complex. Researchers propose an approach of complementing machine intelligence and human intelligence to develop virtual teaching assistant systems [8-10]; that is, using human intelligence to reduce the complexity of system development and applying machine intelligence to reuse human intelligence. This study develops a computer assisted grading system, named ProgramHelper, to support teachers with an interface for marking token patterns on program output and giving corresponding grades (human intelligence). The system also records teacher marking token patterns and grades as grading cases for matching the output of other student programs and providing grading suggestion to reduce teacher load (machine intelligence).

1. System

The ProgramHelper system includes interfaces for teachers to assign and grade program assignments and an interface for students to submit programs; three databases to store program assignments, student submitted programs, and teacher grading cases; and three modules to compile and execute student programs, to match token patterns with the program output, and to suggest grades. Using ProgramHelper, a teacher assigns program assignments by describing program specification, offering inputs and outputs of test cases, and providing a model program. Students submit their programs for assignments and the system collects student programs for teacher grading. When a teacher wants to grade a student program, the system first compiles the program. If the program has compile errors, the system shows the error messages to the teacher. Otherwise, the system executes the program with the inputs of test cases and shows both the outputs of student programs and outputs of model program of all test cases to the teacher. The system will attempt to match the student program output with previous stored teacher grading cases to suggest grades. If the system did not contain any teacher grading cases or the student program output did not match with any teacher grading cases, the teacher needs to grade programs by himself/herself.

The system enables teachers to grade student programs by marking token patterns within the output of student programs and offering corresponding grades to these token patterns. The teacher also needs to assign the matching rule whether each token pattern should be matched at specific location or location-free so that the system can apply the rule to match the token pattern with the output of other student programs. For instance, if the input of a test case of factorial program is 5 and the output of a student program is “The factorial result of 5 is 120”. The teacher can mark “120” as a token pattern, offer 100 points to the token pattern, and assign the location-free matching rule. The system will store the token pattern, grade, and matching rule as a teacher grading case. When the teacher grades another student program and the output of the same test case is “5! = 120”, the system will found “120” matched and suggest 100 points. If a program aims to output all the values of a 2-subscribed array in two-dimensional form in order, the teacher needs to mark each value of the array as a token pattern, offer the corresponding grade, and assign each token pattern at specific location. For instance, the second array value is 4, has 5 points, and is located at the second token of the first row. If the system successfully found the output of the student program matched with some previous teacher grading cases, the system displays all matched teacher grading cases and matched token patterns of student program output in yellow background color, and suggests grades with the total of grades of all matched teacher grading cases. For instance, if the output of student program for calculating factorial contains the token pattern “120”, the system suggests 100 points based on the previous teacher grading case.

2. Summary and Future works

We developed a prototype system of reusing previous teacher grading cases to suggest grading based on the program output. The goal of the system is to frequently and correctly provide grading suggestions to reduce teacher load. The assumption is that the teacher grading cases on the program output can be reused for grading other student programs; that is, teacher grading cases are repeatedly occurred. We will apply ProgramHelper system to several program assignments to investigate the effectiveness of the system.

The ProgramHelper system is designed to assist teachers in grading student programs based on the program output. The system could be applied for other usages:

- Providing students with immediate feedback as formative assessment. After the system collected many teacher grading cases, the system can be used to immediately provide students with system grading when students submit programs. Although the system grading may not be exactly correct and teachers also require to confirm or to modify the grading to offer the final grade, the system grading can be a formative assessment.
- Finding out the programs which have correct functions but do not conform to the output format specifications. The system can compare the output of model program with that of student programs both through strict comparison and through token pattern comparison. If student programs pass token pattern comparison but do not pass the strict comparison, the programs have correct function but do not conform to the output format specification. The system can advise them to revise the program output format.

Acknowledgments

The authors would like to thank the support of the National Science Council (NSC 100-2511-S-155 -004 -MY3).

References

[1] Ala-Mutka, K. M. (2005) A survey of automated assessment approaches for programming assignments, *Computer Science Education*, 15 (2), pp. 83-102.

[2] Douce, C., Livingstone, D., & Orwell, J. (2005) Automatic test-based assessment of programming: A review, *ACM Journal on Educational Resources in Computing*, 5 (3), Article No. 4.

[3] Higgins, C., Hergazy, T., Symeonidis, P., & Tsinsifas, A. (2003) The CourseMarker CBA system: Improvements over Ceilidh. *Education and Information Technologies*, 8, 287 – 304.

[4] Cheang, B., Kurnia, A., Lim, A., & Oon, W.-C. (2003). On automated grading of Programming Assignments in an academic institution. *Computers & Education*, 41, 121-131.

[5] Joy, M., Griffiths, N., & Boyatt, R. (2005). The BOSS Online Submission and Assessment System, *ACM Journal on Educational Resources in Computing*, Vol. 5, No. 3, Article No. 2.

[6] Tang, C. M., Yu, Y. T., & Poon, C. K. (2009) Automated systems for testing student programs: Practical issues and requirements. *Workshop Proceedings of the 17th International Conference on Computers in Education (Workshop on SPECIAL)*, 132-136.

[7] Tang, C. M., Yu, Y. T., & Poon, C. K. (2009). An approach towards automatic testing of student programs using token patterns. *Proceedings of the 17th International Conference on Computers in Education (ICCE 2009)*, 188-190.

[8] Chou, C. Y., Chan, T. W., & Lin, C. J. (2003). Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, 40(3), 255–269.

[9] Lin, C.J., Chou, C.Y., & Chan, T.W. (2008). Developing a computer-supported tutoring Interaction Component with Interaction Data Reuse. *Proceedings of the 9th International Conference on Intelligent Tutoring Systems (ITS 2008)*, 152-161.

[10] Chou, C. Y., Huang, B. H., & Lin, C. J. (2011) Complementary Machine Intelligence and Human Intelligence in Virtual Teaching Assistant for Tutoring Program Tracing. *Computers & Education,* 57 (4), 2303-2312.

Analysis of Qualities of Effective Teachers and Issues on Designing Educational Robots

Ilker YENGIN [a*], Sau-Lai LEE [b] & Martin Saerbeck [a]
[a] *Institute of High Performance Computing, A*STAR, Singapore*
[b] *Department of Marketing, Hong Kong Baptist University, Hong Kong*

Abstract: This paper provides a theoretical analysis of the feasibility of developing a robot that could take the role of an effective teacher by examining qualities and roles of effective teachers in the classrooms. .

Keywords: effective teachers; robot design; human robot interaction; educational robots; instructional technology.

Introduction

Robot technology is increasingly used in classroom settings not only for learning *about* robots but also for learning *from* robots. [1]. However, robots are mainly used as teaching assistants and educational media rather than as a fully autonomous teacher. This study reviews the qualities of effective teachers and the factors that affect student-robot interaction.

1. Modeling Effective Qualities And Teachers' Roles

To approach the problem of designing a robot as an effective teacher and to have an organizational pattern, we categorized the roles and qualities of teachers according to representative literature and examined the qualities that a teacher needs to exhibit to support the role. By integrating the roles and qualities of effective teachers, we formulated a model that links effective teacher roles and commonly expected teachers' qualities (Figure 1).
Roles are represented in blue boxes, sub-roles are in orange boxes, and the matching qualities are represented in yellow boxes. The labels for yellow boxes, which are effective teacher qualities, are given in the legend.

2. Available Robot Technologies for Education

Robots were used in different educational aspects [2]. Educational robots could be used as learning materials; learning companions; and teaching assistants [3] under the classification of "robot subject instruction", "robot assisted instruction" and "robot managed instruction" [4]. The teacher robot may fit in the last category. Learning language with robots is the most widely discussed scenario in research studies [5].

To understand the value of using robots in education, the robots' impact on the cognitive processes of learning should be evaluated [6]. Perhaps, the most unique attributes of robots are their automaticity and physical presence [7]. Robots can listen and talk [8], show facial expression and gestures [9] and affect children's perceptions [20] to have a social presence for supporting instruction [11]. Also, existing intelligent tutoring solutions may be implemented into robots to personalize learning much like a human tutor [12].

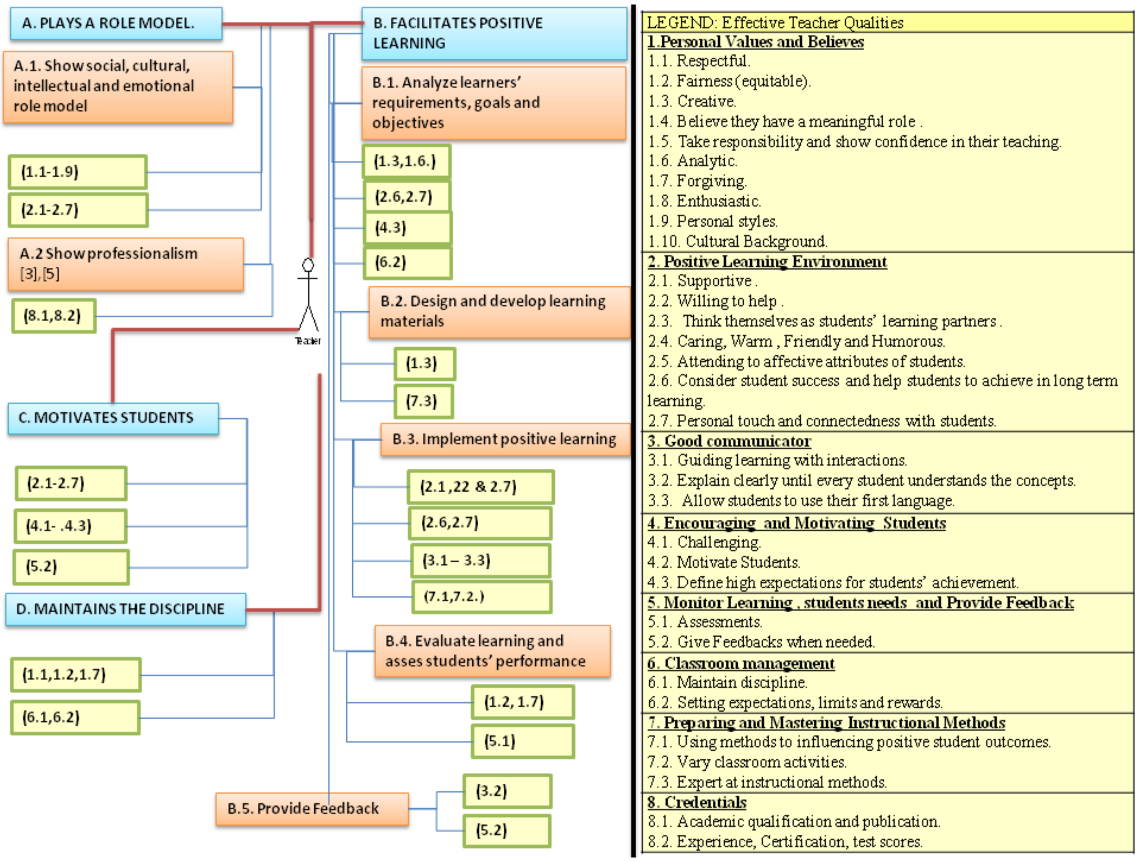


Figure 1. Effective Teacher's Roles and Qualities

2. Analysis of Teacher Roles and Issues on Robot Design

This section discusses to what degree the attributes of effective teachers as listed in Figure 1 can be implemented in an autonomous robot.

Plays a role model: It is generally possible to implement social feedback that approximates human performance on a robot. As these values consider mainly the output modalities of the robot (e.g. in terms of language, movement, sound) the designer has full control over the robot's behavior. Some qualities such as fairness are difficult to implement due to a lack of clear definition and not because of technical constraints. Superficially robots seem fair to all individuals, but in real life there might be situations where the definition of fairness is not as clearly defined as the performance rating. More problematic are the qualities those require human reasoning, such as being creative. Naturally, a robot is limited to its software and knowledge base.

Facilitates Positive Learning: Robot can seemingly create a positive learning environment. A Robots can be considered supportive simply due to its availability to students and its role to help. Robots also can be indefinitely patient to explain the same concept as many times as the students needed. Robots can also be programmed to provide alternative explanations. A shortcoming though is that robots could not create new learning content on its own. It is a question of degree to which detection needs to be implemented. It is an empirical question to what degree a teacher needs to be able to read the emotional state of a student to be considered an effective teacher.

Motivating Students: Robots have the potential to motivate students by giving emotional responses, dynamically selecting the order of learning content, by giving encouraging comments and by providing reward mechanism. Physical embodiment of robots may make them an authentic and engaging device. Robots may monitor students' performance and give feedback similar to computer based adaptive leaning systems.

Maintaining Discipline: Robots may follow and react according to a set of rules. It can set limits and deliver punishment and reward in class. Next to implementation issues there are ethical and political aspects. Being an authority figure might be a controversial topic to implement in a robot. There is a need for further research in this area in order to understand the social acceptance of robots as authority figures in classrooms.

3. Conclusion

This paper reviewed studies on qualities of effective teachers and the possibilities to use robots for education. Following the analysis above, the two teacher roles of being a role model and a motivator can be considered feasible with the current technology. The role of maintaining discipline mainly faces ethical and political issues, rather than implementation issues. Further investigation is needed to determine how much authority should be given to a robot. One of the main questions that remain is the degree to which a robot should perceive unconstrained input such as emotions and subtle behavior cues of students. But even these difficult tasks are becoming increasingly robust to be employed in an educational application. The question of long term interaction effects needs further investigation. An aspect that is often ignored is the challenge to integrate all qualities in a single platform which is necessary to conclusively measure the effectiveness of robotic tutors.

References

[1] A. Druin and J. A. Hendler. (2000).Robots for kids: exploring new technologies for learning. Morgan Kaufmann Pub.

[2] J. Han, M. Jo, S. Park, and S. Kim.(2005)."The educational use of home robots for children," in Robot and Human Interactive Communication. ROMAN 2005. IEEE International Workshop on, 2005, pp. 378-383.

[3] L. Y. Li, C. W. Chang, and G. D. Chen.(2009). "Researches on Using Robots in Education," Learning by Playing. Game-based Education System Design and Development, pp. 479-482.

[4] Q. Bingchen, D. Yanqiu, C. Lina, Q. Wei, and O. Yoshikun. (2009). "The Impact of Robot Instruction to Education Informatization," in Information Science and Engineering (ICISE), 2009 1st International Conference on, pp. 3497-3500.

[5] J. Han, S. Lee, E. Hyun, B. Kang, and K. Shin.(2009). "The birth story of robot, IROBIQ for children's tolerance," in Robot and Human Interactive Communication. RO-MAN 2009. The 18th IEEE International Symposium on, 2009, p. 318.

[6] R. B. Kozma.(1991). "Learning with media," Review of educational research, vol. 61, no. 2, p. 179.

[7] E. Hyun and H. Yoon.(2009). "Characteristics of young children's utilization of a robot during play time: A case study," in Robot and Human Interactive Communication,. RO-MAN 2009. The 18th IEEE International Symposium on, 2009, pp. 675-680.

[8] C. F. Shih, C. W. Chang, and G. D. Chen.(2007). "Robot as a Storytelling Partner in the English Classroom-Preliminary Discussion," in Advanced Learning Technologies. ICALT 2007. Seventh IEEE International Conference on, 2007, pp. 678-682.

[9] T. C. Liu and M. Chang.(2008). "Human-Robot Interaction Research Issues of Educational Robots," in Digital Games and Intelligent Toys Based Education, 2008 Second IEEE International Conference.

[10] E. Hyun, H. Yoon, and S. Son.(2010). "Relationships between user experiences and children's perceptions of the education robot," in Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction.

[11] J. Bates, "The role of emotion in believable agents.(1994)." Communications of the ACM, vol. 37, no. 7, pp. 122-125.

[12] A. C. Graesser, K. VanLehn, C. P. Rosé, P. W. Jordan, & D. Harter, (2001). Intelligent tutoring systems with conversational dialogue. AI magazine, 22(4), 39.

Design of Learning Hyperspace Construction System for Knowledge Refinement in Self-Directed Learning

Yukinori NARUTA[a*] **& Ryo OKAMOTO** [a]
[a] *Faculty of Science, Kochi University, Japan*
*yunaruta@is.kochi-u.ac.jp

Abstract: In self-directed learning, learners often finish their leaning with insufficient or incomplete knowledge. This is a serious problem for learners because they could not have an opportunity to improve their knowledge. One approach to improve their knowledge is to present them to others. In our research, we have focused on both of acquisition and publishing processes of knowledge and propose a method to specify and revise the errors in the acquired knowledge. In this paper, we describe the design and the development of the learning hyperspace construction system which based on the method to specify the incomplete or insufficient knowledge through a computerized presentation.

Keywords: self-directed learning, knowledge refinement, presentation rehearsal

Introduction

In self-directed learning, a learner chooses a learning object and a method for accomplish it. Recently, the learners can learn a many different kind of things via web. Additionally, they need to judge whether the learning is finished or not and select a resources for the learning. Therefore, to success the learning, learners must be able to judge whether acquired knowledge is sufficient. On the other hand, in the learning, it is hard for the learners to meta-recognize as these. Thus, the learners often finish their learning in an insufficient and incomplete knowledge state [1].

For the issue, to refine the learner's knowledge state, some previous studies have been supporting to promote their meta-recognition. Kashihara et al. have pointed out the possibility of supports for a learner's publication and a peer review of externalized knowledge [2]. In the peer review, the learners make an improvement of their knowledge state by criticisms from peers. We selected method of the presentation rehearsal as a type of peer review (e.g. correction of a paper, review on web). Because, this method enhances that the understanding of reviewer's comments through discussion process. Additionally, we focused on the support for a review work to get a lot of reviewer's comments [3].

For knowledge refinement, it is important for the learners to recognize an insufficiency or incompleteness of their knowledge correctly. The reviewer's comments are not necessarily showing the parts of the presenter's incomplete knowledge. Therefore, the presenter can not recognize the cause of the reviewer's comments correctly from the indication obtained in the review. Additionally, the presenter may revise the slides in the insufficient and incomplete knowledge state. Thus, we focused on the back review process to enhance knowledge refinement. In our research, we propose the framework to specify the causes of the insufficiency or incompleteness of presenter's knowledge, and we also have designed a support system based on the framework.

1. Enhancement of Knowledge Refinement through Presentation Rehearsal

1.1. Main Concept

We focused on the back review of presentation rehearsal, because we attempt to enhance the learner's knowledge refinement by effectively utilizing of reviewer's comments. In the back review, the comments from the peers are not necessarily a representation of the parts of the presenter's incomplete or insufficient knowledge definitely. Therefore, the learners should find out what to refine with peer’s comments. If the learners can verify the process to publish and learn, they can refine their knowledge by recognition of the parts of the knowledge. To achieve this, it is necessary to be verifiable their process. On the other hand, in many case, it is hard for the learners to verify them because of the invisible process.

1.2. Process Model for Specification of Insufficient or incomplete Knowledge

We proposed the process model of the knowledge refinement according to the creation process of the slides publishing [4]. This model is consists of knowledge publishing process and knowledge verification process.

In knowledge publishing process, the presenters make explicit representations of their knowledge. These knowledge representations are stored and utilized as resources to compose the presentation slides. In knowledge verification process, the presenters verify the creation process and their knowledge state with the comments has acquired in the rehearsal. If the presenters find defective knowledge, they should refine their knowledge through the relearning process. In this process, the learners relearn about necessary thing to refine their knowledge.

2. Learning Hyperspace

Kashihara et al. is described that annotating the web pages used in web-based learning is the effective method of knowledge reflection [5]. Therefore, when he/her accomplish the learning object, the learner associates a shorts note and bookmarks by a creating concept map. “It called summary of learning” (SL). Additionally, the concept map is used to organize and store the SL, because the learning objects have difference in level of abstraction. In the authoring process of a presentation slides, the learner chooses the concepts which are included the map and assembles them. At first, they make the presentation structure with the titles of the presentation and the concepts of SL. Then, they write a draft by relating a sentences with the structure, and make a script or slides. The learning hyperspace is a network built in such a method. We propose learning hyperspace as a method of representation and management the presenter's knowledge state.

3. Learning Hyperspace Construction System

To develop a prototype of the learning hyperspace construction system, the system consists of three tools. In knowledge publishing process, the learner uses “Knowledge Organizing Editor” and “Presentation Slides Editor”. The first is used in the learning. The learner publishes their knowledge which is acquired in the learning. They store memos or bookmarks with tags which are the keywords of learning contents. When the learners accomplish a learning object, they organize stored resources by creating the SL using the tags. Thus, the resources are organized and stored in hierarchical way as a concept map. The

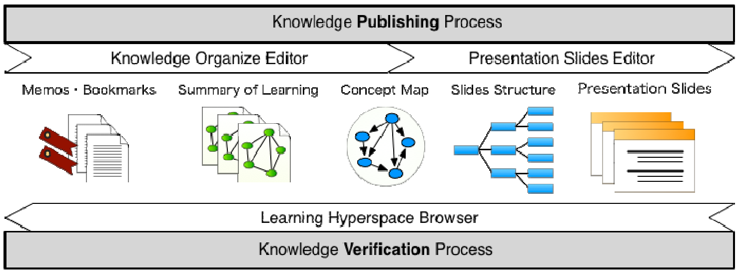


Figure 1: System Configuration

second is used in presentation slides authoring process. This editor can be used to create the presentation structure and compose presentation slides. The learner selects the required concepts from the learning hyperspace and creates the presentation structure using the concepts. Then, they write down details in the structure and create the presentation slides.

In knowledge verification process, the learners use "Learning Hyperspace Browser". The learners recognize incomplete parts in the knowledge by browsing the leaning hyperspace. They can verify the knowledge state by to use learning hyperspace, because learning hyperspace is related with the resource used in the slide authoring process. Therefore, they able to refine their knowledge through slide verifying process and relearning process. We have been developing a prototype system of the learning hyperspace construction system. Figure 1 shows the configuration of the system.

Conclusion

In this paper, we design the learning hyperspace construction system based on the knowledge refinement model through presentation rehearsal. And we also describe the prototype system which we have developed based on the model. This project is ongoing. For consideration of the reflection method, we develop the Learning Hyperspace Browser continuously.

Acknowledgements

This research is supported in part by Grant-in-Aid for Scientific Research (C) (No.22500925) from Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] A.Kashihara & S.Hasegawa (2003). Learning Bench: A Self-Directed Learning Environment on the Web. Proc. of ED-MEDIA, 1032-1039, Honolulu, Hawai: AACE.

[2] A. Kashihara and Y. Kamoshita (2006). Reflection by Knowledge Publishing, Lecture Notes in Artificial Intelligence 4252, 1002-1009.

[3] R.Okamoto, and A.Kashihara (2007). Presentation Review Supporting Environment with Realtime-created Hypervideo Tecnique. Proc. of International Conference on Computers in Education, 433-440, Hiroshima, Japan.

[4] R. Okamoto & Y. Naruta (2011). Knowledge Publishing Method for Knowledge Refinement in Self-Directed Learning. Proc. of International Conference on Computers in Education, 131-134, Chiang Mai, Thailand.

[5] A.Kashihara & M.Sakamoto & S.Hasegawa & J.Toyoda (2003). Supporting Reflection on Self-Directed Learning Process in Hyper-space, journal of Transactions of the Japanese Society for Artificial Intelligence Vol 18, 245-256.

A Course Guidance System based on Learner Characteristic Analysis: Case study of Chiangmai Rajabhat University

Phrimphrai Wongchomphu
Rangsit University, Thailand

Abstract: This study was conducted to study about Course guidance System. The main objective of the research is to develop a guiding system to match student's capability and course characteristic. Currently, we focus on Computer science, Information technology and Web programming and security. We apply ontology concept to define learner profile, and course profile which belong to subject profile in detail. Learner profile assessment and Course profile assessment are introduced to decide the characteristic of user and subject. We apply mapping module to map the characteristic between course and student profile. Finally the summarized recommended results are reported to student.

Keywords: Education Recommender System, Computer Curriculum Recommender System, Learner Characteristic Analysis ,Nearest Neighbor algorithm.

1. Introduction

Understanding the strength of individual student and managing the suitable ways for them are important tasks in Education Guidance. In general, universities provide information such as curriculum, cost, learning environment, the entrance examination range score to assist students for planning their future study with suitable course. A lot of students have selected the unsuitable course because they do not understand the nature of those selected subject. Most students used personal criteria to select course, such as feeling, preference, dream, parents or friends suggestion, accumulative GPA, popular job and so on. Nowadays, computer related field becomes popular among Thai students. There are a lot of computer related course, such as computer engineering, computer science, information technology, management information system, IT for business, and so on. Our assumption problem is "how to advice students to enter the appropriate computer related course?"

Chiang Mai Rajabhat University is a university aiming to develop a skill based students. Students in this university are not top rank students. Most of them do not know about their strengths. There are three computer's curriculum in IT, computer science, and web programming. Based on the statistics of three computer's curriculum on past four years (2008-2011), the dropout rate by regulatory measures over than 30%. Thus, the dropout rate of students by three courses are Computer science 41%, Information Technology 44% and Web programming and security 32%, respectively [4]. After we interviewed the students, most of them informed that they applied to study computer as the end user who use application program such as MS-Office, Adobe Photoshop, entertainment program such as music, video or games. They do not know that it is necessary to understand programming language as a fundamental study. In this paper, we try to develop course guidance system. At Section 2, we explains the related work. In Section 3, we illustrates the overview of system. Section 4 describes the conclusion and future work.

2. Related work

An Education purpose of recommendation system is mostly developed to make a decision and support an operations in different parts of the educational system. There are several related work on recommendation system. Nguyen Thai-Nghe and groups,[3] developed a recommended system for predicting student performance based on Collaborative Filtering and Matrix Factorization techniques and compare to the traditional regression methods that are logistic/linear regression methods. Benjaporn Lerdsakooljinda and Nattavee Utakrit.[2] developed the educational recommender system for high school students as a guide to study in higher education/ academic degree. The case study of Bangalore Mysore District, Karnataka and Delhi State of India are introduced here. They used content-based filtering technique and nearest neighbor algorithm to recommend result which has similarity value with liking and interesting of user. Our work attempts to find the most appropriate technique to map between curriculum and student profile, and recommend based on the student's skill more than the interest.

3. A Course Guidance System based on Learner Characteristic Analysis Framework

3.1 The overall system architecture

We design a course guidance system as shown in Figure 1. It consists of four modules; learner profile assessment module, course assessment module, mapping module and recommendation module. A Learner profile ontology and course profile ontology are designed for conceptual understanding of learner and course. Mapping module is applied to matching between learner profile and each subject in each course to analyze the capability of students with the interested course. Finally, the reports are shown to the students.

3.2 Assessment module

A) Profile assessment module. This module is assigned to collect learner profile data by three processes as shown in Figure 2.
B) Course assessment module. This module is designed to collect course information and to get the characteristic of each subject in the course as shown in Figure 3. In this initial step, this paper concentrated on Computer science , Information technology and Web programming and security course. After we investigate all subjects in three courses, we found that there are 102 subjects in total. After we investigated in detail, the common subject among three courses are 16. The common between CS and IT is 11 subjects;, CS and Web is 13 subjects, and IT and Web is 8 subjects. The special subject, in each course (CS, It and Web) are 6, 9 and 4 respectively.

3.3 Mapping Module

This module is assigned to measure the similarity between student profile and course profile. The process on mapping is shown in Figure 4. We calculate the accuracy based on two strategies, (1) The non-weight strategy uses the percentage of matched subject per total subject. (2) The weighted strategies uses the community of subject as a criteria for weighting the course specific, currently, we set the weight for 3-common, 2-common and non-common are 1/3, 1/2 and 1, respectively. After we normalize the weight, we get the ratio among 3-common, 2-common and non-common are 2/11, 3/11 and 6/11

3.4 Recommendation Module

This module is assigned to present result of recommendation system .The result consists of Name, characteristic, chosen courses and % of accuracy as shown in Table 1.

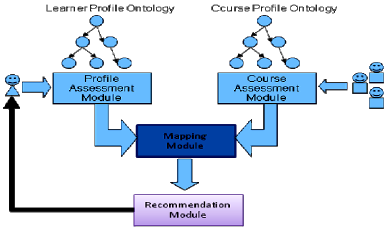


Figure 1: Overview of course guidance system.

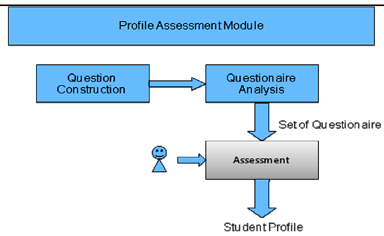


Figure 2: Process of Profile Assessment Module.

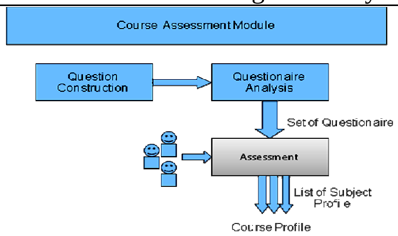


Figure 3: Process of Course Assessment Module.

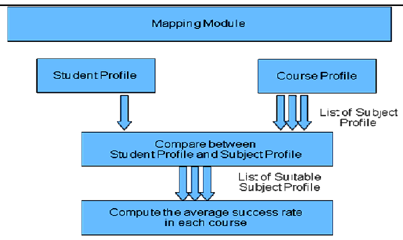


Figure 4: Process of Profile Assessment Module.

Table 1: Recommendation Result

Name	Anantra Maitree.		
Characteristic and Skill	Self-confidence, Individual, Freedom, Honest, Self-discipline, Observant, Serious, Responsibility, Carefully and Diligence, Logical Thinking, Step-by-step Thinking, Connectedness		
Expected course	**CS**		
% of accuracy	**CS**	**IT**	**Web**
Without weight	**93.48**	**81.82**	**85.37**
With weight	**88.64**	**70.31**	**55.36**
Recommendation	Based on your result, we found that the most appropriate course is _CS_. After we consider the information in detail, we found that **Anantra Maitree** tends to do the special subject well on 1 .		

4. Conclusion and Future work

In this paper, we present the concept of course guidance System. This concept consists of four modules as profile assessment module, course assessment module, mapping module and recommendation module. In the future work of this study, we plan to extend the work as follows: (1) we plan to develop a course guidance system and test in the real education environment with some selected school near university, and (2) we plan to conduct an assessment of satisfaction with the developed system.

References

[1] Thanin silpjaru. Research and analysis data with SPSS. Bangkok, SR.printing company co, Ltd ,2551.

[2] Benjaporn Lerdsakooljinda, Nattavee Utakrit , “Academic Degree Recommender System Case Study : Bangalore and Mysore District Karnataka and Delhi State of India By Content-Based Filtering Technique”, The 7th National Conference on Computing and Information Technology(NCCIT).,2011.

[3] Nguyen Thai-Nghe, Lucas Drumond, Artus Krohn-Grimberghe, Lars Schmidt-Thieme." Recommender System for Predicting Student Performance".1st Workshop on Recommender Systems for Technology Enhanced Learning., Procedia Computer Science 01,2010.

[4] Rajabhat Chiangmai University, “Report Statistic”. http://www.academic.cmru.ac.th/web55/Stat.html

Fostering Collaborators and Deep Learners through Knowledge Building among Chinese Tertiary Students

Chunlin LEI, Carol K.K. CHAN
The University of Hong Kong, SAR, China
leichl; ckkchan@hku.hk

Abstract: In this study, we designed and examined a principle-based knowledge building environment that aimed to (a) promote collaboration, deep learning and domain understanding among Chinese tertiary students, (b) characterize knowledge building discourse patterns to understand the changes. The quasi-experimental design involved two groups of students: one in knowledge building environment (KB), the other in technology enriched environment (NKB). Questionnaire survey results showed both environments students improved on conceptions of collaboration and approaches to learning, but only the KB students acquired significant improvement. Also, students in KB environment outperformed their peers in NKB environment on domain understanding. The knowledge building inquiry thread analyses indicated students in KB environment were more engaged in Knowledge Forum (KF) and able to contribute discourse with high-level knowledge advances. The results also showed a trajectory of knowledge building process, suggesting students move toward deep learning, understanding, and emergent knowledge advancement by assimilating knowledge building principles over time.

Keywords: Collaboration, learning approaches, knowledge building principles, inquiry threads

Introduction

Educational reform in China, as articulated by the *National Outline for Medium and Long-term Educational Reform and Development* (*from 2010 to 2020*), emphasizes the needs of establishing a learning society, promoting lifelong learning, and fostering high quality talents with innovative abilities. Elsewhere in the world has witnessed an emerging theme of research focusing on developing people's capacities for knowledge creation, inquiry and collaboration (Chan, 2011). Research on computer-supported collaborative learning (CSCL) has demonstrated its huge benefits to students, for example, higher-order thinking skills, active involvement in the learning process, and improved classroom results (Roberts, 2005). However, problems and challenges remain. For example, implementing innovative approaches in teaching and instruction is difficult in higher education, and simply putting students in a CSCL environment does not necessarily generate deep learning and collaboration (Hewitt, 2005; Kreijns, Kirschner, and Jochems, 2003). Research on students' learning has revealed students' beliefs about learning and approaches to learning may predict their academic performance (Biggs, 1999; Law, Chan, and Sachs, 2008). However, most of the studies are correlation studies and fewer of them have examined designing a CSCL environment that brings about changes in learning approach through instruction. Much more evidence is needed to investigate learners' beliefs on collaboration and how collaboration can be fostered and contribute to learning in complex educational settings. The goal of the study is to design and examine a principle-based CSCL environment to foster collaboration, deep learning, and domain understanding in the context of higher education.

1. Theoretical Perspectives

Over the past few decades, perspectives on learning have emphasized its distributed, social, and collective nature, and learning is often examined when mediated by the use of technology (Bereiter, 2002; Sfard, 1998; Stahl, 2006). This study adopts the knowledge building model (Scardamalia and Bereiter, 2006), a forerunner of CSCL, to examine students' changes in collaboration, learning approaches as well as domain understanding when they are engaged in a designed learning environment. Knowledge building focuses on knowledge creation as a collective work of a learning community. Ideas, which are improvable by means of discourse, are posited at its centre. To support student discourse, Knowledge Forum (KF) is designed to transform classrooms into a knowledge-building inquiry community (Bereiter, 2002; Scardamalia, 2004). KF provides "Scaffold" tools such as "My Theory", "I want to understand" and "Putting our knowledge together" to support and objectify the creation and improvement of ideas manifested in the form of notes (See Figure 1).

Scardamalia (2002) has proposed a system of 12 knowledge building principles to facilitate and examine the socio-cognitive and socio-technological dynamics of knowledge creation. Principles such as *improvable ideas, epistemic agency, collective responsibility, and constructive use of authoritative sources* are often cited by the international community working on knowledge building model to design conditions needed for both individual and collective knowledge growth. A wealth number of studies have been done in the investigation of knowledge building classroom practices (Chan, 2011; Lee et al., 2006; van Aalst and Chan, 2007; van Aalst, 2011; Zhang et al., 2007).These studies revealed the effectiveness of implementing knowledge building model in literacy development and depth of inquiry. However, many previous studies were done in schools and in discipline of science or science-related fields. Fewer empirical studies on knowledge building inquiry were done in the domain of business studies at the tertiary level against the Chinese background. In ijCSCL's first paper from Mainland China, Ge (2011) reported tertiary students major in engineering in Beijing disliked collaboration despite the fact that Chinese people have been traditionally ranked high in the collectivist cultural dimension (Hofstede, 1980). Ge's study also indicated students' strong reliance on teachers even they had been put in a CSCL environment for as long as a school year. This gave rise to the issue as how instruction could be designed to turn agency over to the students and to facilitate collaboration.

Biggs (1999) points out that the concepts of surface and deep approaches to learning are very helpful in conceiving ways to improve teaching and learning. An important goal of twenty-first century education is to cultivate deep learners. Many studies about learning approach have been conducted in traditional classrooms. It is interesting to see how student learning approach may or may not change when they are engaged in CSCL environments. Few studies in the past have linked students' conception of collaboration with student approaches to learning and domain understanding, especially in a principle-based learning environment. Researchers have differentiated students' inquiry discourse on KF as knowledge sharing, knowledge construction and knowledge creation (van Aalst, 2009); it is meaningful to discern how they are manifested among students and how they may be related to students' conceptions of learning, collaboration and domain understanding.

Therefore, two research questions are included: (1) what are the effects of the designed learning environment on students' conceptions of collaboration, learning as well as domain understanding? (2) what characterize students' knowledge building discourse and how they may be related to students' changes in the learning environments?

2. Method and Design

2.1 Participants and procedures

The participants were two intact classes of first year university students in a good business-and-economics-oriented university in Shanghai, China. A quasi-experimental design examined the effects of a principle-based knowledge building environment (KB, n=30) and a non knowledge building environment (NKB, n=30). The KB environment was characterized by KB theory, pedagogy and principles, coupled with KF, the technology platform. The NKB environment was a typical teaching environment which was composed of teacher's lecture and students' discussion; there was no KB principles governed, however, KF was also used due to the school policy for comparable educational experience. More importantly, by adopting KF to both classes, we could examine closely whether it was merely the inclusion of technology or the deeper impacts via knowledge building principle-pedagogy-technology that might contribute to changes. Both classes were taught by an experienced instructor who had taught in higher education for 12 years and observed KB teachers' meetings frequently and accessed a large amount of KB discourse. The course in this study was titled Introduction to Business, which was conducted in two semesters of academic year 2010-11. Each semester had 12 teaching weeks and each week had two lessons, which was 1.5 hours in length. English was used as the medium of instruction. Students had face-to-face discussion and inquiry in class and wrote computer notes on KF after class. For the KB environment students, the above-mentioned KB principles were explicitly mediated, for example, they were asked to give presentations regularly to track, reflect, and think about ways to improve their online discourse; For NKB class, students were encouraged to use KF as a new technological platform to communicate and learn; no explicit KB principles were introduced to them.

2.2 Design of a knowledge building environment

To foster collaboration and deep learning, we designed a leaning environment based on knowledge building pedagogy. (1) *Cultivating a collaborative learning culture*. The first few sessions provided the students with learning experiences which familiarized them with the technology and acculturated them into the practices of collaborative inquiry. A focus was placed on making ideas public on KF *view* (Figure 1) and assuring them psychologically safe in contributing ideas to KF. (2) *Developing knowledge-building inquiry*. The course included some big, core ideas, which were tailored into a progressive curriculum (Caswell and Bielaczyc, 2001). Student through face-to-face and online discourse, elaborated what they know about the topics, set forth their theories, and explored the answers in a cyclical way. (3) *Improvable ideas and emergence*. Students worked continuously to improve the quality, coherence and utility of ideas (Scardamalia, 2002). Agency was turned over to them to create, revise and refine ideas. As inquiry went deeper, they even self-defined goals and activities for solving emergent problems, and the teacher acted as a facilitator and co-inquirer. (4) *Formative assessment*. Concurrent formative assessment played an important role in the design. Students were involved in classroom reflective presentations and group electronic-portfolio assessment during semester 1 and 2 respectively, which were used to scaffold learning, collaboration as well as characterize collective knowledge advancement.

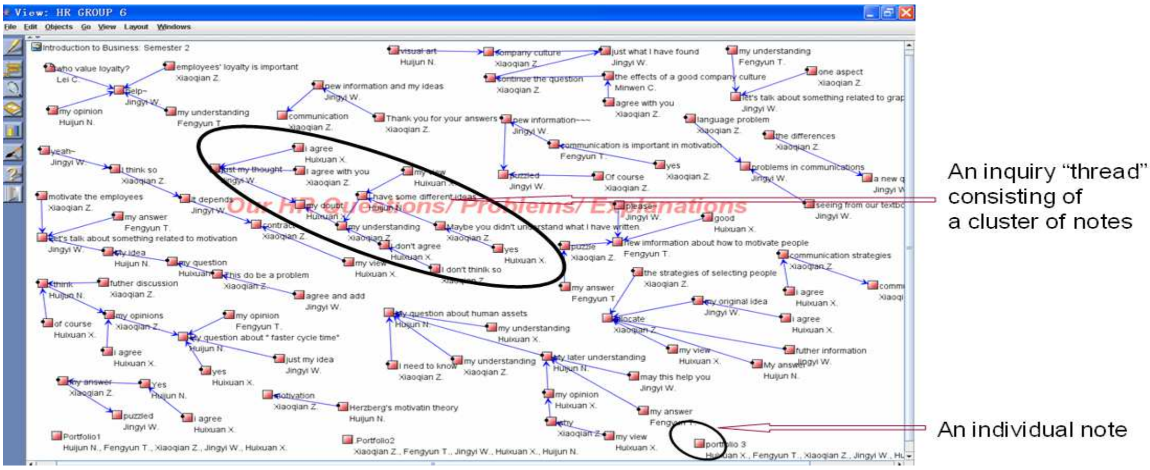


Figure 1. A knowledge forum *view*, including individual notes and inquiry threads

2.3 Data source

We collected multiple-source data for the whole study and included only part of them to answer the above research questions. Two questionnaires, investigating students' conceptions of collaboration and approaches to learning, were done in a pre- post fashion. The questionnaire on collaboration was developed by Chan & Chan (2011), which consists of 12 items and uses 5-point Likert scale examining students' views of collaboration aligned with the notion of knowledge building. The questionnaire on approaches to learning was based on Biggs' (2001) Revised Version of Study Process Questionnaire (SPQ), which is comprised of 20 items, with 5-point Likert scale and two-factor structure of deep and surface approaches to learning. Students' domain test papers were collected to examine their understanding of business concepts. The test paper was comprised of five core, open-ended questions asking students to explain their understanding of business concepts (for example, Do you think Shanghai has a good business environment now? Why?) It was done in a pre- post manner and one teacher blind marked all the test papers according to a marking scheme jointly developed by the teacher and the course leader in the university. Students' KF discourse was included. In particular, we classified their Semester 1 inquiry discourse into threads (Figure 1). An inquiry thread was defined as a cluster of notes addressing the same principle topic or problem in the communal space (Zhang et al. 2007). Thread analyses helped to characterize knowledge building dynamics and track collective knowledge growth.

3. Analyses and Results

3.1 Changes in conceptions of collaboration

The Questionnaire on Collaboration required students to reflect their experiences in their learning environments. The pre- and post Cronbach's Alpha were .84 and .85 respectively (n=60), indicating good scale reliability. The pre-, post scores for students in the KB environment were 3.87 (.50) and 4.23 (.43), and paired t-test showed students improved their conceptions of collaboration significantly with $t\ (29) = 6.14$, $p<.001$, *Cohen's* $d =.72$. While in the NKB environment, the students' pre-, post scores were 3.98 (.51) and 4.16 (.51) respectively, and paired t-test revealed that although student obtained higher scores at the end of the program, their improvement on conception of collaboration was not statistically significant with *t (29) = 1.53, P=.14*.

3.2 Changes in approaches to learning

Students' responses to SPQ were rated and the pre-, post Cronbach's Alpha for deep approach were .76 and .78 (n=60); and .82 and .87 (n=60) for surface approach, which were consistent with previous studies and indicated good scale reliability. In the KB environment, the pre- post scores of deep approach were 3.35 (.53) and 3.73 (.42); and 2.18 (.58) and 1.91 (.49) for surface approach. Paired t-tests showed that students in KB environment improved both on deep and surface approaches (improvement on surface approach means decreased score at the end of instruction). For deep approach, *t (29) = 4.47, P<.001, Cohen's d =.71*, indicating a significant level of improvement. For surface approach, *t (29) = -2.38, P<.03, Cohen's d =.47,* suggesting fair improvement at a certain level. In the NKB environment, the pre- post scores were 3.25 (.51) and 3.40 (.60) for deep approach; and 2.22 (.62) and 2.14 (.71) for surface approach. However, analyses of pre- post scores using paired t-tests showed that students in NKB environment had not obtained statistically significant improvement, for deep approach, *t (29) = 1.65,P=.46*; and for surface approach, *t (29)= -.75, P=.11*.

3.3 Improvement on domain understanding

Domain test paper was designed to tap into students' understanding of key business concepts. The pretest was given at the beginning of the instruction and the posttest was done at the end of semester 1. The pre-, post scores were 28.9 (6.65) and 60.3 (8.27) for KB class; and 28.0 (8.59) and 52.4 (7.89) for NKB class, respectively. Analyses of pre-, posttests using paired t-tests showed both classes improved on domain understanding tests with KB class, *t (29), = 23.5, p<.001*; and NKB class, *t(29), = 14.1, p<.001*. Further repeated measures (Pre-, post x KB, NKB) were conducted and the results indicated the Environment x Time interaction was significant, *Wilks' λ =.85, F (1, 58) = 10.3, p<.005,* η^2 *=.15,* favoring the KB Environment students over their NKB counterparts, *F (1, 58)* = 6.6, *p<.05*, η^2 *=.10*.

3.4 Characterize knowledge building discourse

During semester 1 both KB and NKB classes wrote notes on a big curriculum topic "Business Environment", which was subdivided into six *views* on KF, namely, "What is business", "Egg theory", "Political environment", "Economic environment", "Social Environment" and "Technological Environment". Notes of the six views were retrieved and we coded them into inquiry threads, for example, KB class students wrote 30 notes and NKB class students wrote 18 notes respectively to discuss "definition of business", thus constituting a thread titled "what is business". Ultimately, we identified 57 discussion/inquiry threads from KB class and 29 from NKB class. We then further analyzed these threads in terms of whether it shows continual testing and modifications of ideas; uses authoritative sources to build-on ideas constructively; and demonstrates communal awareness and efforts in advancing collective knowledge. Rated by the level of knowledge advances, three discourse patterns emerged: *(1) Low-level knowledge advances thread* (LKA) which was usually not long, consisting students' quick or naïve ideas and lacking a well treatment of a topic. *(2) Moderate-level knowledge advances thread* (MKA) in which students were able to answer questions from various perspectives; a pool of business ideas were accumulated yet repetitions occurred frequently showing a inclination of task-completion and a lack of community awareness. *(3) High-level knowledge advances* *thread* (HKA) in which students were engaged in a question-explanation-intertwined process, identified and addressed gaps in collective knowledge, negotiated meaning and formulated more sophisticated views toward business theories or concepts. Table 1 and 2 showed examples of LKA and HKA respectively.

Table 1: An example of LKA thread

Student	Description	Excepts from the inquiry thread
FXX	Question	I need to understand How to make a market survey and meet customers' need?
NHJ	Reasonable answer	My theory We can give some leaflets to different people…, listen (to) their ideas. It will tell us what customers need. The leaflet can contain price …advice to the product…
JYH	Digression	My theory It's too early to think this now. What we should do first is learning more about business so that we are able to solve problems…
FXX	Disagree	My theory But I think we should consider questions from all aspects. Also, it's necessary…
JYH	Discourage	My theory But you don't have any professional knowledge to consider questions from all aspects. For example …
SL	Elaborated answer; but ended here	My theory Maybe, firstly you should know … problem to solve. Then you … design a series of… After that, you organize …give out to people of different ages. Finally, you analyze …

Table 2: An example of HKA thread

Student	Description	A few excepts from the inquiry thread
ZXQ	Start with puzzle	I need to understand *PUZZLE*….Foreign countries say that we have the biggest market…we say the population is burden. How do we comment on the population in China?
YSH	Explanation	My theory ... More people there are, larger the market will be, (more) business opportunity...the mobility of population …Population re-aggregation will lead to …
CXF	Elaborated more	My theory More population means…However …the more …, the bigger social pressure. If…, people's living standards will…
ZXQ	Ask for explanation	I need to understand *Puzzle again:* I know everything has two aspects. But …which one is more important …Because…
WQY	Authoritative information	My theory Disadvantage outweighs…Relevant data (from the Internet) showed …a series of problems such as … Due to …it's hard to …we should consider questions from all aspects…
WJY	Strive for coherence of ideas	My theory As we all know…However, behind is the serious population problem. …more negative effects … Firstly, …challenge to resourses. Secondly, …harm to environment …Finally, …burden to governmet and society, such as …
ZCY	Relevant problem	I need to understand …in recent years, we have a problem of aging..
CMW	Explanation	My theory …the birth rates…fewer babies…old people live longer … problem of aging appear(s) … In the long run, it will …
YSH	See the complexity	My theory The aging issue is thorny …human's life expectency is …birth rate…contradict…unable…
SYM	Idea diversity	My theory It seems …But …only one of the aspects. Since the trend of aging…why not seize the opportunity? In fact, it can be taken advantage …instead of…
CMW	Collective responsibility for solution	My theory First, as many calssmates mentioned, we can resort to the postponing year of retirement. Many people … As far as I'm concerned, … new ideas…Furthermore …solve the aging problems…
FTY	Summarize	Putting our knowledge togehter 1 my understanding of your question

	and rise above based on 9 community notes	2 My anwer 3 population 4 My opinion 5 The issue of population in China 6 need high quality population 7 It will take a long term...8 Why the logo 'Made in China' can't become a logo 'Creat in China' 9 my thoughts First, …confused about the question. …involved many aspects…not a clear direction. Secondly,…Thirdly, …large population ..manpower…employment pressure. Then, China… Despite … After that …Finally, it comes to…However, the government …In conclusion, the population is a problem … ways to solve it …

Note: Notes included here are shortened and selected to highlight the key features only.

We rated all the threads in both KB and NKB classes and the results were shown by Figure 2. It indicated KB students were more productive both quantitatively and qualitatively than their peers in NKB environment. For example, the KB class generated 26 HKA threads (28.1% of total threads generated) as compared to 5 (17.2%) in the NKB class.

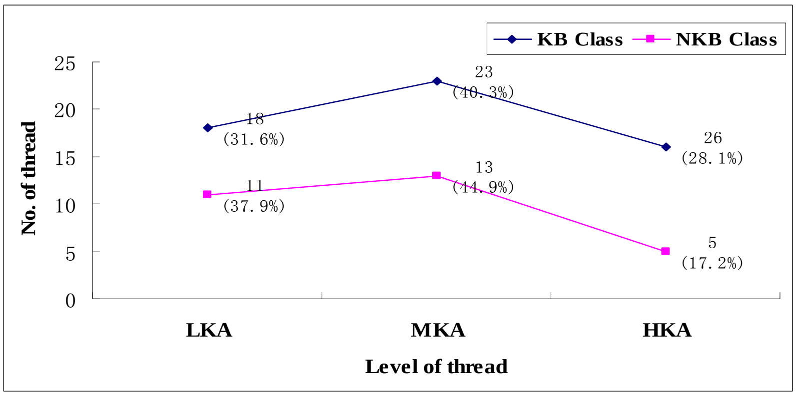


Figure 2: Comparison of threads between KB and NKB classes

Using the forum log information (judging by the ending date of each thread), we classified the threads of KB class into three periods of time, namely Period 1 (week 4), Period 2 (week 8) and Period 3 (week 12 and beyond). As shown by Figure 3, HKA threads increased from period 1(1 and 7.7% in total threads of period 1) to period 3 (7 and 38.9% of total threads of period 3) steadily and LKA decreased across time (from 7, 53.8% of periods 1 to 1, 5.6% of period 3). When students were more immersed into knowledge building culture and principles, they seemed able to refine their discourse toward emergent knowledge creation.

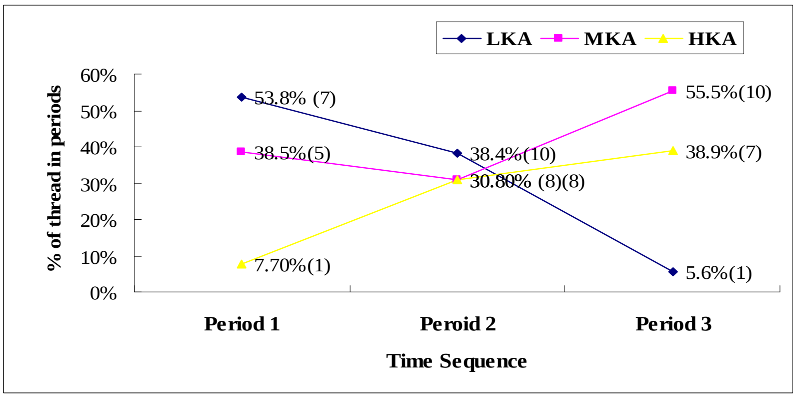


Figure 3: Levels of thread across 3 Periods of time in KB class

4. Conclusion and implications

This study has designed and examined a principled-based knowledge building environment in promoting learning, collaboration and domain understanding among Chinese tertiary students. Following a quasi-experimental design, the study showed both KB environment and NKB (merely-technology-supported) environment students improved their conceptions of collaboration and manifested a tendency of moving to deep learning. However, the KB environment showed salient advantage over the NKB environment in facilitating collaborators and deep learners. The domain test results also suggested both classes strengthened understanding of business concepts, with KB students again outperformed their peers in the NKB environment.

We examined students' knowledge building discourse on KF. Inquiry thread analyses revealed that KB students were actively engaged in the knowledge space by raising authentic problems, using resources constructively, striving for an intertwined question-explanation-based idea improvement, and moving toward high-level meta-discourse. In fact, online forum is now very popular in China and else where in the world, however, many discussion threads have found to be merely chit-chat or on a surface level of knowledge sharing, just like the products of the NKB environment in this study. In the KB environment, we identified a large number of sophisticated discourse patterns which indicated KB students, working in a community, have embarked on a trajectory of progressive problem solving, idea improvement, and moving to emergent knowledge advances. We acknowledge the limitation that there might be some teacher effects contributing to the differences between the two environments. However, we argue that technology needs to be informed with pedagogy and collective views of knowledge building. Only when students have been acculturated in the designed environment and assimilated the principles appropriately can deep learning, meaningful collaboration, and high-level knowledge advances take place.

References

[1] Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71, 148-149.

[2] Caswell, B., & Bielaczyc, K. (2001). Knowledge Forum: Altering the relationship between students and scientific knowledge. *Education, Communication & Information*, 1(3), 281-305

[3] Chan, C. (2011). Bridging research and practice: Implementing and sustaining knowledge building in Hong Kong classrooms. *Journal of Computer-Supported Collaborative Learning*, 6(2), 147.

[4] Chan, C.K.K., & Chan, Y.Y. (2011). Students' views of collaboration and their online forum participation in Knowledge Forum. *Computers & Education*, 57, 1445-1457.

[5] Ge, Z. (2011). Exploring e-learners' perceptions of net-based peer-reviewed English writing. *Journal of Computer-Supported Collaborative Learning*, 6(1), 75-91

[6] Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *The Journal of the Learning Sciences*, 14(4), 567-589.

[7] Lee, E. Y. C., C. K. K. Chan, et al. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-Supported Collaborative Learning* 1: 278-307.

[8] Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal Education in a knowledge society* (pp. 67-98).

[9] Scardamalia, M., & Bereiter, C. (2006). Knowledge building. In R.K.Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97-115). Cambridge: Cambridge University Press.

[10] van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge construction, and knowledge-creation discourses. *Journal of Computer-Supported Collaborative Learning* (4), 259-287.

[11] van Aalst, J., & Chan, C. (2007). Student-directed assessment of knowledge building using electronic portfolios in Knowledge Forum. *The Journal of the Learning Sciences*, 16(2), 175-220.

[12] Zhang, J. W., M. Scardamalia, et al. (2007). Socio-cognitive dynamics of knowledge building in the work of 9 and 10-year-olds. *Education Tech Research Dev*. 55: 117-145.

Creating multimodal texts in language education – an emerging practice at the boundary

Anna-Lena GODHE[a*] & Berner LINDSTRÖM[b]
[a]*Department of Applied Information Technology, University of Gothenburg, Sweden*
[b]*Department of Applied Information Technology, University of Gothenburg, Sweden*
*godhe@chalmers.se

Abstract: In this paper the emerging practice of making multimodal texts within language education is explored, in order to illuminate a boundary in which students negotiate what to include in a digital story produced in an educational setting. The making of multimodal texts originates in contexts outside of education and, accordingly, influences from these practices are incorporated into the activity system of making texts in language education. Through excerpts from interactions between students and their teacher, what is considered to belong, or not belong, in a school context, is illuminated as the negotiations between actors establish what are considered appropriate and meaningful actions. The digital stories created by the students within the activity system of schooling become a hybrid as influences from intersecting activity systems are incorporated into the multimodal texts they create.

Keywords: Multimodal text, boundary, emerging practice, activity theory

Introduction

To write and express oneself through language is an important part of language education. By using digital technology, it is possible today to create multimodal texts where pupils are able to use different modes, such as images, speech and music, to express themselves, which potentially provides for new practices of reading, producing and disseminating texts [8, p29]. Bergman [2] and Olin-Scheller [12] studies concern language education in Swedish schools, and both show that the connection is weak between texts that students consume and produce outside of school and the ones they encounter in school. Creating multimodal texts in language education could be a way of bridging the gap between the different worlds in which students seem to live, and may enable the students to make use of abilities connected to the use of technologies in out-of-school practices.

Using computers outside of school is common to most teenagers in Sweden [11], so it is likely that the pupils are accustomed to using technologies as mediating artifacts in their spare time. Technologies, as mediating tools, impact the way in which learning is mediated as well as the potential practices available for those who use them. Since one-to-one-solutions, where students have individual laptop computers, are becoming increasingly common in schools in Sweden, it is of interest to study emerging practices related to these technologies. As institutionalised habits play a significant role in how a certain technology is used, an important empirical question within educational research, is how individuals are acting with technology in educational settings.

When creating a multimodal story, the students are working with tools related to contexts outside of schooling, which may accommodate the interaction of habits and abilities connected to practices both inside and outside of the school environment. Considering that many youngsters, especially boys[i], watch film clips on sites such as YouTube, it is probable that the students are used to watching, and maybe also producing, home-made short films on the Internet. When the object of an activity in a school context is a multimodal text, more commonly encountered in practices outside of school, this object can be seen as a boundary object which inhabits intersecting contexts.

In the following paper the interaction between two Swedish upper-secondary students, while they make a multimodal text in the shape of a digital story, is analysed to explore how the students create their multimodal story by relating to several references from out-of-school practices. The question raised concerns the reasons as to why some references are incorporated in the multimodal text while others are brought up in the interaction but left out in the digital story they create.

1. Theoretical Framework

Activity theory is used in the analysis as a theoretical framework, applying the activity system, as described by Engeström [5], to the context of schooling in general and to the activity of creating multimodal texts in particular. In what Engeström calls the third generation of activity theory [5, p56], the basic model of activity depicted as a triangle where subject, object, rules, community and division of labour interact in various ways, has been expanded to include at least two interacting activity systems (Figure 1). When several activity systems are involved, the object becomes potentially shared and can then be seen as a boundary object. According to Star & Griesmer [14], a boundary object has different meanings in different social worlds, but because the structure of the object is common enough it is recognized and may hence be a key factor in developing coherence between intersecting social worlds.

Engeström et al [6] sees the transportation of ideas, concepts and instruments from different domains as boundary crossing. Crossing boundaries involves stepping into unfamiliar domains which, in turn, requires the formation of new conceptual resources. Boundary crossing calls for a horizontal expertise where movement across boundaries is necessary. When learning is considered to be a vertical movement where the expert teaches the novice, such horizontal movements are largely ignored.

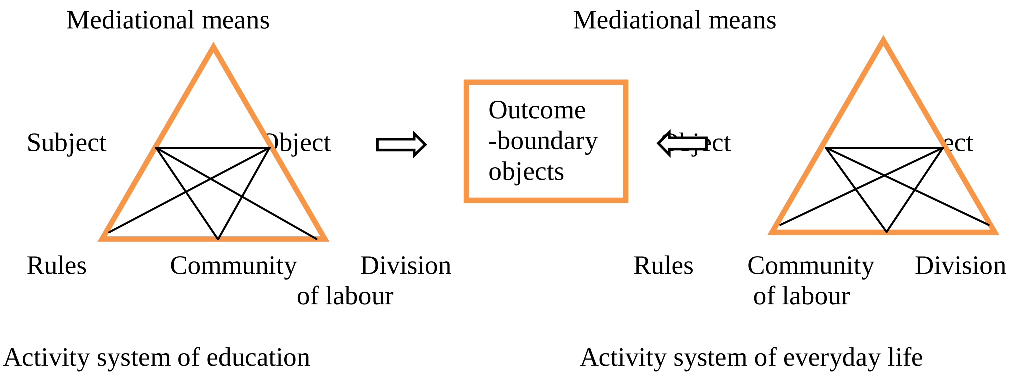


Figure 1. Interacting activity systems and outcomes as boundary objects

Since creating multimodal texts in language education is an emerging practice, it is still not embedded in its own rules and expectations of how it should or could be done. As such it may be compared to what Engeström & Sannino [7, p.2] calls expansive learning, where the learners are "involved in constructing and implementing a radically new, wider and more complex object and concept of their activity". When the students create their multimodal text they partake in a new form of activity in the language classroom, where they simultaneously create and learn what this activity entails. As the outcome of the activity is, to some extent, unknown to everyone there are no experts in such learning processes.

In a review of the literature on boundary crossing and boundary objects, Akkerman & Bakker [1, p.141] concludes that descriptions of boundaries and of people and objects at the boundary show signs of ambiguity as in-between and belonging to both one world and another. This means that boundaries connect as well as divide the activity systems involved. People at the boundary act as bridges between the related worlds, but simultaneously also represent the division between them. Akkerman & Bakker [1] contend that it is because of their ambiguous nature that boundaries have become a phenomenon which is investigated in relation to education.

> "Both the enactment of multivoicedness (both-and) and the unspecified quality (neither-nor) of boundaries crate a need for dialogue, in which meanings have to be negotiated and from which something new may emerge." [1, p.142]

In the literature about boundary crossing and boundary objects, Akkerman & Bakker [1] discern four learning mechanisms one of them being transformation. Transformation involves confrontation and continuous work which leads to profound changes in practice where an in-between practice, or boundary practice, may be created. They see hybridisation, where "ingredients from different contexts are combined into something new and unfamiliar" [1, p.148] as one of the processes involved in transformation. When practices cross boundaries and engage in a creative process, a hybrid emerges.

Depicting activity systems as neat triangles may contribute to a sense that activity systems are stable and harmonious, but instead, Engeström [4, p.72] states that they are characterised by contradictions. It is through tracing troubles and innovations in an activity system that development can be understood. When a component in the activity system acquires a new quality due to influences from intersecting activity systems, secondary contradictions arise between that component and others in the system [4]. The use of computers and other ways of expression than the typographical word, when creating texts in a school setting, means that the mediational means, as well as the object of the activity, acquires new qualities. This in turn leads to secondary contradictions within the activity system of education.

Though the two triangles in the figure are the same size and it looks as though their influence on the outcome is equal, this is not the case when studying the creation of a multimodal text in a school setting. The students are doing a school task and therefore the activity system of education dominates their activity. Although there is a possibility of the intersecting activity system influencing the dominant one, influences from intersecting activity systems will need to be adjusted to the dominant activity system of education. The activity systems in focus here are, of course, part of, and influenced by, other activity systems and also incorporated into the overarching system, which they in turn have to relate and adjust to. It is, however, what is negotiated in a classroom which is in focus here and how certain aspects of a multimodal text are negotiated by students. In a language classroom there are established literacy practices such as reading and writing typographical texts. When making a multimodal text the object of the activity is expanded and incorporates several ways of meaning-making. Drawing on practices both inside and outside of the classroom enables an alteration of the literacy object as well as the practice, which may be expanded but also constrained, as the emerging practice has to relate to the established literacy practice in a language classroom.

2. Method

The study was conducted in an upper-secondary school in southern Sweden in a class where the students had individual laptops. The students were video-recorded when they made their multimodal text. Two cameras were used where one focused on the students and the other on the computer screen. This was done in order to capture the students´ movements as well as their talk and what they did on the computer during the interaction.

Interaction analysis was used when analysing the recordings and when choosing which excerpt to present for further analysis. Jordan & Henderson [9] describe interaction analysis as “an interdisciplinary method for the empirical investigation of the interaction of human beings with each other and with objects in their environment” [9, p.39]. The goal of interaction analysis is to identify regularities in how the participants make use of the resources available to them in that particular situation. It aims to ground the analysis in the empirical material, thereby avoiding ungrounded speculations of what people may think.

Studying the interactions in a classroom is important in order to illuminate what is negotiated and how the local practice through negotiations are made relevant in relation to the task of making a multimodal text. The analysis of excerpts from interactions is, in this study, used to show what is being negotiated as well as what aspects in the negotiations that are made relevant in this particular setting. Linell [10] writes about double dialogicality which means that a situation is in dialogue both with the immediate participants in that interaction but also with the context within which it is set. What is negotiated in a particular situation may thus serve as an example not only of that situated practice, but also of the plausible negotiations within that particular sociocultural praxis.

The analysis is based on interactions between two students which I will here call Isak and Jonas. They are both native speakers of Swedish. They were given an assignment in their course in Swedish at upper-secondary school to alter an existing scene in a book or create a new scene. The assignment was connected to the reading of a book, “Låt den rätte komma in” (“Let the right one in”). Isak and Jonas decide to re-make a scene where a character is attacked by a vampire in a forest.

3. Empirical Findings

In an educational context there are established literacy practices where students write texts with pen and paper or create the text in a word processing program on the computer. When creating a multimodal text in the shape of a digital story, however, both the object of the activity and the mediating artifacts have been altered so that the literacy practices involved as well as the literacy object which the students are expected to create is different from the established literacy practice. While the students also use other mediating artifacts, the computer is vital in the production of the multimodal text. The alteration of mediational means as well as the outcome of the activity could be seen to create a new activity and an emerging literacy practice.

When a new or emerging practice is introduced, the rules and divisions of labour will have to be negotiated in order for the participants to establish what is, and is not applicable in the new practice. The students are hence negotiating what making a multimodal text in school means and they seek clarification as to which rules apply, but they also seek approval once they have started creating their text, to make sure that what they have done is what is expected of them.

In the following section I will present excerpts from the interaction between students where they negotiate what to include in their story. By analysing the interaction between the students and how references to practices not directly related to the school practice, are talked about among peers and with the teacher, I will explore how the students negotiate references at the boundary. As objects at the boundary shows signs of ambiguity they need to be negotiated in order to decide whether they belong to both worlds or not. Through negotiations the students decide which references to include in the object of their activity, the digital story. Some references are, however, left out in their digital story and hence appear to be considered as not belonging to the educational literacy practice.

3.1 Choosing the music

During the first lesson Jonas in particular is preoccupied with deciding which music to use in their digital story. Even though they have not yet decided which scene to create or re-create he considers finding the right music most important.

Excerpt 1, lesson 1[ii]

Isak	But we have to choose a scene first
Jonas	Are you kidding the music is important (.) we only need the music

Jonas here clearly states that what is most important to him is finding appropriate music. By saying that *"we only need music"*, he also questions or chooses to ignore the teacher´s instruction which was to make a film consisting of their voiceover, images and, if they want, also music.

Whilst searching for the music Jonas has in mind, both students suggest using the soundtrack from different contemporary movies but do not come to an agreement to use any of the suggestions made. When Jonas finds the music he has been looking for they quickly agree to use it.

Excerpt 2, lesson 1

Jonas	It´s this one
Isak	That´s good that´s good that´s really good (.) that´s awesome (.) and then boom comes the vampire (2) that was really good (.) yeah but imagine sitting in the tree
Jonas	I hope you have heard it before ((continues to listen to the music))
Isak	It´s perfect (.) and then boom

Isak immediately starts associating the music to what is going to happen in their story when he illustrates the emerging of the vampire with the word "boom" twice. He also enacts their story using his voice and movements, mainly with his hands. Here he uses his movements and the word "boom" to mark where something will happen in the story and how this will coincide with the music.

The music, Requiem for a dream [13], which Jonas has been looking for and which they end up using as the soundtrack to their digital story was originally made as a soundtrack to a movie but has featured in trailers for several well-known movies. It has also been used in video games, TV shows and adverts. The students are likely to have heard this piece of music in various circumstances since it has featured repeatedly in popular culture. This explains Jonas remark to Isak that he hopes he has heard it before, but by making the remark Jonas also appears to be in doubt as to whether Isak is familiar with the music or not. If Isak is not familiar with the music this would indicate that he is not familiar with the popular culture in which it has occurred.

3.2 Using a YouTube clip as soundtrack

In this interaction Isak and Jonas are referring to a film clip on YouTube which according to Wikipedia is the most viewed YouTube video which is not a professional music video, *Charlie bit my finger* [3]. They are contemplating using the boy in the video and what he says as a soundtrack to a certain part of their story.

Excerpt 3, lesson 1

Jonas	Can´t we have that one
Isak	Should we have that one but it will sound really lame ((laughter)) a vampire aw she bit me man ((they listen to the YouTube video)) yeah that one you could have ((continues to listen to the video where the boy talks and screams)) that would work anyway but you have to cut out Charlie then ((the boy screams again)) yeah that and then when he screams (1) that had been really cool (2) when she jumps down if you imagine when she lands ((shows with his hands)) and then the scream she screams
Jonas	Shit it will be so lame ((laughs))

Just before this conversation they have watched the video clip they are referring to and in a humorous manner talked about using it. When Jonas opens up the interaction he also replays the video clip and, as Isak starts talking about how they could use it, he is laughing. As Isak goes on talking Jonas stops laughing and turns toward Isak. Although the conversation still involves a lot of laughter, Isak´s consideration of how to use the video-clip seem to make the suggestion more realistic to Jonas. Even though Jonas earlier has been keen to use the video clip in their film, when Isak now is considering it more seriously he seem to question using it. At the end of the excerpt they both start laughing and then Jonas has some problems with the computer and they do not seriously talk about the video clip again. The sound is not used in their final digital story.

When speaking about and listening to the video clip other students in the classroom are heard to pick-up and imitate the sounds. Other students, hence, indicate that this video clip is something which they are familiar with and can associate to. When discussing whether to use the video clip or not the students do not involve the teacher in their interaction.

3.3 Bloopers

When the students are putting together the different modes in their movie, they start to talk about including what they refer to as bloopers, in the film. Bloopers are short sequences of film where mistakes are made. These scenes are usually deleted but are in some movies shown with the closing credits.

It is not entirely clear in their interaction what the students mean when they talk about bloopers. As they have not been filming they do not have any deleted scenes to add. They do however have the recordings of their voices when they created the voiceover to their digital story and it seems to be parts of these audio files which they intend to use as bloopers. Isak did the talking and he had some problems pronouncing a certain word and ended up overemphasising the last letter which was a T ("medvetslöst").

The students speak extensively about making and including bloopers but they run out of time and therefore say that they will have to leave the idea of using bloopers. However, by adding a big red >T< to their film at the point where Isak overemphasises the letter T, they include their own kind of blooper.

Though both students seem to be pleased with including the >T<, Isak says that they will get into trouble for adding it. He wants it to be green, instead of red, so that it won't be as noticeable, but Jonas does not agree with him. They later explain why they have drawn a red >T< to the teacher and it does not cause any trouble. Although showing their version of a blooper to the teacher they do not actually mention bloopers in their interaction with her. They do, however, mention their intention of adding bloopers to their classmates several times and they also discuss with the students sitting closest to them what the >T< that they add should look like.

4. Conclusion

Since the outcome of the assignment to make a multimodal text is a short film, it is not surprising that the students in their interaction relate their work to popular culture in general and movies in particular. In choosing which music to use they contemplate several alternatives, all of which are soundtracks to contemporary movies. Their final choice is a piece of music which, although originally a soundtrack to a movie, has figured repeatedly in TV as well as in video games. The multimodal text can thus be seen to enable the students to connect different social worlds as they incorporate experiences from movies, TV and video games in the making of their school assignment.

Adding bloopers to their digital story also relates to movies, where bloopers are sometimes added, particularly during closing credits. The discussion of whether to include bloopers, reveal an uncertainty of whether the inclusion of bloopers will be accepted in a school context or not. The fact that they do not mention bloopers to their teacher but talk about it with their classmates, together with their hesitation of adding bloopers to their digital story, suggests an uncertainty as to whether the use of bloopers belongs to an educational setting, or not. In the end Isak and Jonas make their own version of a blooper which they include in their multimodal text. When they show their complete film to the teacher they also explain the meaning of the blooper. A YouTube video is another feature which the students contemplate using in their digital story but which they do not use. The YouTube video is thus not considered to belong in a school setting, but bloopers are seen as belonging to both worlds.

When negotiating what to include in the multimodal text the students elucidate a boundary where some references are included in the text and, hence, are considered to belong to both worlds, while others are excluded and thus considered as not belonging to a school context. The negotiations concerning these influences show signs of the ambiguity related to the boundary where the uncertainty of whether references are both-and or neither-nor gives rise to negotiations. The students perceive some experiences from other social worlds as intersecting with the educational setting and hence they can be referred to in this setting. However, other references they refer to when interacting with each other and their peers but they are neither referred to when interacting with the teacher, nor included in the digital story which they create. Through their actions the students illuminate a boundary between what is possible to include in a multimodal text in an educational context and what is not. In so doing, the students are acting as bridges between worlds but they simultaneously represent the division between these related worlds.

By connecting what they do in the classroom to activities outside of school the students cross boundaries of different domains and use ideas and experiences from one activity system in another. The outcome of their activity, the digital story, is a hybrid as it contains elements from different contexts and a boundary object as it both bridges and divides the activity systems. The literacy object they are assigned to create contains ways of meaning-making which is not usually part of literacy practices in language education. As the literacy object is a school task it primarily relates to established literacy practices in language education and thus influences from other practices are incorporated or rejected based on the students conceptions of what a text created in a classroom may contain. By including some references the literacy object is expanded, but it is also constrained as some references are left out. The development of the multimodal text is characterised by tensions between, and negotiations about, what to include in the multimodal text and what to reject. These tensions elucidate the students´ awareness of boundaries between different literacy practices. The multimodal text, on a general level, acts as a boundary object bridging the different social worlds in which it is recognisable. As the multimodal text which the students create during their lessons in Swedish incorporates influences from different contexts it also becomes a hybrid which is new and unfamiliar in settings both in and outside of the classroom.

References

[1] Akkerman, S. F., & Bakker, A. (2011). Boundary Crossing and Boundary Objects. Review of Educational Reserach, 81(2), 132-169.

[2] Bergman, L. (2007). *Gymnasieskolans svenskämnen : en studie av svenskundervisningen i fyra gymnasieklasser*. Malmö: Malmö högskola, Lärarutbildningen.

[3] Charlie bit my finger (2011). Retrieved 18 okt from; //en.wikipedia.org/w/index.php?title=Charlie_Bit_My_Finger&oldid=455848676

[4] Engeström, Y. (1993). Developmental studies of work as a testbench of activity theory: The case of primary care medical practice. In S. Chaiklin & J. Lave (Eds.), *Understanding practice: perspectives on activity and context* (pp. 64-103). Cambridge: Cambridge Univ. Press.

[5] Engeström, Y. (2009). Expansive learning: toward an activity-theoretical reconceptualization. In: Illeris, K. (ed.) *Contemporary theories of learning: learning theorists in their own words.* New York: Routledge. P. 53.73.

[6] Engeström, Y., Engeström, R., & Kärkkäinen, M. (1995). Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities. *Learning and Instruction, 5*(4), 319-336.

[7] Engeström, Yrjö, and Annalisa Sannino. "Studies of Expansive Learning: Foundations, Findings and Future Challenges." *Educational Research Review* 5, no. 1 (2010): 1-24.

[8] Jewitt, C. (2005). *Technology, literacy and learning : a multimodal approach*. London ;: Routledge.

[9] Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *The Journal of the Learning Sciences, 4*(1), 39-103.

[10] Linell, P. (2009). *Rethinking language, mind, and world dialogically : interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age Publ.

[11] Medierådet (2008). *Ungar och medier 2008: Fakta om barns och ungas användning och uppfattning av medier.*

[12] Olin-Scheller, C. (2006). *Mellan Dante och Big Brother : en studie om gymnasieelevers textvärldar*. Karlstad: Estetisk-filosofiska fakulteten, Litteraturvetenskap, Karlstads universitet.

[13] Requiem for a dream (soundtrack) (2011). Retrieved 18 October 2011 from; http://en.wikipedia.org/wiki/Requiem_for_a_Dream_%28soundtrack%29

[14] Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3), 387-420.

---

[i] In a survey done by Medierådet [11] 75% of the youngsters between 12-16 years of age stated that they watched film clips on You Tube or similar sites while they were on the Internet. According to the study done by Medierådet (2008) 76% of the boys and 64% of the girls, in the age group 9-16 years, watch film clips on sites such as You Tube when on the Internet.

[ii] The following annotation conventions are adapted from *G. Jefferson, Transcription Notation, in J. Atkinson and J. Heritage (eds), Structures of Social Interaction, New York: Cambridge University Press, 1984.*

(# of seconds)	A number in parentheses indicates the time, in seconds, of a pause in speech.
(.)	A brief pause, less than one second.
:::	Indicates prolongation of a sound.
((italic text))	Annotation of non-verbal activity.
xx xxx xx	Speech which is unclear or in doubt in the transcript.

Enriching Multimodality in Learning: Integrating Computer Modeling to Support a Learner Generated Topic

Mi Song, KIM*, Wei Ching, LEE, Xiao Xuan, YE, Beaumie, KIM
National Institute of Education, Nanyang Technological University, Singapore
*misong.kim@nie.edu.sg

Abstract: In this paper, the researchers collaborated with an experienced physics teacher to design a multimodality modeling workshop involving a 3D computer model to facilitate the workshop participants' learning of their self-generated astronomy topic – lunar libration. The modified visual representation in computer model improved the quality of their observation experiences and assisted them to consolidate their understanding on lunar libration. Participatory learning environment created through the Embodied Modeling Mediated Activity (EMMA) encouraged learners to investigate their interest driven and self-generated inquiry. Researchers modified and integrated modeling software (i.e. Astronomicon) to support learners' interests. Multimodality was employed to mediate the self-directed exploration and collaborative discussion in understanding the lunar libration. This finding implies that it is important to establish a sustainable learning community to design and conduct multimodality modeling lessons in informal learning settings.

Keywords: multimodality, learner-generated topic, 3D computer modeling tool, lunar libration, learning community, informal learning

1. Introduction

Since 2008, with an emphasis on multimodality and embodied cognition, we have worked with an experienced physic teacher in Singapore to co-design the Embodied Modeling Mediated Activity (EMMA) workshops for promoting astronomy teaching and learning. Through adopting a design-based research, we focused on designing a participatory learning environment that supports learners' astronomy concept formation in informal settings. The affordances of multimodal modeling guided by effective teacher's instructional strategies were promising and positive in facilitating learning [13]. Learning astronomy through sky observation and multimodality modeling activities engaged learners to integrate new learning from observation into their existing knowledge [12]. In this study, we integrated 3D computer model to embed the motion aspect of the multimodality in learning. We investigated how computer modeling could improve workshop participants' understanding of astronomy phenomena.

The informal learning has more flexibility to address the social cultural aspect of learning since it emphasizes the authenticity in purpose, learning tools, learning context and responds to learners' interests and strengths [4]. Researchers were interested to design multimodal modeling activity workshops to cultivate inquiry-based and participatory learning. The modeling activities were designed such that learners use the best available learning tools and experiences to answer their own generated inquiry. Computer modeling was integrated to support learning. Researchers work collaboratively with school teacher, university professor and workshop participants to design appropriate modeling workshops.

The quality of the workshop design was improving through iterative cycle of refining, reflecting and implementing the workshop.

2. Theoretical framework

Modeling is the process of representing internal abstract ideas or coordinating the structures of the system by ways of simplifying, quantifying and representing with the purpose of explaining, predicting and communicating with others how the ideas work [18]. Drawing upon constructivist learning perspective, recent research in science teaching and learning has recognized the importance of multimodality in students' development of conceptual understanding [9, 15, and 17]. Especially in the domain of astronomy, multiple sensory modalities like visual, verbal, tactile, and kinesthetic are triggered when learners are engaged in the multimodal modeling activity. Many researches [2, 7] have showed the virtual computer modeling illustrated the advantages in facilitating students understanding in spatially related astronomical concepts and improving their visualization of the abstract concept. Furthermore, multimodal modeling also implies the important role of observation that could offer opportunities for learners not only to recognize inconsistencies between an observed experience and their own existing models but also to promote inquiry. Specifically, observation, whether was made in the authentic environment [19] or designed virtual environment [1], provides learners with embodied experiences. This does not only facilitate learners' conceptual learning but also enhance their motivation and interests [14]. Hence, such multimodal modeling calls for a new conceptualization of learning as participating in practice with an emphasis on bodily active engagement and the integration of sensory behavior and cognition, which have not been used very much in formal learning.

The learners generated issues or topics can mediate the learning about what and how learners will study and evaluated what they studied in problem based solving learning [10]. Moreover, the generated topics also helped learners to find the direction of individual learning that improve self-directed learning. Hurk and colleagues found that learners can determine important topics through effective discussion. Thus, it is advisable to engage learner with their self-generated topics in learning when the direction of learning were contextualized and knowledge is build within interaction in a community of peers and experts. Learner-centered design software allowed user to execute certain actions on the tool, to interpret and evaluate the data from the software to achieve learning goal and generate new learning [16]. Educator shall integrate technology tools in a constructive learning environment that promote learners to solve the self generated topic through discussion.

Wenger [20] stated that we all belong to certain communities of practices that allow us to know something or change ability through meaningful experiences by interacting with the members and social resources in the community. The individuals contribute to the practices and the community means to refine the practices to enable learning as meaningful. Researchers believed that community of practices in informal settings can be established through integrating members from different background based on similar interests in promoting modeling learning in Astronomy. The collaboration between informal learning organizations and schools can improve the science learning that encourage authenticity, different disciplines and multimodalities in learning context [4].

3. The Study

Based on embodied cognition and multimodality, we propose EMMA for bridging the sky-gazing practices and understanding of planetary motions/light through active participation. With EMMA, each mode of modeling (2D drawing models, 3D physical models and computer models) engages learners in in-depth inquiry process addressing their prior beliefs and experiences, followed by modeling exploration and discussion to enhance understanding. The curriculum design was situated in informal learning settings where multimodal modeling was employed as main approach to support learning of learners' generated topic. A design-based research has been employed so that we go through iterative cycles of co-designing, implementing and refining EMMA learning activity with participants and Physic teacher. New content was added based on learner generated topics, and modeling materials and tools were refined to meet the need of learners. Three astronomy phenomena were studied in these workshops: lunar libration, Venus transit and lunar eclipse. All topics were authentic and significant events in their observatory.

In this study, researchers worked with five male undergraduate Physic students. They had strong interests in Astronomy and were enrolled voluntary in a university's research-based project that required them to manage an observatory and conduct research on telescope installation and image processing, under supervision of an experienced professor Chen (pseudonym). According to Professor Chen, they were dedicated and diligent students who have good Physics knowledge, able to learn independently as well as to work collaboratively in group. Their names were abbreviated as HQ, KH, RY, CX and KY. Professor Chen was the collaborator in this study who taught Physics and astronomy courses in university. The modeling activities employed in this study were co-designed with an experienced Physics teacher, HJ (pseudonym). He contributed his content expertise and pedagogical ideas in the planning process.

Researchers integrated 3D computer modeling tool, called Astronomicon [6], to provide a conceptually and perceptually meaningful learning experience that might bridge the gap between their perceptions and astrophysical phenomena. Astronomicon was developed through design-based research targeting on beginning learner to explore common astronomy phenomena [8]. However, in this study, researchers modified the visual representation to leverage its' affordance to explore a unique astronomy phenomenon (i.e. lunar libration) that requires detail observation. Modeling with Astronomicon includes creating and manipulating 3D objects, viewing them from multiple perspectives, visualizing and collecting data of the system's process with provided symbolic representations (e.g., orbital plane, numeric data on time, etc.). Models can be created based on user define of the properties of planetary bodies. This allows learners to test their hypothesis by controlling certain parameters of the planetary bodies.

Pre-workshop meeting was conducted to understand learners' background and learning interests. Researchers introduced the concepts of modeling, modeling artifacts and computer modeling tool (i.e.: Astronomicon) in the meeting. Learners were interested to further explore the topic of lunar libration. Thus, researchers refined the lesson plan and learning materials to accommodate learners' generated topic. Learners spent approximately eight hours (in two workshops) in exploring lunar libration, lunar eclipse and Venus transit. The workshops took place at tutorial room. Multiple modeling materials were prepared such as Styrofoam balls, paper plates, globes, wooden sticks, etc. In order to embed the participatory culture, firstly, learners were encouraged to think of what they want to know and generated their own inquiry. Subsequently, they can select appropriate tools and create multimodal models (e.g.: virtual computer model, sketching and concrete model) during their discussion and exploration of the topic. After discussion, learners presented their understanding with their artifacts. Then, a third workshop was conducted for preparing learners to deliver their understanding to new learners. In this paper we only focused on the process of learning lunar libration through computer modeling.

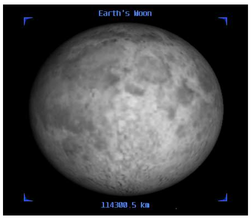
3.1 Data Collection and data analysis

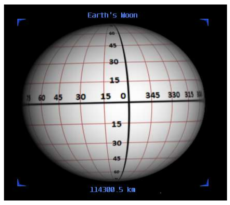
Multiples data sources were collected throughout the process of planning, implementing and evaluating the curriculum. The planning and evaluation meetings and the modeling workshops were video recoded. Screen capture video program-Camtesia© was used to record learners' interaction with Astronomicon. Researchers wrote field notes and reflective journals to document the important learning moments. In addition, learners' artifacts such as pre and post concept maps, sketching were collected to triangulate our understanding about their reasoning and conceptual development.

4. Findings

4.1 Exploring lunar libration through computer modeling

Researchers equipped themselves with the knowledge of lunar libration through group study, supported by Physic teacher, HJ. The team identified that observation was essential to understand the impact of eccentricity on lunar libration. Scientists discovered there are almost 59% (instead of 50%) of the moon's surface can be observed through the telescope. Researchers applied computer simulation to imitate and illustrate the changes of the moon's surface. However, the default moon surface image could not illustrate the minor changes of moon surface (Figure 1 (i)). Therefore, the team designed a visual representation of a "moon-with-grid" graphic to replace the moon's surface image. After modification, the graphic showed clearer evidence of visible moon surface (Figure 1 (ii)).



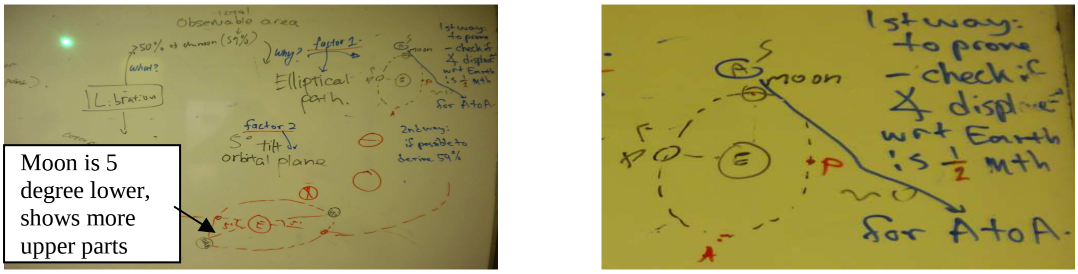


(i) Default moon image (ii) Imposed graphic of "moon-with-grid"

Figure 1. Modified visual representation of the moon.

4.2 Facilitating discussion with multimodal representations

Learners' initial inquiry question was "What is the exact mechanism for libration?" They first created individual concept map of lunar libration to hypothesize the factors causing lunar libration. Then, they co-constructed a concept map (Figure 2) to discuss their understanding and plan their strategies to prove their understanding. Collectively, they stated two factors that cause lunar libration. One of the members, HQ drew a diagram (Figure 2.0 (ii)) to explain how the moon's elliptical orbit causes lunar libration. He explained that when the moon is at apogee node (point A in Figure 2 (ii)), it rotates relatively slower and therefore we can see more surface area of the moon. At perigee node (point P in Figure 2 (ii)), the moon rotates relatively faster. After listen to HQ, another group member, KH described a way to prove their understanding through checking if there an angular displacement with refer to the Earth in the period of half of month (i.e. period from A to A). The above mentioned discussion was heavily based on learner generated visual representations and triggered their visual, spatial and kinesthetic imaginations. Factor 2 was the inclination of the moon’s orbital plane. Another member, CX illustrated the impact of the 5 degree tilted orbital plane in a diagram (Figure 2 (i)) and how it exposes more upper or lower part of the moon. Then, they tried to prove that these two factors allow them to see 59% of the moon. Diagrams and textual expressions were used to co-construct the concept map. Each member employed various modalities to deliver their understanding and they also used their hands gestures to illustrate the motion of the moon.

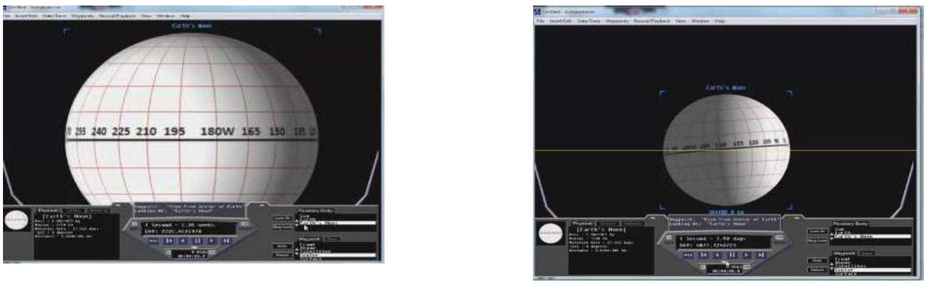


(i) Overview of concept map (ii) Diagram of apogee and perigee nodes

Figure 2. Co-constructed concept map

4.3 Interactive learning with Astronomicon

They were divided into two groups while investigating lunar libration using Astronomicon: group A with two members (CX and KY) and group B with three members (HQ, KH and RY). It was inevitable that learners encountered technical challenges while using new computer software. There was no problem on creating a system that consisted the Sun, the Earth and the Earth’s moon. However, they had difficulties in changing the viewing perspectives (e.g., from center of the Earth or above the moon) to observe the moon. Researchers provided technical guidance, including guiding them to use the modified “moon-with-grid picture” (Figure 1) to improve their observation. After learners became familiar with the software, they started controlling moon’s parameters to test their understanding. In Group A’s first attempt, they set the moon’s eccentricity and plane’s inclination to zero. They presumed that only 50% of the moon would be seen. Their observation confirmed that the moon appeared static all the time, meaning 50% of visible surface (Figure 3 (i)). Thus, they were convinced that eccentricity and orbital inclination caused the lunar libration. In their second attempt, they set the plane inclination as 5.1454 degree and replaced the eccentricity as zero. Then, they observed the vertical displacement of the moon (Figure 3 (ii)). Astronomicon simulated that the moon was moving up and down vertically, which allowed more than 50% of the moon’s surface to be visible. This provided the evidence of vertical displacement of the moon due to orbital plane inclination.



(i) Static moon (ii) Moon with vertical displacement

Figure 3. Evidences illustrated in Astronomicon

4.4 Using observational data as evidence

Learners retrieved vertical and horizontal displacement data from simulation as evidences to prove there were extra 9% of visible moon's surface. In their presentation, they used diagram (Figure 4) to delineate how they simplified the calculation by applying geometry concepts and simple percentage calculation of the angular changes. They made an assumption that the vertical displacement at the middle part of the moon would be equivalent to the displacement at the upper and lower parts of the moon. Astronomicon showed the maximum vertical displacement as 10 degree and contributed of 5.5 % of extra visible surface (i.e., 10°/180° × 100%) and the horizontal displacement as 13 degree, contributed another 7.25% (i.e., 13°/180° × 100%). In total, they deduced 12.75% of extra visible surface, which was more than 9% suggested by scientist. They explained their error was caused by a double counting region (see Figure 4). Due to time constraint they had not solved how to reduce the double counting area.

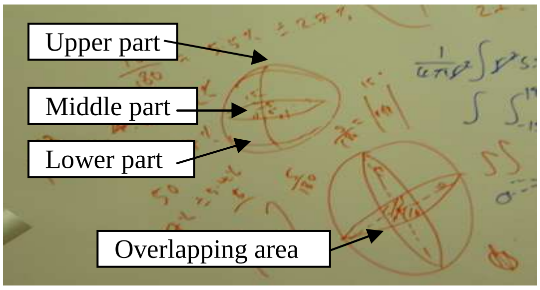


Figure 4. Group A presented their assumption and calculation through diagram

5. Discussion

5.1 Computer modeling improves visualizing and understanding phenomena

Understanding lunar libration required in-depth moon observation of the moon's movement in relation to its properties (i.e., eccentricity and inclination) and its visible surface. Despite of the technical difficulties, learners were engaged actively with the software. Based on various settings, they observed the impacts of different factors on libration and result in different exposure of the moon's visible surface. The modified visual representation afforded in-depth observation and exploration. They collected useful data from the model and applied their mathematical model to derive the extra percentage of visible moon's surface. This computer modeling was effective for these particular learners to execute and evaluate their learning [16] so as to consolidate their understanding. This finding was resonance with other research findings that computer modeling has the advantages to understand the motion of planetary objects in a 3D virtual environment. The features of allowing users to change viewing perspectives, providing virtual observation and supporting interactive modeling were the succeed features for deeper learning [3, 11].

5.2 Accommodate learner-generated topic with multimodality approach

EMMA aims to promote participatory learning that supports learner interests, closely related to sky observation and life experiences. In order to accommodate leaner-generated topics, researchers identified essential learning challenges by working with experienced teacher. Researchers integrate computer modeling to improve learners' visualization and to enrich their multimodality embodied experience. Multimodality was highly recommended based on our pragmatic experience in informal settings [13]. Researchers established multimodality learning environment by providing choices of modeling tools (i.e. drawing, concrete model, computer model and mathematical model), observation data and diagrams. Thus, learners were encouraged to employ diagrams, written explanation, hands gestures, calculate with data to explore the phenomena and represent their understanding explicitly during their discussion. The interactions among members were enhanced in multimodality environment when each member contributed and supplemented each other ideas by various modalities. The learner generated topic encouraged active participation in group discussion and this had effectively scaffold their learning by identifying the causes of lunar libration. This positive finding encouraged educators to adopt learner generated topic in lesson design. The quality of learner generated learning issues can be improved through good facilitation during group discussion [10].

5.3 Community building

The community of practices provided meaningful learning through active participation [20]. In this practice, we noticed researchers, teacher and learners played unique roles in the lesson design and learning process. The learning outcomes were more holistic and inclusive not only for learners but for all the members in this learning community. Learners became active learners by suggesting the topic they were interested and solved the problem more effectively. Teacher contributed his Physic expertise to identify the core concept and essential observation. On the other hand, researchers contributed their expertise by leveraging the affordance of the computer model. In the learning process, learners generated their strategies in solving their problems. Their approaches provided new insights for researchers to improve the pedagogy. The roles of teacher and learner are interchangeable based on the situation [5]. Based on this collaborative practice, we noticed the importance of establishing sustainable learning community in informal context by integrating varied expertise to design meaningful lesson.

6. Conclusion and implication

Multimodality supported learners to illustrate their understanding explicitly and generate effective discussion. We suggested educators to employ multimodality to improve collaborative discussion that encourage each member to contribute their knowledge and experience. Computer modeling with visual and motion enhancement enriched the multimodality in learning environment that facilitated learning by improving learners' visualization based on modified visual representation. The interactive feature of creating model with user defined parameter helped learners to understand the attributive factors and the impacts on lunar libration. We hope this finding encourage teachers to integrate computer software to improve learning on topics that required spatial visualization.

Participatory learning culture was cultivated by designing lesson that accommodate to learners' generated topic. Learners were actively involved to solve their generated topics through discussion. This positive finding encouraged educators to adopt learner generated topic in lesson design. Collaboration between experienced teacher and researchers had improved the lesson design by contributing their expertise in refining the content, pedagogy and learning tools. We suggested some efforts are needed to sustain the collaboration between schools and informal learning organization that encourage authenticity, different disciplines and multimodalities in learning context.

Acknowledgements

The research was funded by grant OER 30/09 from the Office of Educational Research, National Institute of Education, Singapore.

References

[1] Bakas, C., & Mikropoulos, T. (2003). Design of virtual environments for the comprehension of planetary phenomena based on students' ideas. *International Journal of Science Education, 25*(8), 949-967.

[2] Barnett, M., Yamagata-Lynch, L., Keating, T., Barab, S. A., & Hay, K. E. (2005). Using virtual reality computer models to support student understanding of astronomical concepts. *Journal of Computers in Mathematics & Science Teaching, 24*(4), 333-356.

[3] Baxter, J. H., & Preece, P. F. W. (1999). Interactive multimedia and concrete three-dimensional modelling. *Journal of Computer Assisted Learning, 15*(4), 323-331.

[4] Bevan, B., Dilion, J., Hein, G. E., Macdonald, M., Michalchik, V., Miller, D., et al. (2010). *Making science matter: Collaborations between informal science organizations and schools. A CAISE inquiry group report*. Washington D.C.: Center for Advancement of Informal Science Education (CAISE).

[5] Boyer, L., & Roth, W.-M. (2006). Learning and teaching as emergent features of informal settings: An ethnographic study in an environmental action group *Science Education, 90*, 1028-1049.

[6] Cybernet-System. (2004). Astronomicon: Celestial construction tool kit. Ann Arbor, MI: Cybernet Systems Corp.

[7] Hansen, J. A., Barnett, M., MaKinster, J. G., & Keating, T. (2004). The impact of three-dimensional computational modeling on student understanding of astronomical concepts: a quantitative analysis. *International Journal of Science Education, 26*(11), 1365-1378.

[8] Hay, K. E., Kim, B., & Roy, T. C. (2005). Design-based research - more than formative assessment? An account of the Virtual Solar System project. *Educational Technology, 45*(1), 34-38.

[9] Hull, A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. *Written Communication, 22*(2), 224-261.

[10] Hurk, M. M. V. D., Wolfhagen, I. H. A. P., Dolmans, D. H. J. M., & Vleuten, C. P. M. V. D. (1999). Student-generated learning issues: A guide for individual study? *Education for Health, 12*(2), 213-221.

[11] Keating, T., Barnett, M., Barab, S. A., & Hay, K. E. (2002). The virtual solar system project: Developing conceptual understanding of astronomical concepts through building three-dimensional computational models. *Journal of Science Education & Technology, 11*(3), 261-275.

[12] Kim, M. S., Lee, W. C., & Kim, B. (2011). *Modeling the Night Sky: A Case Study of Singaporean Youth*. Paper presented at the Global Learn Asia Pacific 2011 Global Conference on Learning and Teachnology.

[13] Kim, M. S., Lee, W. C., & Ye, X. X. (2012). *Teacher's instructional strategies in multimodal modeling-based learning for understanding the moon phases*. Paper presented at the American Educational Research Association Conference.

[14] Kucukozer, H., Korkusuz, M. E., Kucukozer, H. A., & Yurumezoglu, K. (2009). The Effect of 3D Computer Modeling and Observation-Based Instruction on the Conceptual Change regarding Basic Concepts of Astronomy in Elementary School Students. *Astronomy Education Review, 8*(1).

[15] Lemke, J. (2004). The literacies of science. In E. W. Saul (Ed.), *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (pp. 33-47). Arlington, VA: International Reading Association and National Science Teachers Association.

[16] Quintana, C., Shin, N., Norris, C., & Soloway, E. (2006). Learner-centered design: Reflections on the past and directions for the future. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 119-132). New York, USA: Cambridge University Press.

[17] Roth, W. M., & Lawless, D. (2002). Science, cultures and the emergence of language. *Science Education, 86*, 368-385.

[18] Shen, J., & Confrey, J. (2007). From conceptual change to transformative modeling: A case study of an elementary teacher in learning astronomy. *Science Education, 91*(6), 948-966.

[19] Trundle, K. C., Atwood, R. K., Christopher, J. E., & Sackes, M. (2010). The Effect of Guided Inquiry-Based Instruction on Middle School Students' Understanding of Lunar Concepts. *Research in Science Education, 40*(3), 451-478.

[20] Wenger, E. (Ed.). (1998). *Communities of practices: Learning, meaning and identity*. Cambridge, UK: Cambridge University Press.

Feasibility study of using social networks platform for learning support: an example of Facebook

Chien-Lung Chan[a*], Wen-Erh Fu[a] , K. Robert Lai[b] & Shu-Fen Tseng[c]
[a]*Department of Information Management, Yuan Ze University, Taiwan*
[b]*Department of Computer Science & Engineering, Yuan Ze University, Taiwan*
[c]*Graduate Program in Social Informatics, Yuan Ze University, Taiwan*
* e-mail: clchan@im.yzu.edu.tw

Abstract: In this study, we explored the students' computer-supported collaborative learning behavior based on the Facebook platform. Sixty two senior college students major in Information Management took Decision Support System (DSS) class. Besides the lectures and class discussion, the students participated in the DSS Facebook for collaborative learning. We found that students' characteristics (e.g., gender and mindset of learning) are important factors to affect their Facebook usage behavior and learning performance. The students using DSS Facebook more often get better performance in their final projects, learning satisfaction and the online communication behavior survey. We also found that gender affects the usage of social networks platform. For instance, male students use social networks platform several times per week and get better performance in online communication, learning satisfaction and creativity self-efficacy.

Keywords: Social Networks Platform, Computer-supported Collaborative Learning, Learning Satisfaction

Introduction

The concept of E-learning was raised by Jay Cross in 1999. In 2006, the E-learning 2.0 concept was maturing as anyone could use the information technology (or search engine) to enhance his/her learning efficacy. With the mobile device becoming more popular and the network transmission speed getting faster, we can search and learn valuable knowledge easily anywhere. Even more, students can take the distance learning class in the comfort of their homes. However, previous studies found that the efficacy of e-learning is very limited. This is because the teacher's expectations are unclear, communication lacking, user interface been poor coupled with slow access speed [1-4].

The social networks platforms such as Facebook, plurk, and twitter provide space where users with same interests can gather together and interact. These networks platforms provide faster access speed in both synchronous and asynchronous communication methods [5]. Especially, Facebook provided friendly interface platforms where users interact conveniently. User can share information from other websites and get the latest news immediately as well [6].

In this study, we explored the feasibility of using social networks platform as learning support tool. What kinds of student characteristics can affect their usage behavior? Can social networks platform like Facebook be an effective learning support tool?

1. Materials and Literature Review

1.1 Computer-supported collaborative learning (CSCL)

The collaborating learning is a form of learner and learner interaction. When first used in the industry to promote the productivity in 1930s, collaborative learning has been considered as an effective instructional method in both traditional and distance learning. Based on the Internet and powerful computing techniques, computer-supported collaborative learning can shorten the learning time and the venues. It can enable learning and communicating taking place synchronously or asynchronously. People from different fields can raise different viewpoints through the collaborative learning process as well. Learners can also contribute and share their knowledge [7]. However, several differences can be identified as existing between E-learning and the computer-supported collaborating learning. Although instructors can upload digital teaching materials in an E-learning environment, lacking of face-to-face communication can reduce the students' participation. The digital teaching materials must refresh frequently in E-learning; therefore, the instructors should pay greater attention to update their teaching website. On the other hand, a computer-supported collaborative learning should focus on interactive process, active learning and knowledge sharing. The learning style can be diversified in computer-supported collaborative learning, for example, people communicate face-to-face or at distance, in both synchronously or asynchronously. [7-9].

1.2 Social Networks Platform (Facebook)

Facebook was started by several college students from Harvard in 2004. Until today, Facebook has generated over 800 million users from all over the world. As a community networks platform, users can register by simply entering e-mail, nickname or real name, date of birth, working place and interest. Users also can easily create their own webpage, and interact with other Facebook user counterparts.

When used as a learning support platform, Facebook provides diversity discussion tools for user. It reminds user when a new message is received. Facebook provides synchronous and asynchronous communication functions in graffiti wall. Even more, Facebook provides users with friendly interfaces as well. In Taiwan alone, more than 10 million people use Facebook. The experience of using Facebook tools increases the feasibility in ones usage of Facebook for learning purpose. However, The entertainment application becomes an obstacle while students take Facebook as a learning tool[10]. Compare to traditional learning, like blackboard or PowerPoint slide, students seldom raise their hands to ask questions, most of all, lack interaction in class. In this study, we used Facebook as a learning support tool by taking the advantages of its strength, despite its original social networking purpose.

1.3 Learning performance

The learning performance is the result of a student taking a course or the learning activities. In this study, besides the project scores, we use questionnaire to evaluate the learning performance.

1.3.1 Online communication

Conrath and Zeccola [10] tried to analyze the effect of computer mediated communication (CMC) on student learning, and founded that CMC can provide both positive and negative learning effects on students. For the positive effects, they believe that CMC can help students to get more learning opportunities. For example, education-based Social Networking Sites (SNSs) can be regarded as an effective technological tool for enhancing the quality of learning for students who are in higher level education courses (tertiary educations)

1.3.2 Learning Satisfaction

The previous studies investigated the relationships between various variables of students, including prior computer experience, gender, age, scholastic aptitude, learning styles, and learning experience from a Web-based course. The result showed that most of the students prefer to earn the learning experiences from a Web-based course. Moreover, the effect of learning performance from the Web-based course is almost the same as the face-to-face version [11-13] .

1.3.3 Creativity Self-efficacy

Bandura and Cervone [14] regarded self-efficacy as an important condition for creative productivity and the discovery of new knowledge. For instance, in measuring the effect of personal creative behaviors, they considered self-efficacy as a critical component in their model. Creative self-efficacy appears to provide such momentum in that strong efficacy is believed to enhance the persistence level and the coping efforts individuals will demonstrate when encountering challenging situation. Ford (1996) bring forward self-efficacy as a key motivational component in individual action, and those who have low sense of self-efficacy may be easily discouraged by failure in progress[16].

2. Research design

There are totally 62 students taking Decision Support System(DSS) course. This is a required course for senior college students major in Information Management. The experiment lasts for two months, starting from the mid-term exam to the end of the semester. The students were separated into 12 groups. Each group was constituted of 5~6 members. The group projects' tasks were to determine a decision problem by students themselves and to construct a prototype of DSS for each group of students. Each group has to define their own topics. It needs group member discussion and brainstorming. We applied Facebook to support the students' group projects. The following items were what student used: (a) Graffiti wall: To post articles or ask a question online. (b) Discussion boards: Each group has its own discussion board. (c) Document uploaded: A tool that student can upload articles or news. (d) Great/Good: One of the participant counters. And (e) Response: After the students read one article, they can give their feedbacks.

2.1 Research framework

In this study, the research framework contains three major entities: student properties, use of Facebook and learning performance. When exploring the relationship between student properties and use of Facebook, we use gender, class attendance and the scores of midterm exam and relation to their Facebook usage behavior like contribution and the frequency of use per week. The relationship between student properties and learning performance is investigated by using the questionnaire of online communication, learning satisfaction and creativity self-efficacy. Finally, we studied the relationship between the use of Facebook frequency and the learning performance. The research questions are proposed as follows:

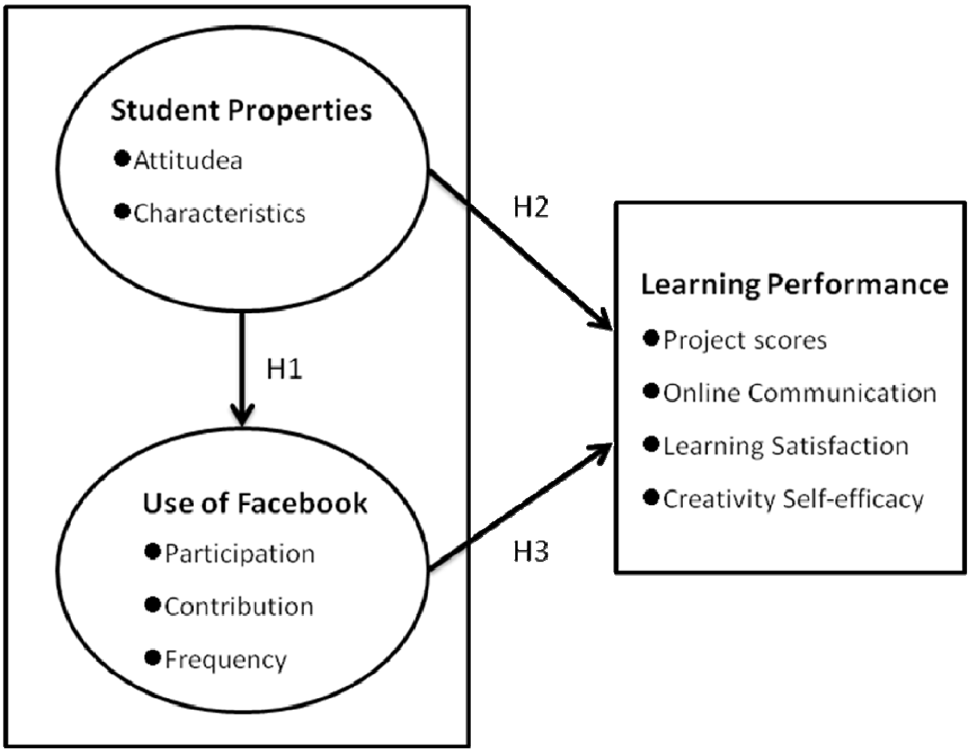


Figure1. Research Framework

1. **What is the relationship between students' properties and the use of Facebook? Who are the heavy users?**
   We try to find the factors (the student properties include gender, the time spend on the internet, the academic learning attitude) which affect the usage of Facebook participation, contribution and what make students use Facebook heavily.

2. **What is the relationship between students' properties and the learning performance? Does using of Facebook play the mediating role?**
   We try to find the factors which affect the learning performance, as well as the final project performance, online communication, learning satisfaction and creativity self-efficacy. However, in this stage, the senior students may change their focus into the job interview or study hard for master program entrance exams that treat the DSS final project grade not so seriously. The previous study showed that learning attitude affects the learning satisfaction and the online communication.

3. **What is the relationship between usage of Facebook and learning performance?**
   Facebook was applied as a discussion platform for the group project. Students can use the tools and share the knowledge immediately. During the process, we recorded both the frequency of usage of Facebook and the content. The students' participation and contribution were evaluated by other peer group members to avoid bias.

2.2 Questionnaire design

The questionnaire was designed in four parts: basic personal information, online communication behavior, learning satisfaction and creativity self-efficacy.

The personal information includes gender, daily using internet hours, actual use of Facebook (hours) and frequency of online discussion.
Online Communication Behavior

- The online communication survey questionnaire was adopted from the Wang's CMC discussion model (2011) [15]. Five positive questions and five negative ones of a 5-point Likert scale. The reliability is 0.82 for positive questions and 0.72 for negative ones. In this study, we use the positive parts because of their higher reliability.

Learning Satisfaction

- The learning satisfaction questionnaire was adopted from Gunawardena and Zittle's study (1997) [2]. The reliability for ten-question construct is 0.86.

Creativity Self-efficacy

- The creativity self-efficacy was adopted from the Tierney and Farmer's study (2002) [16]. The reliability for twelve-question construct is 0.88

3. Results

In the research question 1: What are the relationship between the student properties (learning attitude and gender) and the use of Facebook (participation, contribution and the frequency of using Facebook)? The active student means the student attending discussion more often and getting higher midterm exam score. In this study, we try to figure out the difference between active students and non-active ones until the end of final project.

As in Table 1, the active students get higher participation scores than the non-active students significantly ($p < 0.05$). Regarding gender differences, the male students use Facebook more often than the female ones ($p < 0.05$). The following items were used to evaluate the student learning attitude: Attendance: 20%, Individual homework: 20%, Midterm exam: 60%, Ask questions & Presentation 10% (Extra points). Those scores lower than 60 were defined as non-active. Totally 34 students (17 male and 17 female) were non-active. Those scores higher than 60 were defined as active. Totally 28 students (21 male and 7 female) were active.

Table 1. The student properties and the use of Facebook

Variables	non active n = 28	Active n = 34	t value	p value
Participation	4.20	4.48	-2.206	0.031*
Contribution	4.17	4.46	-1.979	0.052
Weekly use hours	3.11	3.09	0.019	0.985
Variables	**Female n = 24**	**Male n = 38**	**t value**	**p value**
Participation	4.63	4.45	1.523	0.133
Contribution	4.59	4.32	1.978	0.053
Weekly use hours	1.90	3.86	-2.421	0.019*

Research question 2: What are the relationships between the student properties (learning attitude and gender) and the learning performance (project scores, online communication, learning satisfaction and creativity self-efficacy)?

Table 2 shows the relation between learning attitude and creativity self-efficacy ($p < 0.05$). For the college students, measuring their learning attitude by their class performance could be insignificant. This is because the learning performance can be affected by group members or the characteristics. And the small sample size also causes the insignificant situation. Member from the same group gets the same scores. Table 1 shows the male students spend more time in online discussion. And in Table 2, the male students get higher scores in creativity self-efficacy significantly ($p < 0.05$).

Table 2. The student properties and the learning result

Variables	non active n = 28	Ative n = 34	t value	p value
Project scores	85.36	84.18	0.994	0.324
Online communication	20.25	19.65	1.030	0.307
Learning satisfaction	40.68	39.29	1.277	0.207
Creativity self-efficacy	40.25	37.62	2.017	0.048*
Variables	**Female n = 24**	**Male n = 38**	**t value**	**p value**
Project scores	85.33	84.32	0.762	0.451
Online communication	19.21	20.37	-1.968	0.054
Learning satisfaction	38.63	40.74	-1.939	0.057
Creativity self-efficacy	37.08	39.89	-2.115	0.039*

Research question 3: What are the relationships between use of Facebook (participation, contribution and the frequency use Facebook) and the learning performance (project scores, online communication, learning satisfaction and creativity self-efficacy)?

In the Table 3, the student who had been posting and responding gets higher project scores (83.95, 86.09). Those who discuss twice a week also get higher in project scores (83.59, 85.87), online communication (19.38, 20.39) and leaning satisfaction significantly ($p < 0.05$). The result shows that student use online discussion frequently can help them get better learning performance, especially in the learning satisfaction.

Table 3. The usage of Facebook and the learning performance

Variables	Non post/resp n = 40	Post/ response n = 22	t value	p value
Project scores	83.95	86.09	-1.763	0.083
Online communication	19.90	19.95	-0.88	0.930
Learning satisfaction	40.20	39.41	0.695	0.490
Creativity self-efficacy	38.90	38.64	0.188	0.852
Variables	**dicussion < 2 n = 29**	**dicussion > 2 n = 33**	**t value**	**p value**
Project scores	83.59	85.87	1.815	0.074
Online communication	19.38	20.39	1.752	0.085

Learning satisfaction	38.52	41.15	0.528	0.014*
Creativity self-efficacy	39.79	39.70	1.440	0.155

Table 4 shows that the relation between participation, contribution and discussion times are highly significant. The contribution and project scores show no correlation with other variables. Online communication, learning satisfaction and creativity self-efficacy were highly correlated.

Table 4. The relation between usage of Facebook and the learning result

	Participation	Contribution	Discussion times	Project scores	Online communication	Learning satisfaction	Creativity self-efficacy
Participation	1						
Contribution	0.755***	1					
Discussion times	0.356**	0.151	1				
Project scores	0.099	0.175	0.183	1			
Online communication	-0.120	-0.116	0.239*	0.109	1		
Learning satisfaction	-0.153	-0.213	0.314*	0.195	0.681***	1	
Creativity self-efficacy	-0.111	-0.083	0.148	0.084	0.516***	0.546***	1

*means p value < 0.05 ** means p value < 0.01 *** means p value < 0.001

4. Discussion

In this study, we explored the relationship between the student properties, the use of Facebook and the learning performance. We found that the male students spent much time in online discussion, they get higher online communication, learning satisfaction and the creativity self-efficacy. The female students however have higher participation and higher contribution than their male counterparts. And the female students also get higher final project scores.

The online communication provides chances for information and the learning material exchange. Students can get more information in social networks platform than when they search by themselves. As a learning support purpose, the instructor needs to pay keen attention on the platform management and keep the students follow the latest news. In this study, it is observed that students get a good experience when they use DSS Facebook. Their frequencies of use of Facebook are also important factors here. Students discussing frequently enable themselves to gain higher final project scores, online communication and learning satisfaction.

5. Future Development

In the future, we would increase the sample size as well as including different types of courses. While heterogeneity group affects the learning performance, we intend to design an experiment to investigate this phenomenon. The usage experience of Facebook as learning support tool can provide feedback for development of new platform for collaborative learning.

Acknowledgement

This study was supported by a grant from National Science Council, Executive Yuan, Taiwan (NSC100-2632-S-155-001).

References

[1]. Giannousi, M., et al., *Students' satisfaction from blended learning instruction.* Internet, 2009. **1**(1): p. 61-68.

[2]. So, H.J. and T.A. Brush, *Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors.* Computers & Education, 2008. **51**(1): p. 318-336.

[3]. OReilly, T., *What is Web 2.0: Design patterns and business models for the next generation of software.* 2007.

[4]. Shachar, M. and Y. Neumann, *Differences between traditional and distance education academic performances: A meta-analytic approach.* The International Review of Research in Open and Distance Learning, 2003. **4**(2): p. Article 4.2. 9.

[5]. Clough, G., et al., *Informal learning with PDAs and smartphones.* Journal of Computer Assisted Learning, 2008. **24**(5): p. 359-371.

[6]. Baran, B., *Facebook as a formal instructional environment.* British Journal of Educational Technology, 2010. **41**(6): p. E146-E149.

[7]. Lowyck, J. and J. Pöysä, *Design of collaborative learning environments.* Computers in Human Behavior, 2001. **17**(5-6): p. 507-516.

[8]. Cho, H., et al., *Social networks, communication styles, and learning performance in a CSCL community.* Computers & Education, 2007. **49**(2): p. 309-329.

[9]. Gülnar, B., Ş. Balcı, and V. Çakır, *Motivations of Facebook, You Tube and Similar Web Sites Users.* 2010, Bilig.

[10]. Conrath, K. and J. Zeccola, *Does social networking hurt student grades.* American Teacher, 2009. **94**(2): p. 3.

[11]. Hong, K.S., *Relationships between students' and instructional variables with satisfaction and learning from a Web-based course.* The Internet and Higher Education, 2002. **5**(3): p. 267-281.

[12]. Howard, G.S. and R.R. Schmeck, *Relationship of changes in student motivation to student evaluations of instruction.* Research in Higher Education, 1979. **10**(4): p. 305-315.

[13]. Allen, M., et al., *Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis.* The American Journal of Distance Education, 2002. **16**(2): p. 83-97.

[14]. Bandura, A. and D. Cervone, *Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems.* Journal of personality and social psychology, 1983. **45**(5): p. 1017.

[15]. Wang, A. and S. Kong, *A Study of Relations between Students' CMC Behaviors and Perceived Effects of CMC on Learning for Incorporating CMC in Hybrid Learning.* Hybrid Learning, 2011: p. 95-104.

[16]. Tierney, P. and S.M. Farmer, *Creative self-efficacy: Its potential antecedents and relationship to creative performance.* Academy of Management Journal, 2002: p. 1137-1148.

Design for Scaffolding Collaborative Inquiry and Academic Literacy for Chinese Tertiary Business Students

Ke ZHAO [a*], Carol C.K.K. CHAN[b]
[a] *Foreign Language Department, Shanghai University of Finance and Economics, PRC*
[b] *Faculty of Education, University of Hong Kong, PRC*
*zhao.ke@mail.shufe.edu.cn

Abstract: This study investigated how collaborative inquiry and social literacy acquisition took place in a computer-supported collaborative inquiry environment for Chinese tertiary business students. Participants were 102 Year 1 business students in four intact classes. Two classes experienced a principle-based computer-supported environment and two a conventional project-based approach. Data included survey, writing quality, portfolio, focus group interview and online interactions. Quantitative analyses indicated instructional groups outperformed comparison groups on conceptual understanding, inquiry processes and argumentative writing. Interview study characterized change dynamics under four contextual themes-- epistemology, pedagogy, technology, community, pointing to facilitative role of design principles. Micro-level online discourse analysis characterized nature of discourse moves and group patterns suggesting a significant role of socio-metacognitive and explanatory discourse mediated by technology.

Keywords: Collaborative Inquiry, change dynamics, online discourse, business education

Introduction

This study aims to examine the role of a computer-supported collaborative inquiry (CSCI) environment premised on knowledge building principles in fostering collaborative inquiry and social literacy among Chinese tertiary students. Specifically, we examined (a) the effects of the designed CSCI environment on students' conceptions of collaboration, conceptual understanding and academic literacy, (b) change dynamics mediated by the interplay of contextual factors, and (c) the nature of online discourse to understand how changes took place mediated by socio-cognitive and technological dynamics.

Educational reforms in tertiary education now focus on fostering collaborative inquiry and knowledge creation. Innovative models of learning and teaching, such as project-based learning and computer-supported collaborative inquiry learning have become commonplace. However, there is still a dearth of contextual understanding and deep analysis of whether and how these instructional models scaffold students' learning and collaboration in authentic classroom settings.

Computer-supported collaborative learning (CSCL) has received increasing research attention in higher education. Much research has been conducted on designing new computer software systems or examining social and cognitive processes mediated by technology (Koschmann, Hall, & Miyake, 2002; Stahl, 2006). However, a major concern remains regarding how technology can be utilized to enhance learning and collaboration in complex classroom settings. More specifically, deep contextual understanding is called for of how learning and collaboration are impacted by interactions between pedagogy, technology-mediated design and institutional practices in higher education.

Premised on the knowledge-creation metaphor of learning, knowledge building is an influential model in collaborative inquiry (Paavola, Lipponen, & Hakkarainen, 2004). Although considerable research evidence has shown its role in advancing collective idea improvement (Zhang, Scardamalia, Lamon, Messina, & Reeve, 2007), progress inquiry and social literate skills (Sun, Zhang, & Scardamalia, 2010; Zhang & Sun, 2011), the majority of the research has been conducted in science domains among school-aged children (Hämäläinen, Manninen, Järvelä, & Häkkinen, 2006; Zhang et al., 2007). Little has been known about how students co-constructed ideas and advanced their understanding in other domains, such as business education, among tertiary students. Investigating how Chinese tertiary students construct collective understanding of business concepts using English as a foreign language may extend our understanding of the role of and design for knowledge building in a different cultural context.

In addition, project-based learning is now advocated in tertiary business education (Brzovic & Matz, 2009; Eastman & Swift, 2002; Rooney, 2000), yet it is often considered as task-completion with division of labor. How students approach knowledge building of conceptual ideas integrated with project inquiry will be an important area for examination.

To sum up, this study aimed to examine the role of the knowledge-building inquiry environment, to investigate change dynamics in the designed environment, and to further examine meditational role of technology and social dynamics. Three questions were addressed: 1.What were the instructional effects? Specifically, did students in the designed environment perform better than comparison students in conception of collaboration, conceptual understanding, and academic literacy? 2. How did changes take place mediated by the interactions with the contextual factors in the designed environment? 3. What were the discourse patterns distinguishing high- and low-performance groups?

1. Methods

1.1 Participants

The study included 102 Year 1 English for International Business (EIB) students from four intact classes in a University in Shanghai. The four classes had similar achievement levels, generally at low-average levels compared with the same year students in other programs. A quasi-experimental design was employed: two tutors each taught one instructional class and one regular class. 52 students were in CSCI groups and 50 in PjBL groups.

1.2 Design of CSCI environment

The research was conducted in a 12-week core module of EIB to develop students' understanding of concepts in Total Quality Management (TQM) and research and literacy skills. This module was originally featured by a group project investigating TQM implementation in businesses in Shanghai. The PjBL groups mainly did project work after class following the project guideline prescribed by tutors in class.

A CSCI environment was designed for instructional groups to scaffold collaborative inquiry and academic literacy. Knowledge-building pedagogy was considered in the instructional design, as it has been evidenced to promote communal conceptual understanding as well as social-cognitive and social literate skills (Zhang & Sun, 2011). Considering the domain, cultural and instructional specificity of this study, we adapted four of knowledge building principles as design guidelines as follows:

1. *Collective cognitive responsibility*: creating social structures or dynamics for collective knowledge advancement with intentional use of English;
2. *Epistemic agency*: focusing on ideas, problems and co-construction of knowledge and deep understanding, rather than on completion of project tasks;
3. *Authentic problems, improvable ideas*: identifying authentic problems and progressively improving ideas in connection with the broader business community; and,
4. *Community Knowledge*: working together with effective communicative strategies to allows improved ideas and theories to diffuse through the communal knowledge space;

1.3 Data sources

Language Proficiency Students' pre-test English proficiency scores were collected.

Conceptual Understanding and Literacy Students were required to write an essay on TQM. Essays were analyzed using two rubrics; conceptual understanding examined TQM concepts and academic literacy including explanation, organization and mechanics of language. Inter-rater reliability was. 80 and .82, respectively.

Group Learning Portfolio Group portfolios document students' reflection and collaborative learning process. The portfolio was rated on a 6-point scale ranging from fragmentary responses to deep collaborative reflection; the inter-rater reliability was .86.

Collaboration A pre- and post questionnaire survey on conceptions of collaboration was administered using the Collaborative Online Learning Scale (COLS) (Chan & Chan, 2011). Cronabach's alpha coefficients of the COLS were .79 and .76 for collaboration and online-learning respectively.

Focus Group Interview Four successful project groups from both learning environments each participated in a 40-minute interview. Interview questions were designed adapted from Bielaczyc's social infrastructure framework (2006). Four dimensions include: beliefs about learning; pedagogical practices; social-technological dynamics and connection with community. Exemplar interview questions are "*Can you say something about the course design using TQM project? In what way is it different from other courses?*" "*What did your tutor do differently from those in other courses*?" "*How do you like the assessments in this course?*" "*Can you say something about your knowledge Forum (KF) activities?" "Did you use any computer technology in doing the project work? (for RPBL groups) What role do you think technology play on your project learning*?"

Online KF discussion note To further examine how collaboration and social literacy acquisition took place in the designed CSCI environment, students' Knowledge Forum discussion notes were analyzed quantitatively and qualitatively. Following current CSCL discourse studies of quantifying verbal data analysis (Broxel et al. 2000; Hmelo-Sivler, 2003), the online discourse analysis in this study took a multilevel and multidimensional approach, combining both qualitative and quantitative analyses to capture both the cognitive and social processes of collaborative project inquiry.

2. Results

2.1 Effects on Conceptual Understanding, Literacy and Collaboration

Descriptive statistics of the COLS, conceptual understanding and academic literacy are presented in Table 1. The results of MANCOVA analysis controlling for differences in language proficiency showed significant differences between groups; univariate analyses showed significant group differences on conceptual understanding (F (1, 97) =6.77, p<.01, η^2=.07) and argumentation,(F (1, 97) =8.03, p<.01, η^2=.08), favoring CSCIL groups.

Table 1 *Means and Standard Deviations (in parenthesis) for the COLS, the SPQ, Conceptual Understanding, Argumentation, Organization and Language Use*

	Collaboration		Online learning		Conceptual Understanding (Max.=100)	Academic Literacy		
	Pre	Post	Pre	Post		Exp'n	Orga'n	Lang'e
RjPBL (n=49)	3.49 (.80)	3.47 (.56)	4.03 (.89)	3.83 (.78)	47.98 (10.49)	47.59 (10.33)	49.51 (9.89)	48.11 (10.56)
CSCI (n=53)	3.41 (.69)	3.97 (.58)	4.02 (.84)	4.42 (.40)	52.28 (9.10)	52.38 (9.40)	50.48 (10.24)	51.90 (9.40)

Analyses on conceptions of collaboration using repeated measures, controlling for pretest language scores, indicated significant interaction effects. Follow-up analyses indicated significant interaction effects for online learning (F (1, 98) = 8.36, p < .01, η^2 = .08) and collaboration (F (1, 98) = 10.43, p < .01, η^2 =. 10) favoring CSCI groups. These results suggest that instructional groups obtained better conceptual understanding and argumentation, and their conceptions of collaboration tended to become more sophisticated over time than their counterparts.

2.2 Change dynamics in the learning environment

This part reports briefly about key findings from focus group interviews. Analyses identified the salient contextual dynamics in the designed environment that contributed to student changes. Four interdependent and interactive contextual themes in the designed learning environment have been identified as impacting students' changes in learning and collaboration. As shown in Figure 1, four contextual themes were identified in the learning environment: 1) ***epistemological beliefs*** in the instructional design; 2) ***pedagogy***, including the role of teachers and assessment practices; 3) ***socio–technological dynamics***; and, 4) their ***connections with the outside business community***.

First, the design facilitated changes in students' embedded beliefs about learning and knowledge from being initial conflict when confronted with uncertainty of knowledge, to collective constructive use of authoritative information, and then to justification for advancing communal knowledge. Almost all students mentioned their epistemological conflicts at the early stage of project work. For example, one student related her conflict over uncertainty about knowledge: *"The difficult part lay in the different definitions of TQM. At the beginning of the semester, we were total confused by so many definitions of TQM and by the abstract principles. We all expected someone to tell us that this was an authoritative definition and these were useful principles. What we need to do is just note them down. You know, the more we read, asked, and searched, the more we felt confused and hopeless, and had no clear direction for the project." (Student #2, Group #2)*

However, CSCI students mentioned how their epistemological anxieties about information processing were lessened by their collaborative inquiry on knowledge forum, which enhanced their justification and changed their conception of collaboration. As one CSCI student pointed out, justification was not about getting an answer; it was about advancing communal understanding: *"Talking on Knowledge Forum (KF) helped us view different information, different perspectives….But when discussing the framework for the investigation, we got different opinions… relating to the first visit (a pilot investigation) and the principles from TQM theories, we articulated different opinions and tried to convince others with some evidence, reasons, examples to support… In most cases, there seems no definite right or wrong idea, but we may have a better idea after weighing over from different perspectives…Yes, this enhanced our understanding of TQM at that stage. When* *reading back our database, we could see the development (of our understanding).” (Student# 1, Group # 2)*

These analyses suggested that the design of TQM project learning triggered cognitive conflict and changes in deep-rooted beliefs in certainty and authority of knowledge. Successful students, particularly in CSCI environment, who resolved their emerging epistemological conflicts, were simultaneously able to adopt more sophisticated beliefs about learning and knowledge. The CSCI design facilitated the development of high-level collective inquiry skills that could be transferred to other academic and work contexts, including identifying knowledge gaps, making constructive use of information, and cultivating justification.

Second, changes in pedagogical practices, with particular reference to teacher roles and assessment, were conceptualized as contributing to student change. Specifically, new conceptions of teachers as co-inquirers in the knowledge-building community rather than authoritative sources of knowledge may facilitate changes in beliefs about and strategies for learning and collaboration. Use of alternative, formative and collective assessment approaches matures students’ views of assessment; rather than viewing it as the mastery of knowledge through rote memorization, they began to see it as the execution of collective agency and ownership for communal knowledge advancement, thus enhancing their social metacognitive strategy use in collaborative inquiry.

Third, social-technological dynamics scaffolding changes in both the process and the product of project inquiry were unraveled. Designed use of technology went beyond mere communication for knowledge sharing, to a cognitive, metacognitive and collaborative tool. Technology extended knowledge-building discourse through integrating concept learning, project inquiry, and language learning in enhanced socio-technological dynamics.

While only CSCI students had access to the KF online learning platform, most regular project groups reported spontaneous using computer technology such as QQ online instant messaging, MSN Messenger in their project learning. Analyses of interview data characterized their views of computer technology as knowledge-sharing and superficial procedural decision-making. In contrast, In addition to seeing technology as a medium for knowledge sharing or participation, some CSCI students commented on the impact of Knowledge Forum’s scaffolding role on conceptual understanding, collaborative inquiry and literacy development. Their viewed Knowledge Forum as: 1) a cognitive tool for scaffolding thinking and collaborative writing; 2) documenting collective knowledge advancement and enhancing reflection and collective cognitive responsibility; 3) extending knowledge-transforming space linking concept learning and project inquiry; and, 4) expressive space for developing social literacy and group dynamics. Here are some excerpts: *“Unlike our face-to-face discussions, our KF discussions were recorded for future checks or reflection. At the end of the program, we were all glad to see our evolving TQM understandings at the different stages.” (Student # 3, Group# 2); “Writing on KF may help us think and express ourselves in English. It is good practice. … Reading others’ notes and responding to others, though sometimes difficult, is what we did on Knowledge Forum. We noticed that some groups just put on new notes but interactions were quite limited. We tried to respond to, comment on and build on others. …We valued all contributions to the discussion in democratic atmosphere.” (Student # 2, Group# 2)*

Fourth, connections to the business community facilitated the development of students’ beliefs in learning from the simplistic of knowledge assimilation and application, and further to the sophisticated views of knowledge transformation and collective knowledge building.

All the change dynamics under the four identified contextual themes point to underlying knowledge-building design principles, with particular emphasis on epistemic agency, authentic problems and ideas, collective cognitive responsibilities, and knowledge building discourse. These finding explains why CSCI groups experienced more positive changes than their counterparts in PjBL in conceptions of collaboration, collective advances of conceptual understanding and academic literacy. Meanwhile, it identifies social-technological dynamic as a key contextual theme influencing students' changes. Thus CSCI online discourse was examined to unravel how changes in collaboration and learning are mediated by the interplay of social, cognitive, technological dynamics.

2.3 Characterizing Discourse Patterns and Processes

To understand how student engagement in computer-supported inquiry-based learning would promote student's knowledge advances and understanding, contrastive group analyses (group of 5) were conducted; eight groups were selected based on group project learning performance, yet with a comparable number of entries. All the computer notes were analyzed for illuminating collaborative inquiry processes.

Drawing on the theoretical framework of social and cognitive processes of knowledge construction in general (Hmelo-Silver, 2003), and informed by knowledge building notions of questioning and explanation (Hakkarainen, 2003) and meta-discourse (van Aalst, 2009), empirically induced categories were refined. Five major themes emerge including: (a) information-processing in concept learning, (b) information-processing in project inquiry, (c) metacognition, (d) question-explanation and (e) social dynamics (see Table 2). All the notes from the selected groups were multiple-coded based on the scheme illustrated in previous section. To obtain inter-rater reliability, Cohen's Kappa was computed. All the Cohen's Ks in this study were above.75, indicating a very good coding inter-reliability.

Table 2 *Online Notes Coding Categories and Definitions*

Coding categories	Definition
1. Information processing – conceptual learning	
Knowledge telling	Copy and paste from text with little processing
Elaboration	New info treated problematic for elaboration
Collective advance	Evaluate information from different perspectives to advance collective understanding
2. Information processing – project inquiry	
Surface task-based	Take project simply as completion of several mini-tasks
Information-sharing	Provide useful information for project work
Problem-solving	Design project as a problem-driven inquiry into a real business context
3. Epistemic Questioning	
Factual / clarification	Questions on basic facts, literal meaning of a sentence, for help or for clarification
Identifying inconsistencies	Questions identifying inconsistencies for explanation
Explanation-seeking	Questions raised for deep explanation or to seek to address the problem
4. Epistemic Explanation	
Simple claim	Give opinion without explanation or with irrelevant cut-and-paste information
Elaborated explanation	Make a claim supported with reasons, evidence, and examples
Meta-Explanation	Further explanation synthesizing different view(s) in the previous discussion
5. Metacognitive Processing	
Metacognitive-individual	Checking own progress and understanding; Identifying changes by reflecting on understanding and actions;
Co-regulation	Control and adapt strategies as a result of interactions with group members; learn in context with others

Collective regulation	Check ongoing project progress and communal understanding; identify changes and emerging key problems from the discussion discourse; synthesize different ideas and generate new coherent understanding
6. Social dynamics	
Rapport-building	Statements for building up rapport
Facilitating discussion	Statements to facilitate discussion
Making contributions	Statements to make suggestion to project inquiry

To generate deeper understanding of productive discourse moves and online social cognitive dynamics, contrastive analysis was conducted between high-performance (HP) and low-performance (LP) groups. Quantitative analyses were conducted, based on the coding scheme for discourse moves, to examine the groups' differences in the identified discourse moves. To ensure valid comparison of discourse moves across groups, the frequency occurrence was divided by the total number of group notes written to reveal the percentage of notes in which each discourse move occurred. Due to limited sample size and for coherence, the group comparison analyses were only conducted on the higher-level discourse moves for the major categories, namely collective advance in conceptual information processing, problem solving in project inquiry information processing, higher-level epistemic questioning (a combination of questions of identifying consistency and explanation-seeking), higher-level epistemic explanations (including both elaborate explanation and meta-explanation), collective regulation in metacognition and making contribution in social dynamics.

Analyses were made by assigning each individual his or her group percentage score for the various high-level discourse types. Group percentage scores were employed based on the notion that discourse moves emerge collectively in a group rather than belonging to each individual (Stahl, 2006). Mann-Whitney U tests were conducted to examine differences between HP and LP groups on the above-mentioned six discourse moves. Significant group differences were detected in collective advance (Z = -4.83, p <0.01), problem-solving inquiry (Z = -4.22, p <0.01), higher-level epistemic questioning, higher-level explanation (Z = -4.83, p <0.01), collective regulation (Z = -4.83, p <0.01), and making contribution (Z = -4.83, p <0.01).

The results indicate that HP groups more actively involved in knowledge construction and in collaborative project inquiry than were LP groups. HP groups used more individual and social metacognitive strategies to advance collaborative inquiry. During the inquiry process, they showed a higher level of collective epistemic agency by posing explanatory questions and providing elaborated explanations or meta-explanations. Moreover, they demonstrated lively social dynamics conducive to collective knowledge building.

3. Discussion and Conclusion

The study examined the role of CSCI environment for promoting collaborative inquiry in the context of project-learning and identified the nature of change dynamics interacted with social-metacognitive, pedagogical and technological factors and the scaffolding role of principle-based use of technology.

Earlier work related to this research project reported the observed effect of the designed CSCI environment on changes in learning approaches, conceptual understanding and literacy skills. The study extended the inquiry line by identifying the design effect on changes in student conceptions of collaboration. Drawing on focus group interview data, it provided deeper analysis on the change dynamics under the four contextual themes-epistemology, pedagogy, technology and community. More important, all the dynamics reflected the importance of knowledge-building principle-based design emphasizing collaboration. Similarly, online discourse study corroborated evidence highlighting students' engagement in high-level socio-metacognitive discourse moves and meta-explanation contributes to collective advancement of communal understanding and higher-level collaborative inquiry strategy use. These findings from different data sources converged that student changes in CSCI environment are consistent with knowledge building model and epistemology. CSCI design informed by knowledge building principles is evidenced to foster students' development in sophisticated conceptions of collaboration and epistemological beliefs as well as higher-level collaborative inquiry strategy and to cultivate meta-explanation discourse that advanced both individual and collective conceptual understanding.

This study contributes to current literature of promoting student learning and collaboration in CSCL environment in higher education by investigating both the impact of innovative instructional practices and the change dynamics in relation to the innovative learning experience. Moreover, the study highlights the alignment between design, cognition, technology and context. Drawing on knowledge-building principles, this study designed a CSCI environment using Knowledge Forum to enhance collaborative knowledge building inquiry, examined the design effect on learning and collaboration, and investigated contextual change dynamics through analyzing student experience and online collaborative inquiry discourse.

This study has important pedagogical implications. It documents the effectiveness of a principle-based knowledge-building approach to designing an innovative model within the social cultural context of higher education in China. Pedagogically, it sheds light on how social constructivist learning theories can be transformed to promote changes in conceptions and strategies of collaboration and achieve both individual and collective gains among Chinese tertiary students. It provided an example of knowledge building with positive effects in a different domain of business study in a new cultural context. Further investigation will include deep analyses of connected KF discourses at different phrases to unravel how collective conceptual understanding was advanced and academic literacy was appropriated mediated by social-technological, cognitive and linguistic dynamics.

References

[1] Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71*, 148-149.

[2] Eastman, J. K., & Swift, C. O. (2002). Enhancing collaborative learning: Discussion boards and chat rooms as project communication tools. *Business Communication Quarterly, 65*, 29-41.

[3] Hämäläinen, R., Manninen, T., Järvelä, S., & Häkkinen, P. (2006). Learning to collaborate: Designing collaboration in a 3-D game environment. *Internet and Higher Education, 9*(1), 47-61.

[4] Koschmann, T., Hall, R., & Miyake, N. (2002). *CSCL2: Carrying forward the conversation.* Mahwah: NJ.

[5] Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of Educational Research, 74*, 557-576.

[6] Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal Education in a knowledge society* (pp. 67-98).

[7] Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, Mass: MIT Press.

[8] Sun, Y. Q., Zhang, J., & Scardamalia, M. (2010). Knowledge building and vocabulary growth over two years, Grades 3 and 4. *Instructional Science, 38*, 247-271.

[9] van Aalst, J., & Chan, C. (2007). Student-directed assessment of knowledge building using electronic portfolios in Knowledge Forum. *The Journal of the Learning Sciences, 16*(2), 175-220.

[10] Zhang, J., & Sun, Y. Q. (2011). Reading for idea advancement in a Grade 4 knowledge building community. *Instructional Science, 39*, 429-452.

Reasoning and Reformulating for Linguistic Knowledge Improvement: A Comparative Case Study in a CSCL Classroom

Wenting XIE*, Wenli CHEN, Yun WEN & Cheekit LOOI
Learning Sciences Lab, Nanyang Technological University, Singapore
*wenting.xie@nie.edu.sg

Abstract: This paper reports a comparative case study to explore the discrepancies in learning outcomes attained by two student groups in an identical CSCL activity in the language classroom and the differences in group behaviors that contributed to these discrepancies. From micro-analysis, group behaviors that are desired for language development are identified. The findings can inform future pedagogical and technological design to improve language learning in classrooms.

Keywords: CSCL, language learning, comparative study, collaborative dialogue

Introduction

According to Swain, Brooks & Tocalli-Beller [8], Collaborative Dialogue (CD) where learners co-construct language or knowledge about language to solve complex linguistic problems is a legitimate source for language learning. This validates CSCL intervention into language classrooms. This study examined the learning behaviors of two student groups in an identical CSCL activity and investigated their impacts on group learning outcomes. This comparative analysis can inform future learning design.

1. Research Context

This study is part of a 3 year project in a secondary school where researchers and teachers co-design Group Scribbles (GS) (refer to [2] for GS descriptions) supported collaborative learning to promote L1 learning. A Grade 2 class (22 students, randomly distributed into three 4-student and two 5-student groups) participated in the 1st cycle. The two groups each consisted of 4 students (Figure 1). As indicated by the school English test scores before GS intervention, these two groups were equally competent in L1 ($t=-1.050$, $P= .334< .05$, G1: $M=36.5$, $SD=3.69$; G2: $M=39.3$, $SD=4.35$). In GS lessons, group collaboration occurred in dual-interactional spaces, F2F and online. Each student was given a Macbook with GS 2.0 and seated in physical proximity. The GS activity reported was the 2nd one in the <u>Persuasive Writing</u> lesson where students worked together to plan an argumentative essay on the topic of Cyber Bullying. This activity, followed the 1st GS activity on *Content Generation*, focused on *Organizing and Linearizing* the contents produced and pooled [1] (Table 1).

Table1. Activity Design

Activity	Description	Time
Introduction	Teacher shows a video clip about Cyber Bullying	5 mins
Content Generation	Intra-group interaction: brainstorm to mine possible ideas\ arguments \ examples related to Cyber Bullying	10mins
	Inter-group interaction: Gallery walk visit other groups' boards for more ideas	5 mins
Content Organization & Linearization	Intra-group interaction: 1) select, categorize, synergize, and arrange contents 2) devise thesis \topic sentences	20mins
	Inter-group interaction: Presentation 1) present group work ; 2) offer constructive comments;	10mins

Group 1

Fiona	Kim
Michael	Peter

Group 2

Richard	Jack
Ada	Tom

Figure 1. Group seating arrangements

2. Analytical Framework

The data included group artifacts, post-lesson interview transcripts, group audio\video transcripts, student Morae transcripts. In transcripts, student GS act\verbal talk was comprehensively and chronologically documented. Group learning behaviors, perceptions and outcomes mined in G1 and G2 were inclusively mined and put into comparison.

2.1 Comparing Learning Outcomes in Group Work

The learning outcomes included group final artifacts and Linguistic Knowledge Improvement (LKI). The former concerned with whether the artifacts constructed were: 1) suffice to solve the linguistic problem; 2) grammatically accurate and socio-linguistically appropriate. LKI was measured by the instances of constructing understanding on linguistic concepts (including grammar, vocabulary, pragmatics and stylistics) that emerged in group language. As language occurs in collaborative problem-solving is the spontaneous verbalization of cognitive processes, turning them into objects for analysis [3], the language produced in G1 and G2 (spoken and written), was exhaustively coded to mine LKI (Table 2). One turn in verbal talk or GS text(s) that expressed one idea was coded as one unit.

Table 2 Linguistic knowledge improvement

Description	Example
Verbal talk\GS texts in the form of a statement that contained explicit linguistic knowledge (usually with technical terms), including: 1) Grammar 2) Vocabulary 3) Stylistics	1) (Some students are able to deal with the problem, its deal or deal with?) --Deal with, deal with the problem. 2) --Overlook means you just ignore. Oversees means take charge. 3) --A thesis statement. It's the main stand. It's the main stand that you would take.

2.2 Comparing Social Interactions in Group Work

Social interactions are mainly mediated through language. The mediating function of language is identified as: 1) a cognitive tool for meaning making; 2) a social tool for communication [8]. In CSCL environments, language mediated interactions construct two pivotal dimensions of learning: 1) social\psychological dimension, which relates to social-emotional aspects of group forming and group dynamics; 2) educational\cognitive dimension, which relates to group learning [4]. CD embedded in cognition-related interactions is the very source for language development. In a CSCL classroom, social interactions can go for multiple purposes. The fact that classroom learning is a zero-sum game for the limited physical time makes CD more cherished from the perspective of language learning. To explain discrepant group learning outcomes and perceptions, we compared the social interactions occurred in G1 and G2 (Table 3). One turn in conversation\ one GS act (publishing\withdrawing GS posts on group public board) was coded as one unit.

Table 3 Social interactions in group work

Category	Description	Example
Collaborative Dialogue	Interaction units for solving the linguistic task, including: 1) formulating the required linguistic form; 2) reflecting on the linguistic form; 3) consolidating\constructing linguistic knowledge; 4) pooling and organizing ideas\contents;	1) --Cyber bully has negative impacts on the-. (formulating the thesis statement) 2) --What? What is “brokers”? 3) --Overlook means you just ignore. Oversees means take charge. 4) --Are we gonna talk about the community as a whole or are we talk about individuals in the community?
Task Coordination	Interaction units for forming task strategy and regulating group working process, including: 1) distributing roles\work 2) managing group working progress 3) technology related issues.	1) --"Definitions and examples of cyber bullying", ok, I will handle definition. 2) --I think the first, the one… ok, anyway, put it up put it up. We are too slow. Ok, full stop. Ok, now let's talk about the examples. 3) --You need a bigger piece of scribbles.
Group Forming	Off-task interaction units including: 1) topic related jokes 2) discussion on unrelated topics	1) --Well, I feel so cyber-bullied now. 2) --Woo, The last time I ever used my Safari it was like last year.

2.3 Comparing “Linguistic Knowledge Improvement Trigger” in Group Work

LKI enhances language proficiency and are pursued in “advanced” language learning [6]. CD encourages but not ensures such improvements. LKI should be “triggered”, i.e. learners “notice” the problematic linguistic forms [7]. Only when learners’ attention is drawn to the linguistic problems will they “reflect on” these problems and then create knowledge to solve these problems. We call CDs that can attract and divert group attention to those problematic language “Linguistic Knowledge Improvement Triggers” (LKIT).In observation, we noted students behaved differently when confronted with problematic linguistic forms, which in turn induced different responses. In the third level of analysis, we compared the number of LKITs emergent and investigated the distribution of different types of LKITs (Table 4).

Table 4 Linguistic knowledge improvement triggers (LKIT)

Category	Description	Example
Commenting	Collaborative dialogues that contained the questioning and\or rejecting of the whole\ a part of previous linguistic form.	--Why why why it is “overlook”, then “it become part of the school's norm”? I don't get what you mean.
Reasoning	Collaborative dialogues that contained justification for the grammatical inaccuracy and socio-linguistic inappropriateness of the whole\ a part of previous linguistic form.	--(But if you write that) it's kind like you say- You know it's ok to cyber bully-Because you are teaching them how to handle-
Reformulation	Collaborative dialogues rephrased or redevised the whole\a part of previous linguistic form.	--(Cyber bullying has a negative impact to schools and the school should take action.) --On school. On the school community.

3. Results & Discussion

3.1 Discrepancies in Group Learning Outcomes

There was not much difference in the group final artifacts between G1 and G2 in terms of completeness and correctness. Each group had constructed 1 thesis, 4 topic sentences and 1 concluding statements. In G1, all these were in complete sentences, while in G2, one “topic sentence" was actually a "topic phrase". As requested, both groups produced at least one example\fact for each topic sentence. Though in total G2 had generated 4 more examples\facts than G1, but repetition plagued. After combing similar ones, G2 achieved only one more. As for grammatical accuracy and socio-linguistic appropriateness, G1 and G2 performed equally. However, significant discrepancy was observed via Chi-square test in the amount of LKI occurred in G1 (43) and G2 (9), with G1 outperforming G2 ($p= .000$, $p< .01$) (Table 5). Via collaboration, G1 had improved their understanding on linguistic concepts concerning stylistics (e.g. definition of *Thesis Statement*?) and vocabulary (e.g. the differences between *ability, skills, knowledge, maturity*). Though G2 students also expressed doubts on the functions and features of the Thesis Statement, these doubts were not solved. It was concluded that G1 achieved better learning outcomes compared to G2.

3.2 Discrepancies in Group Social Interactions

There were 211 interaction units observed in G1 and 208 in G2.Though the total numbers were almost equal, differences were mined concerning the nature of interaction. G1 highly engaged in cognitive activities (Cognitive: 99.1%), G2 in social-emotional interactions (Social: 21.6%). And G1 were more engaged in CD ($p= .000$, $p< .01$) and less in task coordination ($p= .000$, $p< .01$) (Table 5). Qualitative analysis of group learning processes revealed that G1 and G2 approached the collaborative task differently, G2 counting on "cooperation" while G1 on "collaboration". In G2 labor was divided and each member was responsible for a portion of the problem solving. Yet in G1 mutual engagement to solve the problem together with good coordination was achieved. This was further validated by how G1 and G2 perceived their group work in the interview. G1 found the group learning experience useful as "ideas were shared, pooled and improved in our group" while G2 were not that positive. They felt group work was just "a compiling of individual work to finish the task". Among the 56 task coordination units observed in G2, quite a lot were "role distribution" (e.g. *Ok, I will handle definition.*) and "progress" statements (e.g. *Ok, there we go. Ok I have done my job.*) on individual bases. In G2, collaborative task solving was a 3-staged iterative process on individual basis ("problem division-individual problem solving-individual solution aggregation"). As G2 focused on paralleled individual problem solving without paying attention to others, more regulation was needed to avoid repetitive\overlapping work, resulting in the commonly occurring inquiries where one asked about others' progress (e.g. *Are you devising already?*). Yet in G1, nearly all coordination was at the group level, either in directing group attention (e.g. *Ok, the definition and examples of cyber bullies…*) or monitoring group progress (e.g. *…put it up put it up…we are too slow.*). Moreover, in cooperation, students are only responsible for a specific piece work. This constrains their engagement with the whole task. Though wanting to contribute to other parts of the problem, they are often reluctant to do so as others may not appreciate or even get annoyed. This was observed in our case. In G2, when Richard posted on the "definition of cyber bullying", Tom, who was supposed to take care of this point, explicitly expressed his dissatisfaction (e.g. *Oh stop putting stuff on my area...*). This not only restricted one's engagement with the task but also left more time for Off-task interactions.

3.3 Discrepancies in Linguistic Knowledge Improvement Trigger

Altogether 69 LKIT were observed in G1, and 34 in G2. When the total number of CDs produced was taken into consideration, there was no difference in the number of LKIT occurred (Chi-square test: $P= .269> .05$). This showed that both groups were equally sensitive to linguistic problems. Yet, G1 and G2 handled these problems differently (Table 6). In G2, the most frequent LKIT was pure Commenting (50%), while that was significantly fewer in G1 (Chi-square test: $P= .002< .01$). In G1 Commenting LKITs were often accompanied with Reasoning. Unlike G2 who solely announced the problem area without further actions, students in G1 stretched their linguistic knowledge to defend for\against a certain linguistic form they were uncertain of. Besides discussing and reflecting on the problematic language, G1 were also more willing to offer tentative solutions to fix the problem (In G1, the most frequent LKIT was Reformulating, accounting for 42%). These extra efforts made in G1, though challenging and exhausting, were worthwhile as they brought about better harvests.

Table 5 Descriptive analysis for group social interactions

Group	Cognitive Dimension		Social Dimension	Total
	Collaborative Dialogues	Task Coordination		
G1	86.3%	12.8%	0.9%	100%
G2	51.9%	26.9%	21.2%	100%

Pearson Chi-Square	Value	df	Sig. (2-sided)
Collaborative Dialogue	57.942	1	.000
Task Coordination	13.159	1	.000

Table 7 Descriptive analysis: Linguistic knowledge improvement trigger

Group	Com	Rea	Ref	Com+Rea	Com+Ref	Com+Rea+Ref
G1	20.3%	5.8%	42.0%	21.7%	5.8%	4.4%
G2	50%	3.0%	32.4%	11.6%	3.0%	0%

4. Conclusion

Through comparative analysis, the desired group behaviors that can bring about positive perceptions and outcomes in CSCL language classrooms are identified. As this study was exploratory in nature and specific in context, further experimental studies are needed.

References

[1] Coirier, P., Andriessen, J. E. B., & Chanquoy, L. (1999). From planning to translating: the specificity of argumentative writing. In J. E. B. Andriessen & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 1–29). Amsterdam: Amsterdam University Press.

[2] Chen, W., Looi, C.-K., & Tan, S. (2010). What do students do in a F2F CSCL classroom? The optimization of multiple communications modes. Computers & Education, 55(3), 1159–1170. doi:10.1016/j.compedu.2010.05.013

[3] Gutiérrez, X. (2008). What does metalinguistic activity in learners' interaction during a collaborative L2 writing task look like? *The Modern Language Journal, 92*, 519-537.

[4] Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior, 19*(3), 335-353. doi:10.1016/S0747-5632(02)00057-2

[5] Swain, M. (2002). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.

[6] Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In Heidi Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). New York: Continuum International Publishing.

[7] Swain, M. (2010). Languaging: Private speech and collaborative dialogue. In M. Swain, P. Kinnear, & L. Steinman. (Eds.), *Sociocultural theory in second language education: An introduction through narratives* (pp. 34-50). New York: Multilingual Matters.

[8] Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics, 22*, 171-185.

Method for evaluating discussion status in online text discussions using network analysis

Kanako NIKI[a*], Takehiro FURUTA[b], Takako AKAKURA[a], Takahito TOMOTO[a], Yuri NISHIHORI[c] & Keizo NAGAOKA[d]
[a]*Tokyo University of Science, Japan*
[b]*Nara University of Education, Japan*
[c]*Sapporo Otani University, Japan*
[d]*Waseda University, Japan*
*kanako-niki@ms.kagu.tus.ac.jp

Abstract: We analyze text discussions with the aim of using statement data to assess the status of an in-progress discussion. We assume convergent discussions with three discussion status types: Publication, Active, and Convergent. Then we propose an automatic method for evaluating discussion status by network analysis. Experimental evaluation indicates that the method is useful. We also achieved positive results in comparison with manual analysis.

Keywords: CSCL, Discussion Status, Network Analysis, Time-Series Analysis

Introduction

Analysis of discussion status from text is valuable when designing text-based CSCL. We analyze text discussions with the aim of assessing the status of in-progress discussions. We develop an automatic method for evaluating discussion status using time-series analysis.

There are two main approaches to analyzing statements in discussion: content analysis (CA) and social network analysis (SNA) [1]. CA codes statements according to statement type, focusing on contextual meaning [2]. SNA is a method for determining the relationship structure of components in various objects [3]. Erlin et al. [1] introduced research about CA and SNA, and argued that their integration provides a scientific and systemic way to analyze the quality of discussion. SNA can be applied to evaluate participants' roles by analyzing statement character [4]. In this paper, we apply SNA for such evaluation, and examine the efficacy of our method in comparison with CA-based methods. We also construct a network that indicates statement relativity based on discussion statement data.

Our objective is to propose an automatic analysis method for evaluating discussion status by SNA. We assume convergent discussions that form a conclusion, and that there are at least three status types: *Publication*, *Active*, and *Convergent*. As a preliminary development of the method, we propose a method to detect these status types. We analyze the experimental data by using our method and by manually applying CA, and determine the discussion status. Finally, we quantitatively assess our method by comparing its results.

1. Method for Evaluating Discussion Status Using Statement Data

1.1 Discussion Status

We assume an online discussion in text-based CSCL. Statements are extracted from a chat log along with input time and person. Our target is not divergent discussions that gather various opinions, but rather convergent discussions that form conclusions.

We assume that convergent discussions contain at least three types of discussion status (Figure 1): participants first describe their own ideas (Publication), and then they examine and compile these ideas (Active, and Convergent). We examine relational ties between statements with related content to detect discussion status. Participants disclose their own ideas in the Publication status. Therefore instructions to gather ideas and informative statements are made in this status. Back-channel feedback to these statements is also given in the Publication status, so at this time there are more new ideas than relational ties between statements. Participants examine their own ideas in the Active status, and thus provide questions and answers, as well as statements that connect with other statements. The number of relational ties between statements increases in the Active status. Participants compile their own ideas in the Convergent status, and thus make comments about their interaction and instruction that promote compilation. Thus, statements with many connections to previous ones are added at this time. Our objective is to detect changes in discussion status in progress. We seek alterations in discussion status by analyzing networks consisting of incrementally formed statements.

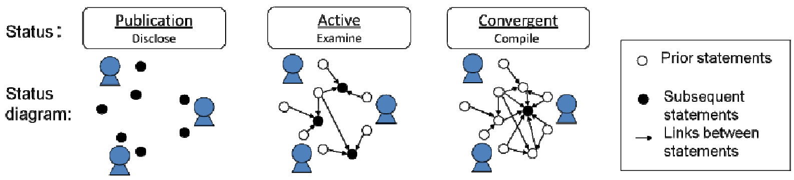


Figure 1 Discussion Status

1.2 Time-Series Analysis

This paper analyzes changes in discussion status by time-series analysis. We adopt statements by increments continuously to construct networks (statement networks), in which the nodes are statements and the links are the relational ties of statements. After making a statement, we adopt a given number of consecutive statements including that statement as a component of statement network, and construct statement network in continuity. Figure 2 shows an example of statement networks consisting six statements. In this manner, we analyze by calculating indexes of statement networks consisting consecutive statements as often as statement is made. We define relational ties according to the co-occurrence of words, adopting links between statements that share same nouns and adjectives. A statement network is digraph in which links are connected from a prior statement to a subsequent statement.

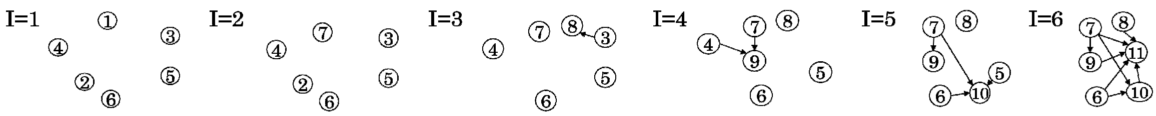


Figure 2 Time-Series Analysis (Number: Statement No., I: Statement Network No.)

To conduct quantitative analysis, we calculate three indexes on statement networks. First, we evaluate the amount of information in statements to detect *Publication*, in which statements present new ideas. Then we count the net number of words in statement networks to evaluate the amount of information in the statements. The net number is counted duplicative words as one word. The more words that statement networks have, the more information they contain. This allows us to detect *Publication*, the status with the most information, by counting the net number of words.

Next we focus on relational ties to detect the *Active* status, in which participants interact to examine their ideas, by calculating network density. Density, the proportion of possible links that are actually present in the network, is an index of the overall network structure [3]. In a statement network, density indicates the number of relational ties between statements, allowing us to detect the *Active* status through the increase in the number of relational ties.

Finally, we search for statements that affect prior statements to detect the *Convergent* status, in which statements concluding prior ones are added. Then we calculate the indegree centrality defined as the number of links directed to the node in the digraph [3]. The more links a node has in a network, the more central the node is [3]. Thus, the value of the indegree centrality of a node in a statement network indicates the amount of influence from the other statements. This allows us to detect *Convergent* status by calculating the indegree centrality of the last node in a statement network. When a statement alters the discussion status from, for example, *Active* to *Convergent*, the network around the statement can be in both statuses. Thus, we assume that statement networks can be in multiple statuses.

1.1 Determining Discussion Status through Social Network Analysis (SNA)

The following describes the method of determining discussion status through SNA. First, we construct statement networks consisting of h statements in a discussion with n statements. A statement network S_k has nodes from the kth statement to the $k + (h - 1)$th statement, and is a member of the statement network set $N = \{S_k \mid k = 1, \cdots, n - h + 1\}$. We use the net number of words, $W(S_k)$, as an index to determine *Publication.* A statement network set *P* whose elements are determined as being in the *Publication* status is defined as

$$P = \{ S_k \in N \mid W(S_k) > \overline{W} \}, \qquad (1)$$

where $\overline{W}$ is the average of $W(S_k)$ of all statement networks. Similarly, a statement network set *A* whose elements are determined as being in the *Active* status is defined as

$$A = \left\{ S_k \in N \mid \frac{D(S_k)}{D(S_{k-1})} > 1 \right\}, \qquad (2)$$

where $D(S_k)$ is the density of S_k. A statement network set *C* whose elements are determined as being in *Convergent* status is defined as

$$C = \{ S_k \in N \mid I(S_k) > \bar{I} \}, \qquad (3)$$

where $I(S_k)$ is the indegree centrality of the last nodes, and $\bar{I}$ is the average of $I(S_k)$ of all statement networks.

2. Experiment

2.1 Outline

In this chapter, we describe the experimental data on which we applied our method, and evaluate our method qualitatively by comparing chat log data. We also evaluate our method quantitatively by comparing the results of applying a CA-based method.

We conducted experiments with the participation of 20 Japanese students. We divided the students into four groups of five participants each. All four groups held a discussion through online text chat for 50 min. The theme of the discussion was the consensus game, the purpose of which is obtaining consensus among group members through discussion [5]. Group members performed an exercise in which they ranked eight items in descending order of importance for survival in a difficult situation [5]. They could use Google Chat as a group chat program and Google Docs Presentation to share information.

2.2 *Results of Proposed Method*

In applying our method to the experimental data, we set the number of nodes h as 35 in a statement network for every group. The 'SNA' column in Table 1 shows the results of the status determination. The number of status determinations in the group 2 discussion was biased, because that group did not have enough time after an extended discussion of policy, and thus decided to determine conclusions by majority vote. Results are therefore biased towards a *Publication* determination. Here we describe the results of the group 1 discussion analysis in detail. Figure 3 shows the statuses of the group 1 discussion. The horizontal axis shows the statement network number, thus progress of discussion. The painted bars describe that network is in the status. The beginning of the discussion was determined as *Publication*, indicating that they were disclosing their own ideas. After that, the number of *Active* statuses increased. Statement network 55 was determined to be in the *Convergent* status. It showed a full flow of discussion: disclosing, examining, and compiling ideas. The early part of discussion showed the discussion cycle defined in Section 1.1, but the latter part of the discussion became jumbled. This illustrates the difficulty of our method in exactly analyzing status along with discussion progress. The discussion status skipped, for example, from *Convergent* to *Active* status when previous statements were mentioned again. We confirmed those statements by checking pertinent sections of the chat log. As a result, our method is effective for evaluating discussion status.

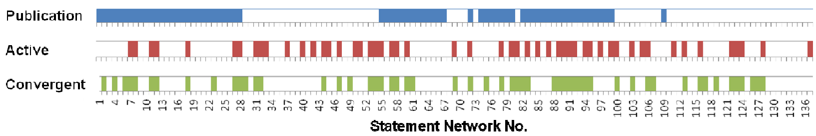


Figure 3 Determinations by our method (Group 1)

2.3 *Comparison of proposed method with CA*

We next compared the results from our SNA-based method with the results from a CA-based method. Following Fujimoto et al. [6], we used six types of statement code for CA: *Information, Instruction, Reflection, Question, Answer*, and *Comment*. This code is an adaptation of Verbal Response Mode Coding System [7], modified to accommodate discussions in Japanese. Three collaborators who did not join the discussions manually coded the chat log data. Coding was decided by majority, and data were coded as *non-codable* when all three opinions differed.

We determined the relationship between this code and discussion status based on the assumptions in Section 1.1. A statement S_k is determined to be *Publication* when there are more statements coded *Information*, *Instruction*, or *Reflection* than the average of those in the entire network. Similarly, a statement S_k is determined to be in the *Active* status when there are more statements coded *Question* or *Answer* than the average of those in the entire network. A statement S_k is determined to be in the *Convergent* status when there are more statements coded *Instruction* or *Comment* than the average of those in the entire network. We examine precision and recall to compare the results. As previously noted, each statement network can have multiple statuses. We therefore separately evaluate each status. Equations (4) and (5) define the precision and recall, respectively, for a status *X*.

$$\mathrm{precision} = \frac{|\{\text{Networks are determined to be } X \text{ by SNA}\} \cap \{\text{Networks are determined to be } X \text{ by CA}\}|}{|\{\text{Networks are determined to be } X \text{ by SNA}\}|} \quad (4)$$

$$\mathrm{recall} = \frac{|\{\text{Networks are determined to be} X \text{ by SNA}\} \cap \{\text{Networks are determined to be } X \text{ by CA}\}|}{|\{\text{Networks are determined to be } X \text{ by CA}\}|} \quad (5)$$

Table 1 Comparison of SNA and CA

	Group 1(137)				Group 2(163)				Group 3(105)				Group 4(316)			
	SNA	CA	Precision	Recall	SNA	CA	Precision	Recall	SNA	CA	Precision	Recall	SNA	CA	Precision	Recall
P	68	69	0.735	0.833	103	71	0.592	0.616	51	66	0.608	0.463	101	111	0.287	0.420
A	52	61	0.423	0.458	59	113	0.542	0.364	43	57	0.465	0.317	95	215	0.516	0.310
C	54	84	0.685	0.440	72	86	0.458	0.384	35	57	0.686	0.421	91	200	0.626	0.285

P:*Publication*, A:*Active*, C: *Convergent* ( ): The number of statement networks

Table 1 shows a comparison of the precision and recall of the results of our method and those of the CA-based method. The ‘CA’ column in Table 1 gives the CA-based determinations. With the exception of group 4, determinations of *Publication* status showed relatively good value. Group 4 had many short statements, making automatic determination difficult by our method. Although values there were poor, the proposed method achieved some positive values for *Active* status. With the exception of group 2, the precision of determining *Convergent* status exceeded 60%. Calculating the feature quantity of networks of consecutive statements in our automatic analysis method is corresponded with manual evaluations to some extent. We could get good values for *Publication* and some positive values for *Active* and *Convergent* status determinations.

3. Conclusion

We proposed an automatic method based on SNA for evaluating discussion status in chat log data, assuming convergent-type discussions with three discussion status types: *Publication*, *Active*, and *Convergent.* We applied time-series analysis using indexes to evaluate each status. Furthermore, we developed a method for determination of discussion status using SNA based on those indexes. Experimental verification using the chat log data indicated the possibility of detecting discussion status, and comparison of the results of CA confirmed that our automatic method corresponded with manual analysis to some extent. In future research, we plan to develop a method for evaluating other discussion status from many perspectives. Furthermore, we will consider what constitutes a smooth discussion when this method is established. We expect that we will be able to immediately give suitable support in CSCL for each discussion status through real-time discussion evaluation.

Acknowledgements

This research was supported by a Grant-in-Aid for Scientific Research (B) (# 24300291) from Japan Society for the Promotion of Science (JSPS).

References

[1] Erlin, Yusof, N., & Rahman, A. (2008). Integrating content analysis and social network analysis for analyzing asynchronous discussion forum. *Proc. of ITSim 2008*,1621-1628.

[2] Krippendorff, K. H. (1980). *Content analysis: An introduction to its methodology*. Sage Publications.

[3] Wasserman, S., & Faust, K. (2009). *Social network analysis: Methods and applications*. Cambridge University Press.

[4] Niki, K., Furuta, T., Akakura, T., Tomoto, T., Nishihori, Y., & Nagaoka, K. (2011). Method for evaluating participants’ roles in online text discussion using network analysis. *Proc. of ICCE 2011*, 199-203.

[5] Hall, J. (1971). Decisions, decisions, decisions, Analysis. *Psychology Today*, *5*, 51-54, 86, 88.

[6] Fujimoto, M., Murayama, A., & Daibo, I. (2004). What ensures a small-group discussion efficient? : Identifying speaker-role in a small discussion. *IEICE Technical Report*, 104(198), 33-38.

[7] Stiles, W, B. (1992). *Describing Talk: A taxonomy of verbal response modes.* Sage Publications.

A Training System that Analyzes the Behavioral Factors of Children with Autism

Yuji Fujii[a], Hisashi Ogawa[a], Junichi Kakegawa[a], Miyuki Takano[a], Koichiro Morihiro[a]
[a] *Hyogo University of Teacher Education, Japan*
m11106c@hyogo-u.ac.jp

Abstract: Teachers of autistic students must observe the overt behaviors of these students and acquire the necessary skills to analyze the causes of these behaviors. Teachers can build a low-stress environment by considering the characteristics of each student, which can be revealed by analyzing their behavioral factors. However, it is not easy for less-experienced teachers to perform such analysis. Therefore, we are developing a training package that aims to help less-experienced teachers acquire the observation and analytical skills necessary to accurately identify the behavioral factors of each student with autism. We carried out our preliminary experiment by using a prototype system based on the hypothesis that it is possible for even an less-experienced teacher to make correct assumptions about behavioral factors such as stress. The results of our preliminary experiment support this hypothesis.

Keywords: autistic person, stress, training package, behavioral factors, behavioral analysis

Introduction

This study focuses on children who have been diagnosed with low-functioning autism spectrum disorder (ASD). ASD is a neural development disorder characterized by impaired social interaction and communication; those characteristics are listed on the diagnostic criteria of the *DSM-IV* [1]. Problem behaviors, such as self-injury, include actions that injure the self or others, such as picking at the skin, hand biting, and head banging; these behaviors are almost always caused by ASD's communication difficulties [2]. The manifestation of problem behaviors is sometimes compared to an iceberg, with the tip representing the overt behaviors (i.e., the results) and the submerged portion representing the underlying differences and impairments (i.e., the factors that create the results) [3]. Therefore, teachers must observe the behaviors of autistic students and analyze the factors that cause those behaviors (hereafter referred to as "behavioral factors"). Problem behavior, which is seen at least occasionally in each person with autism, is caused by behavioral factors in many cases, such as minor changes of environment and surrounding circumstances. To build a low-stress environment, the teacher must take into account the characteristics of each student with autism and must guess at the behavioral factors that influence these characteristics as accurately as possible. However, it is difficult for less-experienced teachers in special-support schools to acquire the skills to accurately guess behavioral factors that are not presented.

Thus, we are developing a system to accompany a training package that will help less-experienced teachers to acquire the necessary skills for observing the behavior of students with autism and to analyze the behavioral factors of those students. Although many behavior-analysis training packages help users to obtain knowledge and to acquire analytical skills, they require users to undergo multiple rounds of self-study or training; as a result, these other packages are not necessarily efficient. Therefore, establishing the framework for an efficient training package is enormously significant.

We have developed a case conference support system equipped with a balloon-type video-annotation function (hereafter referred to as “video annotation”) [4]. It has been suggested that it is possible for viewers to obtain awareness of certain behavioral factors by reading video annotations, inserted by experienced teachers, that express the likely intentions of people with autism, and by comparing these annotations with those inserted by other teachers. However, although this technique is effective in terms of observing certain behaviors, it is difficult to analyze the behavioral factors solely through inserting individual video annotations. In discussions between less-experienced teachers, the topic of behavioral factors also appeared; however, the contents of those discussions were scattered and did not deepen. It was observed that less-experienced teachers tended to gain awareness of behavioral factors only from referring to the annotations from expert and experienced teachers. Whereas it is effective for teachers to share information among themselves regarding their awareness of the behaviors of specific autistic students and to reflect this knowledge in their lessons, this concept cannot be applied to the field of knowledge acquisition or obtaining teaching skills because the characteristics of each autistic person are different. Thus, although it is necessary to carry out training intended to give teachers experience in guessing behavioral factors, it is difficult to accomplish this using only methods outlined by previous studies.

Therefore, we conducted a survey to discover why less-experienced teachers cannot correctly infer behavioral factors from the overt behaviors of students. As a result, it was discovered that students training to be teachers tend to be unable to identify a correspondence between problem behaviors and stress; also, student teachers interpret behavioral factors based on their own experimental rules. Hence, we believe that the less-experienced teachers had made superficial judgments. We also theorize that less-experienced teachers can infer behavioral factors through by learning to identify potential stress in their students and analyzing student behavior with a view toward stress.

1. Compendium of the Training Package

We propose a training package based on the hypothesis that even an less-experienced teacher can make correct assumptions about behavioral factors such as stress. Steps 1 and 5 of this package use methods established in previous study [4].

Step 1: Participants express their interpretation of the intention of the target child in the video by using balloon-type video annotation. It is expected that less-experienced teachers can become aware of behavioral factors through the activity of expressing their interpretation of these intentions.

Step 2: Participants analyze the scene containing these annotations, with a view toward stress. Through this work, participants are expected to be able to focus on the stress level of the target child.

Step 3: Participants analyze behavioral factors with a view toward stress. The system gives a visual representation of the stress analyzed in Step 2 in the form of a graph; participants describe the common points between high-stress and low-stress scenes.

Step 4: Participants observe the differences between the expert teacher’s stress graph and their own. Then, participants describe the common points among the high-stress and low-stress scenes that were pointed out by the expert teacher. It is expected that participants can reflect on the differences between their methods of analysis and those of the expert teacher.

Step 5: Participants refer to the expert teacher’s video annotations.

The training package is composed of a worksheet and a training system. Herein, we describe the prototype system and the procedures of this training package in detail. The prototype system is a client-server system. The client side can be operated in Windows VISTA/7. Figure 1 shows the prototype system interface.

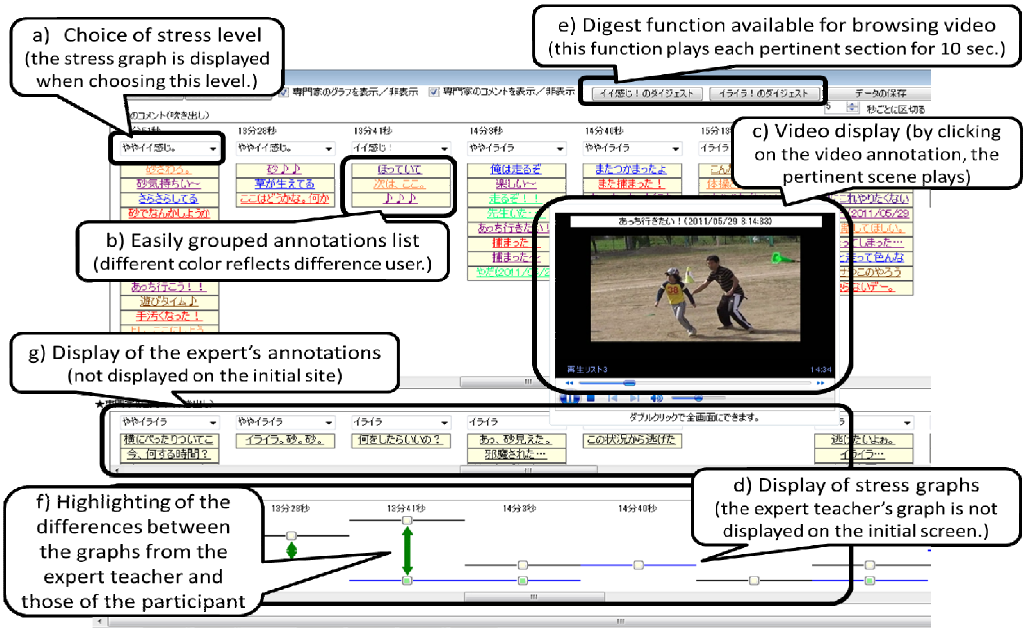


Fig. 1 Prototype system

1.1 Procedures for use of the training package

We developed the training package based on the mode of two-hour exercises using videos of approximately 15 minutes each and a one-hour commentary by an expert. The procedure for this exercise was as follows:

1. Participants insert the annotations in the video. (40 min)
2. The system groups the annotations (which participants have inserted on an hourly basis); participants refer to the video, which is linked by its annotation groups. Then, participants select a stress level. (15 min)
3. Participants describe the differences in their worksheets among their own annotations and the common points among the high-stress scenes and low-stress scenes. (25 min)
4. Participants draw a comparison between their own stress graphs and that of the expert teacher; they describe their burgeoning awareness in the worksheet. (15 min)
5. Participants browse the video digest of the common points chosen by expert teachers between the high-stress/low-stress scenes, and describe the common points in their worksheets. (10 min)
6. The system displays the expert teacher's annotations; the users describe their growing awareness in their worksheets. (15 min)

2. Experimental Methods

2.1 Target video

The target child was a 10-year-old boy who had been diagnosed with ASD and was enrolled in a special-support school. His vocabulary was limited; his method of communication was to use non-verbal instrumental gestures. Many video scenes showed him running around or not participating in class. We video-recorded the target child during one day of school. The edited video was 13 minutes long and was composed of a morning assembly, an arithmetic lesson, and a physical-education lesson during which the target child exhibited problem behavior frequently.

2.2 Participants

Participant A and participant B were graduate-school students who had applied for jobs as special-support-school teachers and had experience teaching in elementary school or junior high school. However, these individuals had not gained practical experience in a special-support school.

2.3 Procedure

We conducted a preliminary experiment to investigate whether less-experienced teachers can correctly interpret behavioral factors in terms of stress. The survey was administered individually. First, we worked with the training package, which was set up according to the prototype system and used the methods we described in section one. However, we did not work on the content of the expert teacher’s lecture. Next, we carried out non-structured interviews and the questionnaire survey.

3. Results and Discussion

3.1 Selecting a stress level

First, Participant A focused on the scene in which the target child touched the sand repeatedly; this participant commented, in her worksheet, “Does he really like to touch sand? I think…probably not.” Then, participant A stated, regarding non-structured interviews, “When I watched the video the first time, I recognized his behavior of touching the sand repeatedly as a fun activity. However, through use of the training system, I began to question my own interpretation of his behavioral factors.” Therefore, it was suggested that participant A can adjust her interpretation by choosing a stress level appropriate to the target child in the video. Similarly, participant B can analyze behavioral factors in terms of stress by observing the differences between his own annotations and those of other teachers. Therefore, the work of selecting a stress level may be effective with a view toward pointing out areas of stress. In addition, it was suggested that participants can acquire awareness of behavioral factors.

3.2 Describing the common points among high-stress and low-stress scenes

In her worksheet, Participant A describes the common points among the high-stress scenes as being “the scene of touching the sand” and “the scene of running around.” In her non-structured interview, she stated, “When this child touches the sand or runs around, he is under stress. Therefore, he overcomes the temptation to get angry.” She was aware of the behavioral factor when she selected the stress level and arrived at her conviction by referring the relevant scene when she described the common point between the high-stress scenes. In her non-structured interviews, she stated, “If the child likes to touch sand, he should continue to touch it” and “I thought that he should figure out the instructions when he has an understanding of what to do next.” Participant B also described the common point concretely. Therefore, it was suggested that the content of the discussion in which participants choose the stress level of the target child leads them to analyze the behavioral factor by describing the common point between high-stress scenes. However, neither participant described behavioral factors by referencing the common point between the low-stress scenes.

3.3 Awareness gained by reference to the stress graph created by an expert teacher

Participant A referred to the expert teacher's stress graph and compared it with her own. However, she has not acquired the necessary awareness to make statements such as "It was unexpected" and "I think that he practices gymnastics in a straightforward manner." Similarly, although participant B strove to gain awareness, as indicated by statements such as "I can understand that touching the sand is not a pleasant sensation for him." this participant did not acquire the same level of awareness as he had in his other description. Nevertheless, both participants strove to acquire awareness by referring to the video digest containing a graph of the high-stress situations and a graph of the low-stress situations (both created by expert teachers) and by describing the common points among the high-stress and low-stress graphs. For example, participant A pointed out behavioral characteristics such as "He jerks away from the teacher in the high-stress scene" and "He looks into the distance in the high-stress scene." She clarified her awareness regarding behavioral factors in statements such as "The high-stress factor is that the teacher allowed the child to exercise forcibly," which specify the common point among high-stress situations. Further, participant A acquires awareness of the common points among low-stress situations: "It is in low-stress scenes that the teacher faces the child and the teacher can hold his interest" and "It is in low-stress scenes that the teacher takes a positive approach toward him and praises his behavior." Similarly, participant B gained awareness, as he can appreciate awareness of the importance of analyzing behavioral factors objectively by observing the child's actions before and after the behavior highlighted in the video. Therefore, participants are unable to understand the true meaning of experts and not able to gain awareness by only comparing the stress graph However, this result suggests that participants can gain awareness and a new point of view by referring to the video digest of high-stress and low-stress scenes.

4. Summary

In this article, we report the results of our preliminary experiment using a prototype system. This system was developed based on the hypothesis that less-experienced teachers learn to accurately infer behavioral factors in terms of stress by being given a framework of stress analysis to help them learn to spot stress occurrence in people with autism. Although our experiment is preliminary, participants were able to gain a great deal of awareness." In future work, we will improve this system based on the results of the preliminary experiment and on its use in actual training sessions.

References

[1] American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders (4th ed.)*. Washington, DC.

[2] Schopler, E. (2004) *Parent Survival Manual: A Guide to Crisis Resolution in Autism and Related Developmental Disorders*, New York: Plenum Press

[3] Mesibov, G. B., Shea, V., & Schopler, E. (2004) *The TEACCH Approach to Autism Spectrum Disorders (Issues in Clinical Child Psychology)*, New York: Springer-Verlag Press

[4] Ogawa, H., Fujii, Y., Kakegawa, J., Takano, M., Morihiro, K. (2011) Practical Consideration of the Support Function for Analyzing the Behavioral Factors of People Suffering from Autism Spectrum Disorder in a Video Annotation System", *Proc. of ICCE* 2011

A Case Study of Critical Thinking Behavior in an Online Collaborative Inquiry

Jing LENG
Faculty of Education, The University of Hong Kong, China
lengjing@hku.hk

Abstract: This paper reports on a case study to investigate whether the scores obtained from various critical thinking instruments were correlated to the extent to which student exhibited critical thinking behavior in authentic problem-solving situations. On the one hand, students' critical thinking skills and dispositions were examined using three instruments. On the other hand, student participants were introduced to use an online discussion environment to conduct collaborative inquiry. The students' online discourse reflects their engagement in authentic problem-solving practices as they tackled the learning tasks assigned by the teacher. Content analysis was performed to identify instances of critical thinking behavior in the online discourse using Perkins and Murphy's framework. The results suggest that students' context-free critical thinking skills and dispositions have been improved after the collaborative inquiry. And the context-specific critical thinking test is useful as a predictor of students' critical thinking behavior exhibited in authentic contexts.

Keywords: collaborative inquiry, critical thinking instruments, critical thinking behavior

Introduction

Critical thinking is often considered as requisite ability for citizens in the 21st century [1]. The concept of critical thinking has maintained a strong focus on the micro-skills related to critical thinking [2, 3], which have been criticized as inadequate. Recent research studies give more emphasis to developing habits of mind to think critically, i.e. nurturing the dispositions towards critical thinking [4-6]. Another trend in critical thinking research is the realization that good performance on generic critical thinking skills tests does not guarantee the ability to exercise critical thinking in authentic problem-solving contexts. Existing literature suggests that engaging students' in knowledge building discourse could be a satisfactory and effective way to improve critical thinking [7, 8]. However, limited empirical studies have investigated students' critical thinking while they engage in extended collaborative inquiry. This paper examines students' critical thinking using different instruments, including: (1) two standardized surveys on generic critical thinking skills and dispositions, (2) a constructed response instrument designed to measure context-specific critical thinking skills. Further, this paper explores whether students' scores on the three critical thinking instruments were correlated in any way to the extent to which the students exhibited critical thinking behavior in the online discourse.

1. Literature Review

1.1 Critical Thinking

The conception of critical thinking can be traced back to John Dewey's views on reflective thinking [9]. In the ensuing years, many philosophers have taken interest in developing the conception of critical thinking based on Dewey's ideas [10]. Ennis is one of the most influential philosophers who defines critical thinking as a set of skills and dispositions [11]. Moreover, McPeck points out that critical thinking is always thinking about something, and the instruction for critical thinking should engage students in certain activities with context-specific knowledge [12]. In line with the trends in critical thinking development, educators have shown growing interest in fostering students' critical thinking skills and nurturing critical thinking dispositions through engaging them in collaborative inquiry in authentic problem-solving contexts.

1.2 Ways to Measure Critical Thinking

This paper reviews some popular instruments of critical thinking and reviewed different ways to measure critical thinking, including multiple-choice tests, constructed response tests, and the online learning discourse. In the following, three ways of testing are introduced and discussed in detail.

In the multiple-choice tests, generic cognitive skills are commonly included, such as inductive or deductive inference [13]. Students are presented with "general scenarios" which do not require specific knowledge. Based on the given passages, they identify the assumptions in what other people said and make judgments on the credibility of the provided sources or other's observation. However, these tests in multiple-choice style are only recognition tasks. So they only examines the evaluative aspects of critical thinking rather than the productive aspects [14]. And they fail to reveal the thinking processes of the students.

In terms of the tests allowing constructed responses, students are asked to respond to a set of semi-structured questions. So this kind of tests contains both the evaluative and productive aspects of critical thinking. The constructed response tests generally involve content, either problems in daily life or context-specific topic. Comparing with multiple-choice tests, these tests are more open-ended. The disadvantage of constructed response tests is that the grading task is difficult and time-consuming. Guidelines and trainings must be provided on making flexible and reasonable complex judgment in grading the test.

The online discourse is a new way to examine a continuous flow of evidence of students' critical thinking behavior that autonomously exhibited during the learning process. The online discussion environment can be designed to support students' collaborative inquiry and sustained knowledge building, such as sharing useful information, exchanging ideas with others upon the specific course topic. Although the online discourse can demonstrate how students may exhibit critical thinking in complex problem-solving contexts, the open-endedness and authentic context make it difficult to measure any critical thinking skills or dispositions.

In summary, these multiple-choice tests can be easily administered and scored. But it is not appropriate to use the multiple-choice tests to assess students' ability to employ several skills together when working on a complex problem. Hence, this kind of test should be only part of the assessment. Comparatively, essay-format tests are considered to provide more leeway in answering which can gather more information on students' productive thinking. For example, in the Ennis-Weir test, students are prompted and pushed to use certain critical thinking skills. But still they are not sufficient to reveal students' critical thinking dispositions in solving authentic problems. So it is argued that critical thinking assessments need to provide opportunities for students to use any critical thinking skills they want to apply in an open-ended yet focused problem situation [15].

2. Design and Method

2.1 Research Context

This paper selected a school which has established a strong curriculum focus within the Integrated Humanities (IH) subject on developing students' critical thinking and designing inquiry-based task to promote children's critical thinking skills. The humanities teacher believed that the use of an asynchronous online environment will enhance the collaborative inquiry process and foster critical thinking. The teacher's class at secondary two was invited to participate in this study. This is the first year for the teacher and his students to use Knowledge Forum® (KF). The thirty-two students were involved in the humanities module and guided to engage in problem-solving inquiries on different topics related to the task of designing a new tourist attraction in Hong Kong. The inquiry extended over a period about 3 months.

2.2 Instrumentation

All participants completed two standardized tests on critical thinking skills and dispositions, and they also filled out a peer critique form which can reflect their critical thinking ability within the subject-matter context. More importantly, instances of critical thinking behavior in the online discourse were used as evidence of students' ability to exercise critical thinking in authentic contexts. The details of the instruments and the framework employed to measure critical thinking are described as following:

1. The adapted Cornell Critical Thinking Test, Level X (CCT-X) [16]

The test measures context-free critical thinking skills, such as induction, deduction, observation, credibility and assumption.

2. The Inventory of Belief and Critical Thinking Disposition (IBCTD) [17]

This instrument is a Likert-scale questionnaire to measure critical thinking dispositions, including (a) systematic and analytic; (b) open-minded and empathetic; (c) intellectual; inquisitive, and (d) holistic and reflective.

3. The Constructed Response Test on Critical Thinking Skills: Critique Form

The peer critique form is a customized version of the Ennis-Weir essay test [18]. This test can measure context-specific critical thinking skills, which includes identifying relevance, appropriate definition, appropriate use of authoritative sources, appropriate reasoning, consideration of different possibilities, and overall strength of the argument.

4. Critical Thinking Behavior as Exhibited through the Online Discourse

Students' online discourse was analyzed to using Perkins and Murphy's framework, which consists of four critical thinking processes: clarification, assessment, inference and strategies [19]. The unit of analysis used to code the online discourse of students was an "idea unit", a segment of discourse containing an identifiable idea or a single coding category. There can be more than one critical thinking process in a note. The inter-rater reliability was conducted on 20% of the total number of notes. The Cohen's kappa was 0.86, indicating a good level of agreement between two raters.

3. Results & Analysis

3.1 Critical Thinking Exhibited in Students' Online Discussions

All of the discussion notes were coded, but not all notes could receive a code. Overall, 591 of the 1017 (58% of the total number of notes) were identified as containing evidence of critical thinking processes, resulting in a total of 652 idea units being coded. 21 out of 32 students were found to exhibit all the four critical thinking behaviors during the online inquiry. In general, the critical thinking skill most frequently demonstrated by the students was clarification, while assessment was demonstrated least. Table 1 presents the findings broken down by stage.

Table 1 *Distribution of idea units with critical thinking processes in different views*

	Clarification		Assessment		Inference		Strategies		Total idea units	
	N	%	N	%	N	%	N	%	N	%
Critical thinking behavior	341	52%	72	11%	142	22%	97	15%	652	100%

3.2 The Relationship between the Critical Thinking as Measured by Three Instruments and Critical Thinking Behavior exhibited in the Online Discourse

Table 2 shows the results of correlation analyses between the scores obtained from three critical thinking instruments and the number of units coded as critical thinking processes. There was a stronger correlation between the post-test scores for CCT-X and IBCTD test with the four exhibited critical thinking behaviors in the discourse. Take the CCT-X test for example, while no significant correlation was found between the pre-test score and the four critical thinking behavior in the discourse, the post-test score was significantly correlated to the number of clarification (r = 0.42, p < 0.05), assessment (r = 0.50, p < 0.01), and strategies (r = 0.59, p < 0.01). On the other hand, the contextual critical thinking test (Critique Form) showed a significant correlation with critical thinking behavior, such as the number of processes of inference (r = 0.41, p < 0.05), and strategies (r = 0.48, p < 0.01). Therefore, these findings suggest that Critique Form is a useful predictor on students' critical thinking behavior exhibited in the discourse. And engaging students in critical thinking in authentic problem solving can potentially improve their context-free critical thinking skills and dispositions.

Table 2 *Correlations between scores of three critical thinking instruments and critical thinking behavior exhibited in the online discourse*

	Online Critical Thinking Behavior			
Total score	clarifications	assessments	inferences	strategies
CCT-X pre-test	0.30	0.13	0.22	**0.43***
CCT-X post-test	**0.42***	0.35	**0.50****	**0.59****
IBCTD pre-test	0.23	-0.09	0.24	0.30
IBCTD post-test	**0.38***	0.03	0.29	0.36
Critique Form test	0.28	0.28	**0.41***	**0.48****

Note. * $p < .05$, ** $p < .01$

4. Discussion

In this paper, three instruments were employed to measure students' critical thinking skills and dispositions. And students engaged in an extended online collaborative inquiry where they were faced with authentic problems and tackling the learning tasks assigned in the humanities module. Content analysis was conducted to examine students' critical thinking behavior exhibited in the online discourse. The results suggest that most students can apply the four critical thinking processes, with the highest percentage of units coded as clarification and the lowest percentage coded as assessment. The correlation results suggest that students' context-specific critical thinking skills, as measured by the Critique Form, are valuable as a predictor of students' critical thinking performance in authentic contexts. Also, students' critical thinking skills and dispositions can be improved as reflected by their scores in context-free tests through engaging in an extended collaborative inquiry. This paper has implications for curriculum developers and practitioners who seek to design authentic learning tasks to promote collaborative inquiry and foster critical thinking in secondary students. And it contributes to the literature on the predictive value of critical thinking instruments on students' ability to exercise critical thinking in authentic problem-solving contexts.

References

[1] D. F. Halpern, Thought and knowledge: An introduction to critical thinking 4th ed. Mahwah, NJ: Erlbaum, 2003.
[2] E. M. Glaser, An experiment in the development of critical thinking. New York: Teachers College, Columbia University, 1941.
[3] R. H. Ennis, "A concept of critical thinking," Harvard Educational Review, vol. 32, pp. 81-111, 1962.
[4] H. Siegel, Educating reason: Rationality, critical thinking and education. New York: Routledge, 1988.
[5] R. W. Paul, Critical Thinking: What every person needs to know in a rapidly changing world. Rohnert park, CA: Center for Critical Thinking and Moral Critique, 1990.
[6] P. A. Facione, et al., The California critical thinking disposition inventory test manual. Millbrae, CA: California Academic Press, 1996.
[7] J. E. McPeck, Teaching critical thinking: dialogue and dialectic. New York: Routledge, 1990.
[8] C. Bereiter, "Critical thinking, creativity, and other virtues," in Education and mind in the knowledge age, C. Bereiter, Ed., ed Hillsdale, NJ: Lawrence Erlbaum, 2002, pp. 341-381.
[9] J. Dewey, How we think. Boston: D. C. Heath and Co., 1910.
[10] J. A. Blair, "Recent developments in critical thinking in anglophone North America," Thinking: The Journal of Philosophy for Children, vol. 7, pp. 2-6, 1987.
[11] R. H. Ennis, "A taxonomy of critical thinking dispositions and abilities," in Teaching Thinking Skills: Theory & Practice, J. Barton and R. J. Sternberg, Eds., ed New York: W. H. Freeman and Company, 1987, pp. 9-26.
[12] J. E. McPeck, Critical Thinking and Education. New York: St. Martin's Press, 1981.
[13] T. Haldyna, Developing and validating multiple choice test items. Millbrae, CA: The California Academic Press, 1994.
[14] S. P. Norris and R. H. Ennis, Evaluating critical thinking. Pacific Grove, CA: Midwest Publications, 1989.
[15] S. P. Norris, "Testing for the disposition to think critically," Informal Logic, vol. 68, pp. 157-164, 1992.
[16] L. Mao, et al., "The revision of Cornell critical thinking test - Level X (CCT-X)," Journal of Elementary Education, vol. 1, pp. 1-28, 1992.
[17] Y. C. Yeh, "A study of substitute teachers' professional knowledge, personal teaching: Efficacy, and teaching behavior in critical-thinking instruction," Journal of Nantional Chengchi University, vol. 78, pp. 55-84, 1999.
[18] R. H. Ennis and E. Weir, The Ennis-Weir critical thinking essay test. Pacific Grove, CA: Midwest Publications, 1985.
[19] C. Perkins and E. Murphy, "Identifying and measuring individual engagement in critical thinking in online discussions: An exploratory case study," Educational Technology & Society, vol. 9, pp. 298-307, 2006.

Positive and Negative Learning Impacts from Technological Social Agents

Jonathan S. HERBERG[a,b*], Daniel T. LEVIN[c] & Martin SAERBECK[a]
[a]*Institute of High Performance Computing, A*STAR, Singapore*
[b]*Department of Psychology, National University of Singapore, Singapore*
[c]*Department of Psychology and Human Development, Vanderbilt University, United States*
*herbergjs@ihpc.a-star.edu.sg

Abstract: Computerized learning environments often include implementations of simulated social agents, incorporating the reasonable design assumption that learning interactions that are more social are more effective. However, recent evidence suggests that computerized social agents can in some circumstances fail to promote or even hinder learning. The current paper outlines evidence both supporting and arguing against the utility of computerized social agents for learning. We propose a framework reconciling this evidence by delineating impacts from a social agent that impinge upon different points of learning, from shallow to deep phases. Shallow social impacts when ineffectual carry little to no potential to actively impede learning, but any potential positive impacts on shallow learning phases are relatively limited in the absence of positive deep impacts. Deep social impacts, on the other hand, carry the potential to strongly drive deep learning, but when ineffectual carry the risk to impede it. The paper concludes with a proposal for future research based on this framework.

Keywords: Social agents, pedagogical agents, social cognition, theory of mind

Introduction

The notion of the beneficial impacts of a social interaction context has been incorporated into various socio-cognitive frameworks for learning. On this basis, computerized learning environments often include simulated social agents, which can be used either to fill the role of the tutor [1] or, to generate learning-by-teaching, to fill the role the role of the agent that the student is supposed to teach [2].

However, a number of tests suggest limitations to the effectiveness of social agents for learning, in particular of those that fail to provide adaptive explanatory feedback [3]. This suggests there may be added layers of extraneous processing associated with a simulated social agent, which in the absence of cues to channel the behaviors for effective learning can sum to a net cost rather than benefit [4]. Related to this, situations in which people must use their representation of a social agent partner's differing knowledge states typically require extra cognitive effort [e.g., 5, 6]. Thus, in the context of computer based learning environments, it is important to consider whether the cognitions and behaviors that can be elicited by social agents sometimes result in net learning costs when they function as diversions from, rather than reinforcements of the learning context.

A recent study has suggested that a social interaction context may in some situations generate a greater cost than benefit to learning even when the social agent seemingly elicits more active learning behaviors [7]. In this study, participants' transfer of their knowledge of how to solve the 3-ring Tower of Hanoi problem in order to solve the 4-ring problem was less robust after they demonstrated their 3-ring solution to a social agent, a person, as compared to those who demonstrated to a non-anthropomorphic computer. Those demonstrating to the social relative to non-social agent engaged in more behaviors such as highlighting solution steps for the agent while demonstrating. Therefore certain social interaction behaviors may in some contexts actually impede learning. In the current paper, we develop a framework that classifies different social agent impacts as beneficial versus costly to sustained, deep self-regulated learning. This framework provides general parameters on when social interactions should help and when they should hurt deep learning resulting from knowledge construction, suggesting fruitful avenues of future research.

1. Learning Effects from Social Situations and Computerized Social Agents

1.1 Learning Behaviors and Depth of Learning

We begin by examining behaviors that may stem from an interaction with a social agent that may enhance learning. A framework by Chi separates out active, constructive, and interactive learning behaviors [8]. Active learning behaviors are comprised of physical movements and engagement in learning. Constructive behaviors are seen in the learner's verbal reasoning, logical elaborations and links, and explaining of solutions. Cognitively, the learner now recruits knowledge to infer and construct new knowledge. Finally, interactive behaviors can be observed when two learning partners listen to each other, respond to each other, and argue or collaborate with each other in their problem solutions.

From this sketch of learning behaviors at different depths it is possible to classify specific mechanisms that can drive beneficial learning effects from a social agent. Social agents may generate learning benefits by, on the one hand, eliciting from the learner more attention and engagement in the task and by increasing the activation of already-existing task-relevant knowledge. On top of such active learning behaviors, social agents may shape a path of constructive learning by pushing the learner to engage in deep reasoning, question seeking and answering, and inference of new knowledge from prior knowledge. Social agents may induce this form of deep learning by prompting the learner to explain, reason with, and build conceptual links between different pieces existing knowledge, and to make inferences of new knowledge on the basis of existing knowledge. Social agents may, furthermore, push the learner to reflect on what he or she knows and what needs to be learned. Such metacognitive behaviors, crucial for deep abstract learning, occur most vigorously when the learning interaction is rich and sustained [8, 9]. At the deepest level of interaction between two learning agents, the two will build on each other's knowledge and make joint inferences of and elaborations on new knowledge [10]. Some examples of the most effective computer-based learning environments in generating robust learning outcomes use simulated social agents for dialogic interactions, whether in the form of a kind of Socratic tutor who scaffolds learning [1], or in the form of a peer-agent the student builds concept maps for which the agent in turn indicates understanding on the basis of [2].

1.2 Self-Regulated Learning and Depths of Learning

Ultimately, the goal of computer or robotic implementations with simulated social agents should be toward deep knowledge-building through self-regulated learning (SRL). SRL embodies a set of purposeful, self-directed learning behaviors and cognitions, leading to deeper understanding and better repair of misconceptions than externally-regulated learning. A successful self-regulated learner continuously monitors what he or she knows and how much he or she is learning on a given learning task, and on that basis sets and updates goals and strategies to learn what needs to be learned [1, 11]. Thus, a learner engaging in effective SRL is constantly devoting significant cognitive resources and effort toward metacognition, that is, monitoring and thinking about his or her own knowledge and learning. Successful learning environments that employ social agents should therefore motivate and guide SRL.

It is possible to identify four phases of SRL as follows [11]. 1) Task definition: The initial phase where the learner, for effective direction of resources, sets in mind the conditions of the overall learning task including the context and materials. 2) Goal-setting: This phase involves setting specific learning goals, including setting standards of adequate versus insufficient understanding, and constructing plans for reaching these goals. 3) Studying-tactics: The learner enacts plans to guide his or her learning. 4) Metacognitive adaptations: The effective self-regulated learner self-evaluates his or her learning by comparing products of learning to the initial standards and on this basis alters the task definition, standards of learning, learning goals, and plans for achieving goals. This last phase is where metacognitive processes are most relevant, which highlights their role in directing the setting of goals for correcting misconceptions.

1.3 Separate Kinds of Social Learning Impacts at Separate Depths

Based on this learning-depth distinction, we may further distinguish between the effects a social agent may have on more shallow learning behaviors, tied to basic engagement in a task, versus deeper learning impacts from a social agent, tied to metacognition and misconception correction. That is, tweaking for the current context the general idea from Social Impact Theory of various distinct measurable social impacts that can impinge on task performance [12], we propose that separable effects arise from shallow versus deep social impacts. We propose that social agents have the potential to generate a basic facilitation of learning from automatic or simple mechanisms such as boosted neural arousal [cf. 13], selective attention to task dimensions [14], and potentially more emotion to increase motivation [15]. When such impacts on earlier learning phases fail to generate robust outcomes, however, they do not actively *interfere* with learning. On the other hand, we propose that deeper social impacts have the potential to either *enhance* metacognitive and related deep learning processes, or *divert* cognitive resources away from such processes.

2. Potential Negative Impacts on Deep Learning from a Social Agent

With our developed framework of separable social impacts we are now in a position to better understand the results of the recent study broached in the introduction, which suggested a net negative learning impact from a social versus non-social agent [7]. Recall that in this study participants first learned how to solve the 3-ring Tower of Hanoi problem, then demonstrated the solution to either a social or non-social agent, and then as a measure of their learning solved for themselves the more difficult 4-ring problem. Participants' less optimal solutions to the 4-ring problem after demonstrating to the social agent presents a mystery, in that the social and the non-social agent were both represented by a simple picture, and subjects' demonstrations for the social agent were more active (i.e., there were more looks to the picture and more social highlighting, i.e., pointing behaviors, for the social agent). In particular, this presents a challenge for the Chi framework, which conceptualizes active behaviors as always being potentially better, never worse for learning, than passive engagement [8]. However, in the context of the framework we have developed for shallow and deep social impacts, the limitations of shallow social effects on learning become easier to see, as does the potential traps of deep social impacts that are not guided by the interaction to enhance learning, thereby carrying the capacity to function as an *opposing* rather than merely ineffectual force. This is made clear by a multiple regression analysis that suggested that pointing in the 3-ring demonstration adversely affected performance in the 4-ring task, while looks at the agent were actually linked to better performance. That is, the increased looking behaviors represent effects confined to shallow levels of learning, that while improving 4-ring performance, are swamped by the negative impacts that can directly interfere with deep learning cognitions. Pointing behaviors on the other hand seem tied to deep but negative social impacts. That is, with increased metacognition for the imagined social agent's insufficient task knowledge and the resulting increase in action highlighting behaviors, resources are diverted away from building a richer representation of the solution that is more readily transferable to solving the subsequent more difficult problem. The main contrast making the metacognition here detrimental to learning-by-teaching versus the strong learning benefits found in computerized teachable-agent implementations is that in the latter case the social interaction guides the student's metacognitive behaviors to cohere with correcting and building the student's own knowledge, rather than interfering with this end. Thus such implementations channel the student's metacognition for the teachable agent toward the student's own knowledge construction [2, 16]. However, with no guidance from the appropriate social interaction, ironically the deep social impact arising in the context of demonstrating to an imagined social agent results in *shallow* learning behaviors [7].

3. Conclusions and Future Research Directions

These considerations suggest paths of future research for testing the relative strength of different social effects, positive and negative, on both shallow and deep learning, and for thereby generating insights for maximizing the learning impacts of simulated social agents. In our sketch of learning program implementations we traced out a number of different kinds of social impacts that a social agent can carry. As we noted there are simple effects of increased arousal and engagement that come more or less automatically from a social interaction. Future work can be carried out to investigate whether distinct kinds of effects at this level can be differentiated, and whether the *social* character of such cues are necessary for such effects. For instance, in tasks which benefit from cues that elicit attention at specific timepoints in learning, is there a difference in the cue's effectiveness if it is a *social* cue as opposed to a non-social signal (e.g., a disembodied beep)? Furthermore, is the generalized learning impact of a watchful social agent dissociable from specific cues for directing attention? Turning to deeper social impacts, our overview highlighted two distinct metacognitive effects positive for learning. Social agents may dialogically guide a learner to notice misconceptions, by pushing the learner via prompts and hints which provide just enough information for the learner to figure out misconceptions on his or her own [1]. As a distinct pro-metacognitive and learning-enhancing impact, social agents can lift a natural impediment for acknowledging and correcting misconceptions, when the learner engages in these essential deep learning behaviors for the agent rather than his or herself [4, 16]. On the flip side, however, limited social interactions may push metacognitive behaviors that are *detrimental* to learning, resulting ironically in shallow-learning behaviors, with resources for constructing a deep understanding of the problem solution being diverted instead toward actions like highlighting solution steps for the social agent [7].

A related line of follow-up research suggested by our framework might, given an empirically determined set of both positive and negative social impacts on learning, test the relative *strength* of each impact by keeping the social learning situation a constant except for a manipulation of one impact across conditions at a time. For instance, experiments may be designed to test the strength of social agent behaviors which push learning-relevant metacognition in their feedback, versus the strength of the general metacognitive impact that comes from the creation of knowledge representations for a teachable agent. The strength of these effects can in turn be compared to a social interaction that entails a *negative* learning impact from the social agent. It may also be possible to determine if each of these sorts of impacts, whether positive or negative, affect a learner's performance in a pattern similar to that well established for a wide range of social impact factors, such as the impact of a group's opinion's on conformity [12]. For instance, if after learning how to solve a 3-ring Tower of Hanoi problem the task were to demonstrate to, across conditions, one, two, or three social agents, would the number of participants' looks and points fit with the pattern suggested by the psychosocial law, of a diminishing increase in these behaviors for each subsequent agent? Would the net (negative) impact on subsequent performance for the 4-ring task in terms of time and number of solution steps also fit this pattern? Would other principles of social impact theory be applicable in studying the effects of social agents on learning and learning-behaviors? These considerations highlight the potential of developing from the current framework methods of pinpointing the various components of positive and negative learning effects that may arise from a social agent and a social interaction context.

Acknowledgement

This work was supported in part by an NSF grant (#0826701) to the second author.

References

[1] Graesser, A. C. & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist, 45*(4), 234-244.

[2] Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education, 18*(3), 181-208.

[3] Moreno, R. (2004). Animated pedagogical agents in educational technology. *Educational Technology, 44*(6), 23-30.

[4] Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychological Review, 19*, 309-326.

[5] Keysar, B., Shuhong, L., & Barr, D.J. (2003). Limits of theory of mind use in adults. *Cognition, 89:1*, 25-41.

[6] Epley, N., Boven, L. V., Keysar, B., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. Journal of Personality and Social Psychology, 87(3), 327–339.

[7] Herberg, J. S., Levin, D. T., & Saylor, M. M. (2011). Social audiences can disrupt learning by teaching. *Journal of Experimental Social Psychology, 48*(1), 213-219.

[8] Chi, M. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*, 73-105.

[9] De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Science, 14*(10), 441-447.

[10] Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science, 32*(2), 301-341.

[11] Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist, 45*(4), 210-223.

[12] Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343-356.

[13] Zajonc, R. B. (1965). Social facilitation. *Science, 149*, 269-274.

[14] Huguet, P., Galvaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology, 77*(5), 1011-1025.

[15] Locke, E. A. & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist, 57*(9), 705-717.

[16] Chase, C. C., Chin, D. B., Opezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology, 18*(4), 334-352.

A Study in Negotiation-based Peer Assessment: Natural Language Applied in Assessment Representation

Chung-Hsien Lan[a*], K. Robert Lai [b], Chih-Yueh Chou[c] & Kuo-Hung Chao [d]
[a]*Assistant Professor, Taoyuan Innovation Institute of Technology, Taiwan*
[b]*Professor, Yuan Ze University, Taiwan*
[c]*Associate Professor, Yuan Ze University, Taiwan*
[d]*PhD. Student, National Taiwan Normal University, Taiwan*
*chlan@tiit.edu.tw

Abstract: This study presents a conceptual framework for providing intelligent supports through agent negotiation, fuzzy constraints and natural language processing to enhance the effectiveness of peer assessment. By using fuzzy constraints, it not only provides a flexible marking scheme to deal with the imprecision and uncertainty for the representation of assessment. Additionally, a fuzzy constraint-based negotiation mechanism is employed to coordinate the cognitive differences between students. Through iterative agent negotiation, students can reconcile the differences and reach an agreement on the assessment results. Owing to the difficulty in reading fuzzy sets of assessment results we incorporate the technique of linguistic approximation to translate fuzzy sets into natural language to facilitate students to understand assessment feedback. Experimental results indicated that students were able to acquire more meaningful and readable feedback to reflect upon and revise their work.

Keywords: Peer assessment, linguistic approximation, natural language, agent negotiation

Introduction

Peer assessment supports group learning by motivating students in deep thinking, comparison, discussion and critical judgment of peer work. Numerous researchers have investigated the effectiveness of computer-based peer assessment systems in various learning scenarios [1][2]. In our previous study, we constructed a negotiation-based peer assessment system (NePAS) for providing intelligent supports through agent negotiation and fuzzy constraints to enhance the effectiveness of peer assessment. In this framework, assessments are represented as fuzzy membership functions to deal with the inexactness of marking and its subjective nature. Additionally, a fuzzy constraint-based negotiation mechanism is employed to coordinate the cognitive differences between students. Through iterative agent negotiation, students can reconcile the differences and reach an agreement on the assessment results. The proposed framework can provide more detailed, informed, and make students more inclined to accept the results and to reflect upon and revise their work. However, human are often led to use words in natural language instead of numerical values. Interpretation of fuzzy sets often involves the use of linguistic approximation that assigns a linguistic term to a fuzzy set based on the predefined primary terms, linguistic modifiers and linguistic connectives [3]. This study incorporates the techniques of linguistic approximation and natural language processing to enhance assessment representation. Fuzzy linguistic techniques that can help allow representing qualitative phenomena from a quantitative approach and even deal with incomplete information [4]. According to the mapping of linguistic approximation and fuzzy sets of assessment results, the graphics of assessment results can be transformed into sentences to express peers' suggestions.

1. Enrichment of Peer Assessment Process

Previous studies reveal that students can have a more in-depth contact with the course material for knowledge interpretation, prolonged interaction between peers for provision of constructive feedback based on multiple observations of performance and opportunity to develop critical reasoning skills and self-directed learning during peer assessment [5][6]. Through a student-involved and interactive process, students' interpretation and reflection can be enhanced, and instructors also can improve their understanding of students' performance by observing students' interaction. However, students may not have the control over the assessment process, and thus they possibly disagree with the assessment rating given by instructors or other peers. Students have difficulties in comprehending how to reflect on their work if assessment results are only given as scores without textual feedback [7]. To alleviate the aforementioned weakness, we have presented a conceptual framework for the enrichment of a peer assessment process as shown in Figure 1.

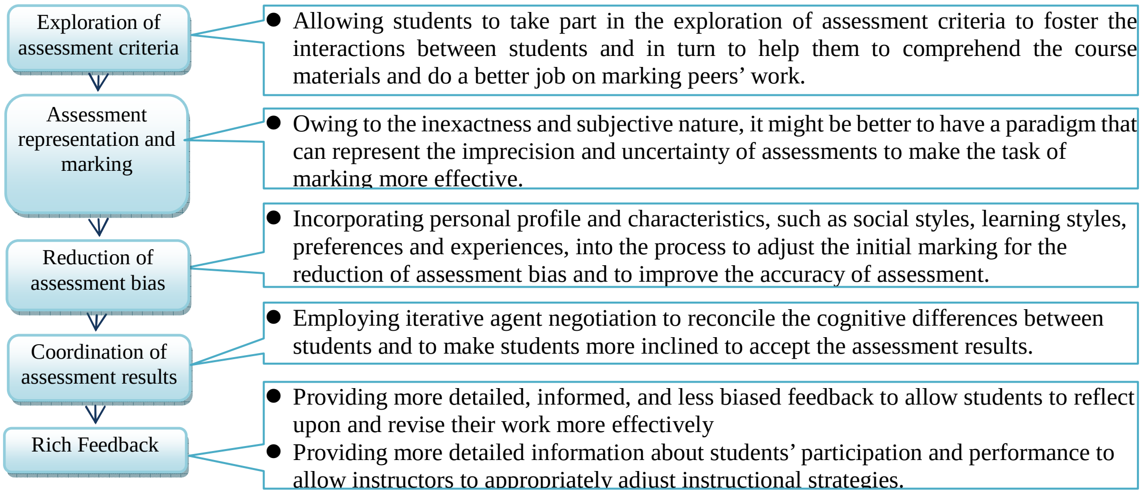


Figure 1. A conceptual framework for the enrichment of a peer assessment process

In this study, we focus on the aspect of rich feedback. Receiving accurate and complete feedback is correlated with effective learning [8]. In the stage of rich feedback, previous proposed method uses defuzzification technique to generate scores to represent assessment results. However, the type of assessment results cannot provide students rich feedback. This study proposes an enhanced framework to incorporate linguistic approximation and natural language processing to produce meaningful representation about assessment results. The enriched peer assessment process enables students to enhance course interpretation, frequently interact with peers, represent their thinking and reflect their work. Through the interactive process and enhancement of assessment representation, assessment accuracy and quality can be improved. The overall process facilitates students in fostering critical thinking skills and reflection as well as promoting meaningful learning.

2. Representation of Assessment in Natural Language

In our previous study, the assessment process and results were represented in graphics and scores. It is difficult to facilitate students to understand peers' assessment and intent of assessment results. Therefore, this study incorporates fuzzy linguistic techniques and natural language processing to represent assessment results in words in order to produce meaningful feedback. Fuzzy linguistic techniques that can help allow representing qualitative phenomena from a quantitative approach and even deal with incomplete information [9]. The linguistic approach is an approximate technique which represents qualitative aspects as linguistic values by means of linguistic variables, that is, variables whose values are not numbers but words or sentences in natural language. The label is a word or sentence belonging to a linguistic term set and the meaning is a fuzzy subset in a universe of discourse. The information processing includes the following three steps:

- Definition of the linguistic term set with its semantic. It consists of establishing the linguistic expression domain used to provide the linguistic performance values according to the different assessment criteria.
- Computation of linguistic approximation. Linguistic approximation can be formalized in the terms of re-translation rules that correspond to the translation rules [11]. The fuzzy membership function of assessment results can be translated into linguistic quantifiers that can be used to represent assessment in natural language.
- Decision of the best representation. It consists of choosing the best semantic according to the linguistic approximation provided.

According to the definition of Zadeh [10], a linguistic variable is characterized by a quintuple (L, $H(L)$, U, G, M) in which L is the variable (which is the assessment criteria); H(L) (or simply H) denotes the term set of L, i.e., the set of linguistic values of L, with each value being a fuzzy variable denoted generically by X and ranging across a universe of discourse U which is associated with the base variable u; G is a syntactic rule (which usually takes the form of a grammar) for generating the names of values of L; and M is a semantic rule for associating its meaning with each L,$M(X)$, which is a fuzzy subset of U.

In the first step, in order to reduce the complexity of defining a grammar, we use an approach based on an ordered structure to define linguistic terms. A set of seven terms H could be given as H = {H_0 = *None*, H_1 = *Very poor*, H_2 = *Poor*, H_3 = *Ordinary*, H_4 = *good*, H_5 = *very good*, H_6 = *perfect*}. The semantic of the linguistic term set is defined by an ordered structure and fuzzy sets represented by triangular and trapezoid membership functions. These membership functions are uniformly distributed. The semantic representation is achieved by four parameters. The first two parameters indicate the interval in which the membership value is 1; the third and fourth parameters indicate the left and right width. For example, H_0 = *None* = (0, 30, 0, 10),…, H_3 = *Ordinary* = (60, 60, 10, 10),…, H_6 = *Perfect* = (90, 100, 10, 0).

In the second step, the problem of linguistic approximation can be defined as mapping from a fuzzy set X of assessment results for one assessment criterion into a set of terms H. The approximation of fuzzy set H_i and X can be defined as follows:

$$\sum Count(H_i/X) = \frac{\sum Count\ (H_i \cap X)}{\sum Count(X)}$$

where $\sum Count(X) = \sum_{j=1}^{N} \mu_{X(j)}$ and i represents the number of linguistic terms. A solution of linguistic approximation is a linguistic description HA composed of linguistic primary terms A and linguistic connectives c such that it is most meaningful to describe a possibility distribution of a linguistic variable. For example, a given possibility distribution of a fuzzy set X describing the assessment criterion may be linguistically approximated to "Content is good or very good", i.e. $HA(X) \equiv X$ is (A_1 c A_2) where $X \equiv$*Content*, $A_1 \equiv$*Good*, $A_2 \equiv$*Very good* and $c \equiv$or. Finally, the best semantic can be represented. The results can be represented as "Content is good or very good" for providing appropriate feedback in natural language.

3. System Realization and Illustrative Example

A walk-through example then is used to illustrate a peer assessment process. First, each student accesses the NePAS through an assessment agent which provides intelligent supports for various assessment activities, including criteria exploration and ranking, characteristics detection, self-assessment, making peers' work and feedback. Coordination agent adopts a fuzzy constraint-based negotiation mechanism to resolve the cognitive differences among the assessors and learner himself. Assessment database includes students' assessment log and coordination results. The coordination agent incorporates the techniques of linguistic approximation to translate the fuzzy sets of negotiation results into natural language. The architecture of coordination agent is shown as Figure 2.

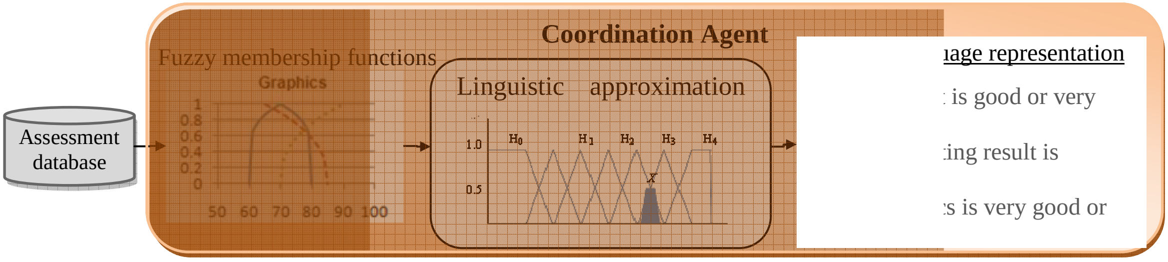


Figure 2. The architecture of natural language representation

During the peer assessment process, students complete and submit their projects to the system, and they then move on to perform peer assessment activities. Assessors select fuzzy membership functions (i.e., triangular, trapezoidal and Gaussian) for each criterion first and then fill out the required (e.g., supports) and optional parameters (e.g., satisfaction degrees). Afterward, a graphical representation of the fuzzy membership function is displayed on the right for reviewing and can be changed literally, if necessary. By using fuzzy membership functions for assessment representation, it provides not only an effective approach for dealing with the uncertainty and impreciseness, but also allows the students to express the confidence of their assessment. A negotiation is automatically performed to coordinate the cognitive differences among students. Agreement is achieved when all participants agree. The communication protocol for agent negotiation is adapted from [11]. The curves indicate the acceptable ranges when students propose their own offers by lowering the threshold. If an overlap exists between acceptable ranges, an agreement can be expected. Otherwise, agents need to revise their assessments prior to a new negotiation process.

Finally, after several rounds of negotiation between agents I, J and K, it has arrived at an agreement on the assessment results of student K's web site design as shown in Figure 3 (solid areas). However, students are difficult to understand the intents of the graphics, and thus we use the technique of linguistic approximation to translate them into natural language to facilitate students to realize assessment feedback. For the assessment criterion "*Content*", the assessment representation in natural language is "*the content is good or very good*". The feedback and the linguistic approximation are represented in Figure 3. At the same time, students can also examine the satisfaction value (0.6, 0.6, 0.7, 0.6) for each criterion. The closer the satisfaction value is to 1, the higher is the acceptance for the assessment results. Therefore, the system can offer rich feedback with two dimensional representation and foster deeper reflection and thinking. Additionally, the system also employs a defuzzification technique to render numerical scores for students' performance.

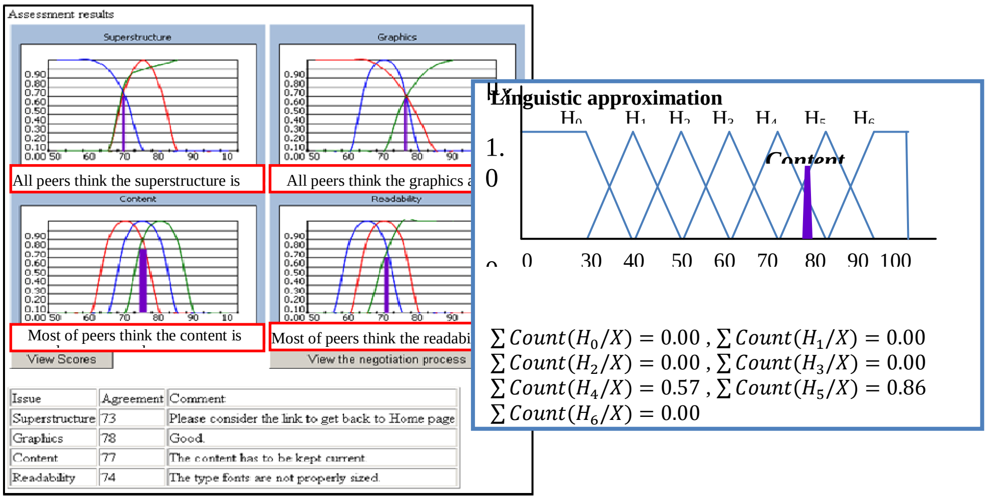


Figure 3. Representation of assessment results and linguistic approximation

4. Conclusions

This study has presented how to use linguistic approximation to represent assessment results in natural language. The difficulty in reading assessment results can be resolved and thus students can understand peers' feedback more inclined to accept the assessment results and to reflect upon their own work. Although the proposed methodology has yielded promising results in promoting the effectiveness of peer assessment, considerable work remains to be done, including further large-scale classroom experiments and system improvement.

References

[1] Lin, S. S., Liu, E. Z. F., Chiu, C. H., & Yuan, S. M. (2001). Web-based peer assessment: attitude and achievement. *IEEE transaction on Education, 4* (2), pp.211.

[2] Davis, R., & Berrow, T. (1998). An evaluation of the use of computer supported peer review for developing higher-level skills. *Computers and Education, 30*, pp.111-115.

[3] Kowalczyk, R. (1999). On numerical and linguistic quantification in linguistic approximation, *Proceeding of IEEE Systems*, pp. 326-331.

[4] Cabrerizo, F. J., P´erez, I. J., & Herrera-Viedma, E. (2010). Managing the consensus in group decision making in an unbalanced fuzzy linguistic context with incomplete information, *Knowledge-Based System, 23*(2), pp. 169–181.

[5] Hattum-Janssen, N. V., & Lourenco, J. M. (2006). Explicitness of criteria in peer assessment processes for first-year engineering students. *European Journal of Engineering Education , 31* (6), pp. 683–691.

[6] Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education , 27*, pp. 427-441.

[7] Lan, C. H., Graf, S., Lai, K. R., & Kinshuk. (2011). Enrichment of Peer Assessment with Agent Negotiation, *IEEE Transactions on Learning Technologies*, 4, pp.35-46.

[8] Bangert-Drowns, R. L., Kulick, C. L., Kulick, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research , 61*, pp. 213-238.

[9] Herrera, F., & Herrera-Viedma, E. (2000). Linguistic decision analysis: Steps for solving decision problems under linguistic information, *Fuzzy Sets System, 115*(1), pp. 67–82.

[10] Zadeh, L. A. (1975), The concept of a linguistic variable and its applications to approximate reasoning. *Part I, Inform. Sci. 8*, pp. 199-249.

[11] Lai, K. R., & Lin, M. W. (2004). Modeling agent negotiation via fuzzy constraints in e-business. *Computational Intelligence , 20*, pp. 624-642.

Combining Facebook and Open Learner Models to Encourage Collaborative Learning

Mohammad Alotaibi*, Susan BULL
University of Birmingham, UK
*maa715@bham.ac.uk

Abstract: In this paper we describe the use of a social network application (Facebook) with individual open learner models at university level, to support collaborative learning. Results suggest that Facebook and open learner models can be used together to support collaborative learning. It also suggests that Facebook and open learner models can support sociability in collaborative learning: both were found to be useful.

Keywords: Open Learner Models, Online Social Networks, Collaborative Learning

Introduction

Support for collaboration between learners is one of the aims of Computer-Supported Collaborative Learning (CSCL) [1] and Open Learner Models (OLM) [2]. In adaptive educational systems, an OLM is a learner model that is accessible by the learner. It is the representation of a learner's knowledge and skills, and sometimes their misconceptions or other attributes. Promoting learner metacognition through reflection; giving learners more control over their learning; helping learners to plan and monitor their learning and promoting collaboration and interaction between learners are some of the reasons behind opening the learner model [2]. Learners can make use of their OLM to identify their knowledge and misconceptions, etc. Then they can choose a suitable way to proceed in their learning. This addresses the first three reasons (above) for using an OLM. On the fourth (collaboration), three approaches to using OLM to support collaborative interaction and learning have been described: individual learner models available for peers to view; a group model comprising data from individual team members; and a combined group model which is available to group members [3]. We here focus on the first approach where students are presented with individual learner models (using OLMlets [4]). We aim to support both user collaboration around their OLM and social interaction by using online social networks.

In addition to collaboration, work has emphasized the importance of support for social interaction and communication between learners in CSCL environments [5],[6] [7]. Online social networks were developed to facilitate social interaction between users. However, the current generation of Web 2.0 social networking applications also offer tools that can support collaborative learning activities. Facebook, for example, provides several communication channels that can be used by learners to communicate with peers or instructors. It can support both synchronous and asynchronous discussion. It can also be used to create study groups by using the group tools to create online learning communities. Research in CSCL has raised the possibility of using online social networking applications to support CSCL aims like learner interaction and collaboration [8].

In this paper we describe the use of an OLM that can support collaborative learning and prompt learner discussion [4],[9] with the Facebook online social network application.

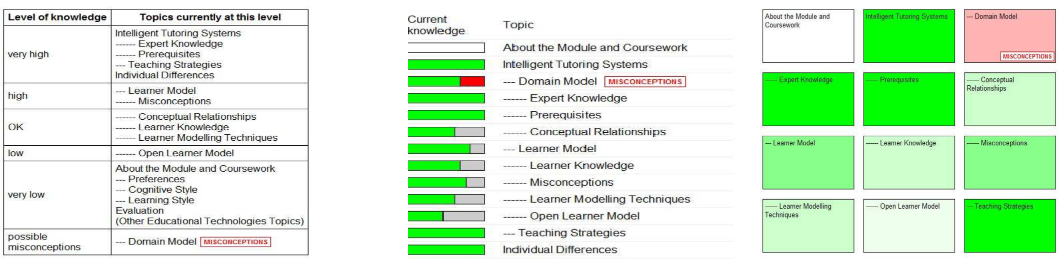
1. Supporting Collaboration and Social Interaction with OLM and Facebook

The fact that Facebook is used by many students in higher education motivated several researchers to explore the educational use of Facebook [10],[11]. Recent work on educational use of social networks suggests that in addition to support for social interaction, they can also be used to enhance critical thinking skills by providing tools that facilitate communication, interaction between learners and collaborative learning [12],[13].

OLM can support collaboration and discussion [4],[9]. Giving learners the option to release (open) their learner model to their peers and instructors has resulted in spontaneous collaboration, discussion and peer help [4]. Furthermore, students reported that they sensed a feeling of community and togetherness when they were using OLM at the same time, as one of the students describes: *When several people were using OLMlets at the same time, most notably in one of the small computer rooms, there came to be almost a community feel. Students were comparing their model against those of people in the room, and discussions were occurring spontaneously all the time* [4]. Here we aim to make use of this collaboration prompted by students' individual OLMs, and extend the social interaction from this collaboration using an existing social networking application that can provide tools for collaboration and social interaction between learners, to support collaborative learning.

2. The Use of OLM and Facebook to Support Collaboration and Social Interaction

Previous research considered Facebook and face-to-face discussion of an OLM, finding that students use both approaches [14]. We here consider the Facebook interactions in greater detail. OLMlets [4] is a domain-independent web-based OLM. It uses students’ answers to multiple choice questions to construct a simple learner model. It then externalises learner models (consisting of knowledge level and possible misconceptions) in different formats, for example: tables with ranked list of topics, skill meters, boxes indicating knowledge level by color (Figure 1). OLMlets gives learners the option to open their learner models to other users such as peers and instructors. They can release their learner model either with their names visible or anonymously, in which case they cannot be identified. By releasing the learner model, learners can compare their learner models with those of their peers.



a **b** **c**

Figure 1: a: table view in OLMlets; b: skill meter view in OLMlets; c: boxes view in OLMlets

2.1. Participants, Materials and Methods

Participants were 15 third year students in the School of Electronic, Electrical and Computer Engineering at The University of Birmingham, taking an Adaptive Learning Environments module. They used Facebook alongside OLMlets for 5 weeks as part of their course. OLMlets was introduced to students in a two hour lab session, and then they used it as they wished in their own time. Students joined a secret Facebook group page for the class, to allow them to optionally discuss their learning with their peers. Students' participation in Facebook and their use of OLMlets was not assessed in this course. At the end of the fifth week, students returned a questionnaire about using Facebook with OLMlets, with response options on a five point scale (strongly agree – strongly disagree).

2.2. Results

Students' interaction moved from no interaction in the first week to higher level of interaction in weeks 2-4, and then decreased in the final week (see [14]). The nature of students' posts on Facebook varied. Students posted questions about different things related to their learning, such as questions asking for clarification about some aspects of adaptive learning environments. For example, a student asked on Facebook: *What is the domain any one?* They also posted questions related to their assessed course assignments, for example*: I'm a bit confused... My subject doesn't really have rules... just facts.. so how do I portray this?* Most of the posts on the Facebook wall were about the assessed coursework. In both cases students attempted to answer peers' questions on Facebook and ask for help from peers or the instructor when they were not sure of their answers. The examples below relate to one of the more common misconceptions identified in students' use of OLMlets, posted by the instructor to encourage discussion: "an intelligent tutoring system does not understand the domain model".

<u>*Example 1. Student question and peer responses*</u>
S1: *Can a domain model contain images or animation as well as text explanation?*
S2: *[The instructor] said if you wanted you could by you have to justify why.*
S3*: I think what she wants in for the domain model to just contain expert knowledge, then the images will be held in the system. The teaching strategy would then choose when and how to use these images.*
S4*: Yeah this is what S5 and I were discussing and I'm not 100% convinced either way. I maintained that you needed to keep in mind when "modelling" something like a domain … you'd keep just the expert knowledge there and maybe not worry about the format (i.e. images, text files, etc). When I type my ideas out and read them back they sound even more confusing than when they're in my head. S5? [Instructor?]*

<u>*Example 2. Peer and instructor responses to a question about domain pre-requisites*</u>
S1: *i'm under the impression it's domain model*
S1: *i've changed my mind, i'd say the teaching strategies. the teaching strategies have the pre-requisites for the learner, the domain model has pre-requisites for how the domain links itself together*
S2: *I think it is under teaching strategies.*
Instructor: *In OLMlets it comes under domain model.*
S3: *I said it had to be stored in the domain model. For all the other (more advanced) knowledge in the domain to be true, it was all linked and built upon the prerequisite knowledge – so for it to be true, it had to include the prerequisite stuff too.*

Some students also posted messages to show their frustration, for example: *a part of me just died :(*. Other students posted messages showing sense of humor, for example, when the instructor used the Facebook Like button a student comments: *What I said is correct. I know this because [the instructor] liked it!*

Questionnaire responses (Table 1) show that 3 students released their learner model named to peers and 7 students released their model to peers anonymously. Most reported that they used Facebook to find peers to work with. Students also reported that they used Facebook to seek help from peers. A few students stated that they tried to work out some OLMlets questions in Facebook and about half claimed the Facebook interaction to be helpful for answering OLMlets questions. Finally, most found Facebook and OLMlets helpful for their module, with 14 out of 15 students for Facebook and 13 for OLMlets.

Table 1. Questionnaire responses

Questionnaire item		**Questionnaire item**	
Released OLM named to peers	3	Released OLM anonymously to peers	7
Used FB to find collaborators	12	Used FB to find help from peers	8
Worked out OLMlets questions in FB	4	FB helped answer OLMlets questions	7
OLMlets useful for the module	13	FB useful for the module	14

2.3. Discussion

The different types of student postings show that Facebook can contribute positively to several factors when it is used alongside OLM. The first relates to learning and completion of coursework. Students used Facebook to ask specific questions about the subject (e.g. *what is the domain model?).* This student considered interacting with peers on Facebook as an option to find an answer to his question. It may also suggest that the student wants to benefit from having the instructor in the group to receive further explanation about this specific topic. Students also asked several questions about assessed work. Although working individually on their assignment, they interacted to discuss and ask questions about their work. This shows that students also considered their interaction with peers on Facebook as a way to support their learning during preparation of their coursework. We also find that students tried to construct knowledge together by commenting on each other's posts and answers, and this is one of the primary aims of CSCL. In the first example, when student S1 asked a question, 3 students attempted to answer even though they were not sure about their answer. The comments also show that S4 tried to find help from the instructor. The second example illustrates how students may give extended explanations even after 'the answer' has been given. From the above we suggest that students' interaction in Facebook, with the availability of the instructor, can support collaboration. The other important factor is the social factor. Students did not conceive Facebook as a formal learning environment; they often used both learning and sense of humour when commenting on peers' posts. We also find that Facebook can be used to support learning with OLM as some students tried to make use of Facebook to ask questions about their OLMs or find people to collaborate with.

When using OLM, students are expected to identify their level of knowledge and find out what misconceptions they have, and then they can choose their preferred way to improve their knowledge or overcome misconceptions. Giving them such control over their learning is one of the primary aims of OLM [2]. In the questionnaire responses we find that students tried to seek collaborators using the Facebook group when it was made available to them even though they were not instructed to do so, and their models represent their individual knowledge and misconceptions. We also find that there are students who did not share their model with peers, but who still tried to work with peers. This suggests that OLM can support collaboration and interaction in line with previous findings [4] [9]. We also find that the majority of students (12 out of 15) tried to find peers to work with on Facebook. This suggests that students considered Facebook interaction as a way to seek collaborators even though their coursework was assessed individually, and possibly considered Facebook as a way to extend their search for people to work with. Fewer students reported that they used Facebook to work out OLMlets questions with peers. This may be because they tried to find peers who can give them immediate feedback, while in Facebook they would have to wait if they and their peers are not online at the same time (as suggested for other asynchronous discussion [15]). Indeed, previous work has considered use of Facebook, face-to-face discussion and OLMlets, finding that students use both discussion approaches [14]. We see that students show a positive attitude towards using Facebook with OLMlets to support their collaboration and discussion as 14 reported that they found using Facebook helpful for their learning in this module, and 13 found OLMlets helpful for the module.

3. Summary

This paper has described the use of Facebook to help support student interaction about their understanding and open learner models in a university context. It finds that both social and knowledge-related aspects of social networking were used, and the OLM is available to provide a focus for discussion. This extends previous research of face-to-face discussion prompted by OLMs available to peers [4], to a context where discussion can continue easily when students are not physically or temporally together.

References

[1] Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-Supported Collaborative Learning: An Historical Perspective. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 409-425). Cambridge: Cambridge University Press

[2] Bull, S. & Kay, J. (2007) Student Models that Invite the Learner In: The SMILI Open Learner Modelling Framework, *IJAIED* 17(2), 89-120.

[3] Bull, S. & Vatrapu, R. (2011). Supporting Collaborative Interaction with Open Learner Models: Existing Approaches and Open Questions, *Computer Supported Collaborative Learning* 2011.

[4] Bull, S. & Britland, M. (2007). Group Interaction Prompted by a Simple Assessed Open Learner Model that can be Optionally Released to Peers. In: *PING Workshop*, UM2007.

[5] Kreijns, K., Kirschner, P., Jochems, W., & van Buuren, H. (2007). Measuring Perceived Sociability of Computer-Supported Collaborative Learning Environments. *Computers & Education* 49(2), 176-192.

[6] Vatrapu, R., Suthers, D., & Medina, R. (2008). Usability, Sociability, and Learnability: A CSCL Design Evaluation Framework. *International Conference on Computers in Education* (ICCE 2008), (CD-ROM).

[7] Kreijns, K., Kirschner, P., & Jochems, W. (2002). The Sociability of Computer-Supported Collaborative Learning Environments. *Educational Technology & Society*, 5(1), 8-22.

[8] Greenhow, C. (2009). Social Network Sites and Education: Emerging Research within CSCL. In A. Dimitrakopoulou (Ed.), *Computer-supported Collaborative Learning 2009*, Rhodes, Greece.

[9] Johan, R. & Bull, S. (2010). Promoting Collaboration and Discussion of Misconceptions Using Open Learner Models, in A. Bader-Natal, E. Walker & C.P. Rose (eds), *Proceedings of Workshop on Opportunities for Intelligent and Adaptive Behaviour in Collaborative Learning Systems*, ITS'10, 9-12.

[10] Lampe, C., Ellison, N., & Steinfield, C. (2008). Changes in Use and Perception of Facebook. *Proceedings of ACM Conference on Computer Supported Cooperative Work* , ACM: New York, 721–730.

[11] Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook 'Friends': Social Capital and College Students' Use of Online Social Network Sites. *J. of Computer-Mediated Communication* 12, 1143–1168.

[12] Ajjan, H., & Hartshorne, R. (2008). Investigating Faculty Decisions to Adopt Web 2.0 Technologies: Theory and Empirical Tests. *The Internet and Higher Education*, 11(2), 71-80.

[13] Mason, R. (2006). Learning Technologies for Adult Continuing Education'. *Studies in Continuing Education*, 28(2), 121-133.

[14] Alotaibi, M. & Bull, S. (2012). Using Facebook to Support Student Collaboration and Discussion of their Open Learner Models, *Workshop on Web 2.0 Tools, Methodology, and Services for Enhancing Intelligent Tutoring Systems, Intelligent Tutoring Systems 2012.*

[15] Wang, Q.Y. & Woo, H.L. (2007). Comparing Asynchronous Online Discussions and Face-to-Face Discussion in a Classroom Setting, *British Journal of Educational Technology* 38(2), 272-286.

Analysis of Socratic Method in Computer Supported Dialog Games

Ilker YENGIN[a*] & Sebastian FELLER[a]
[a] *Institute of High Performance Computing, A*STAR, Singapore*

Abstract: In this paper we discuss the Socratic Method as a theoretical framework for a possible dialog game application in ICT. We assess if and if so to what extent the Socratic Method facilitates leaner motivation by highlighting key pedagogical aspects like facilitating self-regulated learning and high-level thinking in the learner.

Keywords: Socratic Method, Dialog Games, ICT supported Dialog Games, motivation, deep learning, knowledge building.

1. Introduction

Ravenscroft [1], amongst others, argues that dialogic interaction through dialog games is an essential part of information and communication technologies (ICT) based education. We discuss to what extent the Socratic Method can be integrated in dialog games. The focus here is on the potential of dialog games as an application of the Socratic Method. For a detailed research on the efficacy of ICT based dialogue game, one can refer to Ravenscroft & Cook [2].

2. Use of Socratic Method in Computer Supported Dialog Games

Our line of argument is thereby based on the relation between the Socratic Method and dialog games. The use of the Socratic Method in computer supported dialog games is effective because it may help to enhance the user's awareness of what is going on in the interaction regarding to reasoning [3] and it supports having a perspective of the other participants [4] which is essential for building social relationships [5] and for being more flexible on incorporating other's mental models [6].

The Socratic Method works in a dialectical and dialogical way where the teacher asks questions to help learners to examine their existing ideas and the validity of their statements. Learners are required to think about their own reasoning and their responses [7]. Teachers may thus use the Socratic Method to improve argumentation skills of learners and help them to be independent and critical thinkers. The teacher becomes a learning partner [8] with a goal to help the student to discover knowledge. Socratic interaction in a dialog may show knowledge gaps across students. While in the original version of the Method questions are always of a disjunctive yes or no type, a new version of the method should run along questions that make the learner find answers which go well beneath the surface of the subject matter.

So what kinds of questions should the teacher ask then? The pertinent literature suggests a number of different options. As space is limited, we cannot discuss all of them here but select Erdogan & Campbell's [8] taxonomy based on Graesser and Person's [10] work in Table 1. Taking a look at this taxonomy, questions 6 to 12 are especially interesting for remodeling the original Socratic Method in terms of active knowledge building and deep reasoning. Questions for definitions, interpretations, or causal consequences, etc. activate the learner's reasoning in terms of larger cognitive patternings like cause and effect, general and specific, or the whole and its parts relations, amongst others.

Table 1: Question taxonomy based on Graesser & Person [10]

Question Type	Description	Example
Open Ended-Question		
(6) Definition	Determining meaning of a concept	What do you think that is ?
(7) Interpretation	A description of what can be inferred from a pattern of data	Now what is another thing that can happen ?
(8)Casual antecedent	Asks for an explanation of the consequences for an event of state	What happens to a balloon if you keep blowing it up?
(9) Casual consequence	Ask for an explanation of the consequences for an event of state	Molting… they slow down and they do not move very much if you have a lazy worm it probably preparing to molt or what else it might be preparing to do?
(10) Enablement	Asks for an explanation of the object or resource that allows an agent to perform an action	So what are we going to do with these potatoes? Where is the seed then?
(11) Expectation	Asks about expectations or predictions	So when do we take it back out ?
(12) Judgmental	Asks about value placed on an idea, advice or plan	

These question types lay the foundation for computer supported dialog games that are based on real-life teacher/learner interactions (Figure 1). For example, the teacher uses the questions in section [p] and [q] (underlined) to probe further into the learner's mindset, getting him to specify his knowledge about the problem at hand in more detail. The first question in line 275 is of the type 'enablement'. The learner is to explicate what kind of action a voltage reading is. He might therefore refer to its function, and prerequisites, for example.

275 *TUT: what do you remember about makiing voltage readings?
276 *STU: you must make sure that the leads are connected to the correct part
277 of the circut and that you must make sure that the positive and
278 negative ends are matched up correctly.
[p]
279 *TUT: good, you correctly mentioned that you should observe polarity,
280 although technically that is not an issue in ac circuits since
281 polarity alternates back and forth.
282 *TUT: but what did you mean about connecting the leads to the correct
283 part of the circuit?
[q]
284 *STU: you have to make sure that the leads go onto the part of the circut
285 that the power flows through so that you can get an accurate
286 measurement of the energy going through whatever it is you are
287 measuring.
[p]/[r]

Figure 1.A real-life teacher/learner interactions in a dialog [11]

In lines 282-83 the teacher requests clarification as to what the learner meant about connecting the leads to the correct part of the circuit in his answer to the initial question. This question is of the 'definition' type. Generating these sorts of questions and structuring the cognitive processing of the learner along these lines in a ICT based dialog game seems to be a promising step towards an effective ICT application for teaching and learning. Dialog lines 282-83 in Figure 1 can be converted into a computer supported dialog game application as in Figure 2. The game provides a list of meta-cognitive moves (e.g. challenge, agree, etc.) from which the tutor can choose. The game then provides a set of questions and openers for the tutor's questions that lay the basis for Socratic questioning. The openers are pre-given expressions that indicate types of cognitive moves expected from the learner.

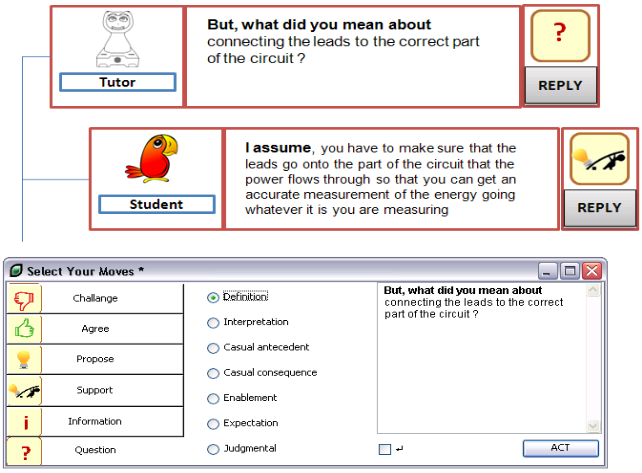


Figure 2. Dialog in Figure 1 can be converted into a computer supported dialog game application

3. Conclusion

Following recent research, we discussed the Socratic Method as a possible dialog game application in ICT. To conclude, we argued that the Socratic Method has potential to support learner motivation and deep learning. Moreover we consider dialog games suitable applications for integrating the Socratic Method into ICT based education.

References

[1] Ravenscroft, A. (2001). Designing E-learning Interactions in the 21st Century: revisiting and rethinking the role of theory. *European Journal of Education*, *36*(2), 133-156. JSTOR.

[2] Ravenscroft, A., & Cook, J. (2007). New horizons in learning design. *Rethinking Pedagogy for a Digital Age-Designing and delivering elearning, H. Beetham and R. Sharpe (Eds), Routledge Taylor and Francis Group*, 207-218.

[3] Ravenscroft, A. (2007). Promoting thinking and conceptual change with digital dialogue games. *Journal of Computer Assisted Learning*, *23*(6), 453-465. Wiley Online Library.

[4] Wegerif, R. (2006). A dialogic understanding of the relationship between CSCL and teaching thinking skills. *International Journal of Computer-Supported Collaborative Learning*, *1*(1), 143-157. Springer.

[5] Turnbull, W., & Carpendale, J. I. M. (2001). Talk and the Development of Social Understanding. *Early Education & Development*, *12*(3), 455-478. Routledge. doi:10.1207/s15566935eed1203_8

[6] Ravenscroft, A., Sagar, M., Baur, E., & Oriogun, P. (2007). Ambient pedagogies, meaningful learning and social software.

[7] Ball, W. H., & Brewer, P. (2000). *Socratic seminars in the block*. Eye On Education, Inc.

[8] Bean, J. C. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. Jossey-Bass.

[9] Ravenscroft, A., & Pilkington, R. M. (2000). Investigation by design: Developing dialogue models to support reasoning and conceptual change. *International Journal of Artificial Intelligence in Education*, *11*(1), 273-298.

[10] McBurney, P., & Parsons, S. (2001). Dialogue games in multi-agent systems. *Informal Logic*, *22*(3).

[11] Real-life teacher-learner interactions (n.d.). Retrieved 1 Sep. 2012, from http://talkbank.org/browser/index.php?url=Tutoring/Circle/Electricity/25.cha

A Case Study of Creativity Development through Scientific Boat Construction

Ruey-Shyy SHIEH
Dept. of Information, Kainan University, Taiwan, ROC
rueys99@gmail.com

Abstract: Fostering students' creativity helps enhance their problem-solving abilities and hands-on skills. The purpose of this study was to cultivate middle-school students' creativity and problem-solving abilities through collaboratively constructing a scientific boat. In addition to boat artifacts, this study also collected interview and classroom observation data. The results indicate that the students were able to exert their imagination to create various types of boats with different materials. It was also found that the students' development of creativity appears to be associated with their preference for free-style thinking and boat construction concepts. Those who were more successful tended to appreciate the free style of imagination for constructing the artifact, and were also more self-motivated. On the contrary, those who were less successful in completing the task expected clearer instructions from the teacher. To more effectively help students develop their creativity, differing amounts of instruction provided to students of various backgrounds is important.

Keywords: Scientific boat construction, creativity development, collaborative learning

1. Introduction

Creativity in education has been a topic of concern since the 1950s [1] as it is seen as an important component for developing students' problem-solving and other cognitive abilities [2]. It is also a core element in advancing science [3]. The purpose of this study is to foster students' creative thinking through constructing an imaginary amphibious boat.

2. Literature Review

Most theories addressing creative thought are grounded on one of the following two assumptions: (1) viewing it as an unconscious, undirected incident, or (2) seeing it as a conscious, controlled occurrence that depends upon some mechanisms, such as analogical reasoning [4]. However, some researchers have found that creative effort involves cycles of conscious and unconscious processing occurring at various stages [5]. Once conscious processing has been triggered, the cycles of information processing and the associated strategies are activated to integrate the existing knowledge structures with new understandings of ideas, which are later translated into action [6].

Rowlands [7] argued that creativity is intrinsically subject to the teaching of the academic disciplines, in addition to novel ideas. Classrooms are generally not seen as creativity-cultivating places due to the lack of appropriate curricula [8] or to the preoccupations of the traditional classroom setting [9]. Amabile [10] emphasized that intrinsic motivation, such as self-achievement, is a prerequisite condition for creativity, and argued that a good learning environment helps cultivate such motivation.

The purpose of this study is to explore what and how students gain from being involved in a creative context and instructional activities aimed at helping them develop their creative thinking skills. Two research questions are addressed in this study:

1. How have the students benefited from engaging in constructing an imaginary boat?
2. What challenges have they encountered in the construction process?

3. Methodology

3.1 Research Context and Research Design

A teacher and 24 7th-grade students in a middle school in Taiwan were involved in the study. In 2011, one of the science projects the teacher asked the students to engage in was constructing an amphibious boat. The students were scheduled to meet every Wednesday for one hour in a designated classroom. The knowledge domain of the project was sciences and life technologies. To allow the students more room for creativity, only basic materials were provided to them, including 1 flat motor, 2 axles, 2 AA batteries, and 1 battery holder. In other words, the students had to design and decide on the remaining materials required to produce a boat, including the material for the hull and wheels, the size, and shape, etc.

3.2 Data Collection and Data Analysis

Three sources of data were collected, consisting of:

- Performance-based artifacts: to examine the students' learning outcomes. A course was built for the boat contest held at the end of the project to assess the students' creativity.
- Class observations: to gather the students' dynamic interactions during their participation in the project.
- Group interviews with the students: to obtain information about the students' learning experiences.

The interview data were transcribed verbatim and were analyzed. The analysis was triangulated with the classroom observation data and the contest results.

4. Preliminary Results

The classroom observation data revealed that the teacher adopted various instructional strategies to help the students fulfill the objectives, including mind mapping activities, six-hat thinking activities, lecturing on associated concepts, and requiring students to submit an end-of-project reflection report. The students, in general, were rather active in participating in the project activities.

Based on the presented artifacts, the students seemed to be able to exert their imagination to create various styles of amphibious boats using different materials, including styrofoam, plastic bottles, pearl board, and cardboard. In the contest, four groups (Groups 1, 3, 4, and 6) out of the eight made it right through the course and took 7, 5, 3, and 9 seconds, respectively. The boats in first and second place are displayed in Figures 3 and 4.

The interview data revealed that the students who succeeded in getting their boat through the course tended to appreciate the free style of imagination for designing and creating their boats; most were also more self-motivated. On the contrary, more than half of the students whose boats failed to successfully complete the course said that they had expected clearer instruction from the teacher, such as the shape of the boat, the size of the wheels, the place for putting the motor, and the area for locating the axles. They also showed less desire to excel, and were found to have devoted less effort to producing their boats.

Figure 3. The 1st-place boat Figure 4. The 2nd-place boat

5. Conclusion

Constructing a scientific boat allows students to engage in both imagination and problem-solving skills. In this study, the students' development of creativity appears to be associated with their preference for free-style thinking and boat construction concepts. As Rowlands [7] contended, in addition to innovative ideas, creativity is constitutionally subject to the teaching of academic knowledge. Creating a moving boat does indeed require some academic disciplines of floating, force, and balancing concepts. Those who anticipated more instruction from the teacher might have reflected that they were less proficient in the associated knowledge, and thus were less devoted to the project. In contrast, those who were able to integrate the attained concepts were also able to translate them into action, in accordance with Martindale's [6] assertion. The study discloses that it is important to provide an adequate amount of instruction to students of various backgrounds. For more advanced, self-motivated students, minimal instruction with maximal room to engage in free thinking is suggested, whereas for less advanced, less confident students, more detailed guidance is necessary.

Acknowledgements

The study was supported by the National Science Council (NSC100-2511-S-424-002- MY2).

References

[1] Craft, A. (2001). Little c creativity. In A. Craft, B. Jeffrey, & M. Leibling (Eds.), *Creativity in education.* London: Continuum.
[2] Plucker, J.A., Beghetto, R.A., & Dow, G.T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39(2),* 83-96.
[3] Moravcsik, M. (1981). Creativity in science education. *Science Education*, *65*, 221-227.
[4] Mumford, M.D. (1998). Creative thought: Structure, components, and educational implications. *Roeper Review, 21(1)*, 14-19.
[5] Sawyer, K. (1992). Improvisational creativity: An analysis of jazz performance. *Creativity Research Journal*, *5*, 253-264.
[6] Martindale, C. (1990). *The clockwork muse: The predictability of artistic change*. New York: Basic Books.
[7] Rowlands, S. (2011). Discussion article: Disciplinary boundaries for creativity. *Creative Education, 2(1),* 47-55.
[8] Archambault, F.X., Jr., Westberg, K.L., Brown, S.W., Hallmark, B.W., Zhang, W., & Emmons, C.L. (1993). Classroom practices used with gifted third and fourth grade students. *Journal for the Education of the Gifted*, *16*, 103-119.
[9] Furman, A. (1998). Teacher and pupil characteristics in the perception of the creativity of classroom climate. *Journal of Creative Behavior*, *32*, 258-277.
[10] Amabile, T. M. (1996). *Creativity in context*. Oxford: Westview Press.

Online Discussion Forums for Mass Events in a Medical PBL Curriculum

Stefan P. MINDER[a*], Michele NOTARI[b] & Sissel GUTTORMSEN SCHÄR[a]
[a]*Institute of Medical Education, University of Bern, Switzerland*
[b]*PHBern, University of Teacher Education, Bern, Switzerland*
*minder@iml.unibe.ch

Abstract: The web-based discussion forums used in the Bernese medical curriculum serve as a means of communication between faculty and students and are well frequented for questions and answers relating to mass events. Using this example, our article takes a look at the critical success factors in implementing e-learning in university curricula and discusses the advantages and disadvantages of a uniform and binding e-learning culture. We address the questions of what influence the discussion forums have on the mass events, and what influences are important in the establishment and maintenance of an e-learning culture. We discuss the online votes from faculty and students on use and significance of the discussion forums.

Keywords: Online discussion forum, medical curriculum, problem-based learning, e-learning culture

Introduction

In a Problem-Based Learning (PBL) curriculum, the communication between teachers and students plays an expanded role, yet for the (remaining) mass events, too, the lecturers are involved in a role change [1], since the students develop and expand on constructivist approaches in the tutorials as well as in the phases of self-study [2]. In this learning process, questions arise that can be in part answered by consulting with textbooks and other sources; on the other hand, uncertainties remain because those sources do not answer all the questions conclusively, and / or because the students are struggling under their misconceptions in the respective field which they cannot readily shake off [3].

Many students show an enhanced need to communicate on professional issues: they want to consult with their lecturers and fellow students to get more input [4]. To that end, exchange opportunities should be institutionalized in PBL. In particular, talking with lecturers who teach exclusively in mass events can only occur outside their lectures.

Communication institutionalized in PBL between students, for the purpose of developing a constructivist approach, takes place in the mandatory tutorials as well as in voluntary, self-organized study groups; in tutorials, small groups (8 to 10 people plus tutor) work out learning objectives based on case reviews. Institutionalized communication between lecturers who teach in mass events and their students occur in specialist consultations (approximately 1-2 hrs. per week and lecturer) which are listed in the class schedule. The lecturers are available during these time windows in their offices for consultations. Another form of contact some students exercise is a brief consultation following a lecture. However, for more than 100 students per year (and event), this time window opens all too rarely and is too short.

The expert consultations have, in addition to the potentially positive social effects of personal contact with lecturers important to some students [5], major organizational disadvantages. Therefore, discussion forums on the internet platform studmed.unibe.ch were launched where students can ask questions and lecturers or other students can answer them. Lessons learned: Contextually appropriate structuring of the discussion forums is crucial for their success [6]. The amount of work for moderators in the forums can be substantial when many students write articles and can only be accomplished through a dedicated approach [7]. When participants perceive themselves as socially exposed through their posts, the need for anonymity in the forum increases [8].

1. Importance of the Online Discussion Forums from the Perspective of the Lecturers and Students

In an anonymous online survey we asked the lecturers and students of the Bernese medical curriculum questions about the online discussion forums, their types of use and frequencies as well as their importance within the curriculum. The responses by lecturers and students in undergraduate studies show (1st to 3rd year, table http://ufive.ch/i2) that the discussion forums as a teaching and learning aid are mostly considered useful and important.

The data concerning the undergraduate studies are of special interest to us since there, the use of discussion forums is unified and interwoven with the conduction of mass events. Do the posts in the discussion forums have a direct impact on the consolidating mass events (syntheses)? 21 (47.7%) lecturers responded with “yes, often”, 10 with “yes, rarely”, 4 with “no”, and 9 could not answer the question. Of 87 lecturers with at least two mass event lectures in the undergraduate curriculum of the past academic year, 44 filled out the online survey. Of the 449 students who were invited twice by email, 150 participated in the survey, of which 103 stated that they deposited technical questions in the forums. 46 students (30.7%) detected an effect of these technical issues on the synthesis-mass events, 11.3% denied that effect, and 26.7% did not know.

It is gratifying to see that the students praised the quality of information flow: the vast majority stated that their questions were answered within 24-48 hours and that the answers were helpful. As a result, it is not surprising that over two thirds of the students rated a hypothetical termination of forum operation as a loss. It is interesting that the students attach more value to responses by lecturers than responses by fellow students. The lecturers underestimate the importance of forums for students: the statement, "the studmed forums are important to the students" was judged by only 31.8% of lecturers as being "largely true". Yet, about half the lecturers rate the forums as fully or partially important to them, and 43.2% believe all or part to be of great educational value.

2. Discussion

We distinguish four dimensions that contain acceptance-promoting elements: the individual will, the social imperative, the personal skills, and situational enablement [9]. These dimensions are supported by cultural elements, usability elements and technology elements. A basic prerequisite for the effectiveness of these success factors, however, is that users have the necessary skills to operate e-learning technologies [10], which is presumed to be the case with students entering university nowadays. Among the lecturers such knowledge and skills vary [11]. Considering all that, e-learning and mass events have the potential to closely work together in synergistic integration. Based on the use of web-based discussion forums in the study of medicine at the University of Bern, we have shown that electronic communication and collaboration, which we generally see as key components of e-learning can cover needs that are not being met in mass events. This applies to both students and teachers: our mass events, each with over a hundred students leave little room for questions from students, and lecturers do not know the students' shakiness with the subject matter. These communication gaps can be closed by the activities in the discussion forums.

The synergistic integration of mass events with e-learning in the Bernese medical curriculum is complemented by the systematic publication of the lecture slides and the release of 60% of the lectures of the third academic year as video podcasts. This kind of unified, mandatory use of e-learning has great potential for success if the participation (use numbers and penetration of the curriculum) is taken as indicator of achievement. This raises the question of how to weight the participation as success indicator; participation is the critical factor in the use of discussion forums, it is a multiplier of effects: increased motivation for learning, learning success, and efficiency through discussion forums.

It is apparent from some free-text responses by lecturers in the survey that teachers comply with an e-learning culture put in place, even if they prefer other solutions, not least because of the expectations held by students and the resulting social pressure on the lecturers. This brings us to a difficult situation: The success of the unified, mandatory e-learning implementation in the curriculum versus academic freedom of the lecturers. We support the critically evaluated, unified, mandatory use of e-learning applications, but emphasize that they must remain subject to continuing assessments, and that the teaching staff within the practiced e-learning culture must have enough degrees of freedom. The existence of a wide range of both dedicated supporters and dedicated opponents of a practiced e-learning culture is almost culture-immanent, but not an insurmountable obstacle to the implementation of a uniform e-learning strategy.

References

[1] Haith-Cooper, M. (2000). Problem-based learning within health professional education. What is the role of the lecturer? A review of the literature. Nurse *Education Today, Volume 20*(4), *267-272.*

[2] Dochy, F., Segers, M., Van den Bossche, P., Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction, 13*(5), *533-568.*

[3] Greening, T. (1998). Scaffolding for success in PBL. *Medical Education Online, 3, 4, http://www.meded-online.org.*

[4] Dolmans, D., De Grave, W., Wolfhagen, I., van der Vleuten, C. (2005). Problem-based learning: future challenges for educational practice and research. *Medical Education, 39*(7), *732-741.*

[5] Solimeno, A., Mebane M., Tomai, M., Francescato, D. (2008). The influence of students' and teachers' characteristics on the efficacy of face-to-face and computer supported collaborative learning. *Computers & Education, 51*(1), *109-128.*

[6] Brace-Govan, J. (2003). A method to track discussion forum activity: The Moderators' Assessment Matrix. *The Internet and Higher Education, 6*(4), *303-325.*

[7] George, S., Labas, H. (2008). E-learning standards as a basis for contextual forums design. *Computers in Human Behavior, 24*(2) *Part Special Issue: Cognition and Exploratory Learning in Digital Age, 138-152.*

[8] Suzuki, L., Calzo, J. (2004). The search for peer advice in cyberspace: An examination of online teen bulletin boards about health and sexuality. *Journal of Applied Developmental Psychology, 25, 6, Developing Children, Developing Media - Research from Television to the Internet from the Children's Digital Media Center: A Special Issue Dedicated to the Memory of Rodney R. Cocking, 685-698.*

[9] Schönewald, I. (2006). Die Gestaltung von Supportstrukturen als Element des Change Managements an Hochschulen. Universität St. Gallen, Lecture, Bonn, 9 November, 2006.

[10] Eleven, E., Karuovic, D., Radulovic, B., Jokic, S., Pardanjac, M., (2012). Development of distance learning, independent learning and modern education technology. *Technics Technologies Education Management 7*(1), *111-121.*

[11] Al-Fudail M., Mellar H. (2008). Investigating teacher stress when using technology. *Computers & Education, 51*(3), *1103-1110.*

Students' understandings of collaborative discourse: A knowledge building approach

Feng LIN[a] **& Carol K. K. CHAN** [a*]
[a]*Faculty of Education, The University of Hong Kong, HK*
*ckkchan@hku.hk

Abstract: The study assessed students' understanding of collaborative discourse and examined the role of knowledge building environment augmented with reflective assessment in fostering more sophisticated views of collaborative discourse among students. One language teacher and 31 high school students in Hong Kong participated in this study. Data were collected from students' pre-post essay writings about "*what is good discussion*". This study identified four patterns of understandings about collaborative discourse ranging from less to more sophisticated, and found that students' understandings of collaborative discourse became more sophisticated after their engagement in knowledge building.

Keywords: Knowledge building; views of collaboration; epistemic cognitions

Introduction

Computer-supported collaborative learning (CSCL) has received much attention in the recent decade. How can we improve students' online participation and enhance the quality of their online discourse is one of the major concerns in this line of research. Among the few studies that address this issue, some study (e.g. [2]) found that students' views of collaboration predict their online participation. The more students' views of collaboration are aligned with knowledge building, the more likely they are going to participate in the Knowledge Forum (a computer supported platform for knowledge building). This study shed light on the importance of examining and influencing students' views of collaboration from a knowledge building perspective. Examining how student understand good collaboration, particularly collaborative discourse, is important also because it reflects a kind of epistemic cognition that focuses on the social aspects of knowledge and knowing. "Epistemic cognitions"(personal epistemology) is an area that studies individual's views about the nature of knowledge and knowing. Its importance has been established by a growing number of literature, as it is constantly reported to be related to students' learning process and outcome [3][5]. With the development of learning theories that became more focused on the social aspect of knowledge and knowing, epistemic cognitions researchers also proposed to take a socio-cognitive and socio-cultural perspective to understand individuals' beliefs about the nature of knowledge and knowing [6]. Some studies responded to this call by addressing the cultural relevance of epistemic cognitions (e.g., [4]). However, the social aspect is still neglected. Therefore, one of the main purposes of this study is to understand the social aspect of individual's epistemic cognitions, in particular, to examine how students understand the collective aspect of knowledge and knowing by looking at how they understand collaborative discourse. Meanwhile, the study also tries to examine the change of these epistemic cognitions in knowledge building environment. Two research questions are addressed in this study: (1) what are middle school students' understandings of collaborative discourse? (2) Do students' views of collaborative discourse become more sophisticated after engaging in knowledge building?

1. Methodology

One high school language teacher and their Form 6 students (n=31) in Hong Kong participated in this study. The teacher had 7 years Language teaching(Chinese) experience, and had used knowledge building pedagogy in language teaching for 6 years. The intervention lasted for half a year (2010-2011). It was implemented once every week; each time it lasted for about 2 hours. The topics covered in knowledge building discussion included "Chinese culture and literature" and "current issues". The learning environment was designed based on the transformed knowledge building pedagogy[1] aligning with reflective assessment. The teacher adapted this pedagogy for language learning: (1) Nurturing collaborative culture. Students read the learning material and discuss it in the classroom. The ideas are recorded in public places (e.g., papers) and became the objects for inquiry; (2) Developing knowledge building inquiry. Questions of interest were moved to knowledge forum where they made collaborative inquiry into emergent ideas and made constructive use of authoritative information to improve the community knowledge; (3) Deepening the inquiry. Students were encouraged to synthesize their ideas and deepen the collaborative discourse; (4) Portfolio assessment. Portfolio assessment was used to capture and also to scaffold the advancement of the community knowledge. Students identified the best cluster of notes and justified the reason behind their selection according to knowledge building principles. In order to assess the change of students' understanding of collaborative discourse, each student was asked to respond to a question in writing "*what is good discussion*" before and after their engagement in knowledge building.

2. Analyses and Results

2.1 Students' Understandings of collaborative discourse

Both students' pre and post writings about "*what is good discussion*" were examined to identify the patterns of understandings about collaborative discourse. Four levels of understanding were identified and developed into a 4-point scale ranging from simple to more sophisticated views.

Level 1: simple view about discussion. At this level, students have some general ideas about the behavioral features of good discussion, such as listening, responding, communicating, and having good attitude: *"In order to have good discussion….we need good communication, answer other people's question politely, have eye contact, respect each other….."(pre12)"*

Level 2: Students elaborate on ideas such as give examples while you discuss, find some relevant references, and question other's response: *"……For a good discussion, there must be someone in charge of taking notes and recording the conclusions and questions. Members need to finish the discussion on time, and find some relevant references before the discussion." (6bpre09)*

Level 3: Students mention multiple unelaborated KB (knowledge building) ideas. They have some limited and fragmented understanding about the importance of "new information", "summary", "clarification", "questioning", "diversity of idea" for good discussion, and did not know how they work together for improving the community knowledge: *"what is a good discussion: (1) use more examples and authoritative information; (2) bring up some small questions to lead classmates to think, deepen the discussion; (3) extend others' response to deepen the discussion; (4) we can do summary in the process to make the discussion clear; (5) point out other' misconception, so that we can have a focus on the discussion…." (6bpost20)*

Level 4: At this level, students not only mention about some KB ideas, but also make connections between them. They have more coherent understanding about the role of "shared goal", "constructive use of information", "diversity of ideas", "rise above", and "social metacognition (clarifying, questioning, reflection)" in improving the community knowledge. They have more awareness of the collective nature of knowledge and knowledge construction process: *"I think, for a good discussion, there must be a clear theme, even though every one thinks differently, for example, all the group members need to have a shared goal, but they can bring up different methods and ideas. Meanwhile, it entails synthesis of different ideas, and a good conclusion. The discussion does not necessarily need a summary, but it needs to show signs of analyzing and synthesizing other people's ideas, and then deepening their ideas….." (6bpost15)*

2.2 Changes in Students' Understandings of collaborative discourse

Students' pre and post essay about "*what is good discussion*" were analyzed and rated on a 4-point scale as illustrated in session 2.1. A second rater will be employed later for establishing the inter-rater reliability. Paired t- tests showed that students made statistically significant changes towards more sophisticated views of good discussion, $t(30)= 6.875$, $p<0.001$ [pre test:M=1.74 SD=.68; post test: M=2.77 SD=.61]. Specifically, most students held simple views of good discussion in the pre test (87.1% are at level 1 and 2), after the knowledge building instruction, most students' understandings of good discussion were at level 2 and 3 (90.3%).

3. Discussion

The purpose of the study is to provide an initial look at the characterization and development of students' understandings of collaborative discourse in the knowledge building environment augmented with reflective assessment. Four patterns of understandings of collaborative discourse were identified, ranging from simple to more sophisticated views which are more aligned with knowledge building. Quantitative analysis showed that students did make significant changes in their understanding of good discussion after experiencing knowledge building. More work is needed to examine how the change happens. To conclude, this study is trying to contribute to our understanding about the collective aspect of epistemic cognitions, as students' views of good discussion reflected how they understand the collective nature of knowledge and knowing. It also shed light on the possibility of influencing students' epistemic cognitions with knowledge building pedagogy and reflective assessment. Future studies can be conducted to understand whether and how students' understandings of collaborative discourse predict their online discussion.

References

[1] Chan, C. K. K. (2009). classroom innovation for the Chinese learner: Transcending dichotomies and transforming pedagogy. In C. K. K. Chan & N. Rao (Eds.), *Revisiting the Chinese learner : changing contexts, changing education* (pp. 169-210). Hong Kong Comparative Education Research Centre, The University of Hong Kong,

[2] Chan, C. K. K., & Chan, Y. Y. (2011). Students' views of collaboration and online participation in Knowledge Forum. *Computers & Education, 57*(1), 1445-1457.

[3] Dahl, T. I., Bals, M., & Turi, A. L. (2005). Are students' beliefs about knowledge and learning associated with their reported use of learning strategies? *British Journal of Educational Psychology, 75*(2), 257-273.

[4] Hofer, B. K. (2008). Personal epistemology and culture. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: epistemological studies aross diverse cultures* (pp. 3-22). Dordrecht: Springer.

[5] Phan, H. P. (2008). Exploring epistemological beliefs and learning approaches in context a sociocultural perspective. *Electronic Journal of Research in Educational Psychology, 6*(3), 793-822.

[6] Tabak, I., & Weinstock, M. (2008). A Sociocultural Exploration of Epistemological Beliefs. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: epistemological studies across diverse cultures* (pp. 177-195*)*. Dordrecht: Springer.

An Extraction Technique for Presentation Schema embedded in Presentation Documents

Shinobu HASEGAWA[a*] & Akihiro KASHIHARA[b]
[a] *Center for Graduate Education Initiative, JAIST, Japan*
[b] *Graduate School of Informatics and Engineering,*
The University of Electro-Communications, Japan
*hasegawa@jaist.ac.jp

Abstract: The main topic addressed in this paper is to help a novice graduate/undergraduate student compose his/her presentation document by means of presentation schema that represents heuristics for presenting research contents to be shared by laboratory members. The key idea is to propose a model of presentation structure, which represents roles of and sequences among presentation slides included in the documents with metadata. Following this model, the presentation schema is defined as a typical presentation structure for the laboratory members. This paper accordingly introduces a technique based on association rule mining for automatically extracting the presentation schema from the repository of the documents accumulated in the laboratory. In addition, we report case studies for investigating how to configure the thresholds of the mining and how the schema extracted is valid in comparing the ones between different laboratories.

Keywords: Presentation Schema, Presentation Semantics, Association Rule Mining

Introduction

In our daily research activities, composing presentation documents is one of the most important skills so that researchers and students can report the research findings with well-organized representation. However, it is quite difficult for novice graduate/ undergraduate students in the laboratory to compose the well-organized presentation documents since they have few experiences and heuristics of constructing the presentation structure to be shared by the laboratory members [1]. We call such presentation structure a presentation schema. The presentation schema would often vary according to diverse factors about presentation context, such as presentation time limitation, audiences, research domain, and presentation philosophy in the laboratory. Therefore, it is not so easy for the researchers to prepare such schema for the novices as tangible standards in advance. The final goal of our research is accordingly to help the novices develop the skill for composing the presentation documents by means of the presentation schema that could be extracted from the presentation documents accumulated by the laboratory members.

The main issue addressed in this paper is how to extract the presentation schema automatically from a certain number of the presentation documents the laboratory members have composed. We first introduce a model of the presentation structure with metadata corresponding to each presentation slide [2]. Following this model, we second utilize a machine learning technique to analyze and extract the presentation schema, especially role of and sequence among the presentation slides, from the repository of the documents attached with the metadata in advance. The extracted schema could become a scaffold for the novices to learn the presentation composition skill practically because they can become aware of the typical presentation structure of the laboratory to compose their documents. Scaffolding with presentation schema is accordingly viewed as a part of laboratory education [3]. In addition, each laboratory has its own presentation philosophy. It means that the laboratory would also have its own presentation schema. There is accordingly a great need for extracting the schema automatically with machine learning technique from presentation documents accumulated by the laboratory members.

1. Framework

1.1 Presentation Document Composition Skill

A skill in composing presentation documents is an important research one for brushing up the research itself in laboratory meetings and for reporting the research findings in international/domestic conferences. They generally include a number of slides, which present the research contents. In order to compose the presentation document well, it is necessary (1) to divide the research contents into the slides and (2) to sequence the slides in an understandable way as shown in Figure 1.

On the other hand, it is difficult for the novices to divide their research contents into a number of the presentation slides and to sequence the slides since such presentation structure is often embedded in each document and they have little knowledge about the structure specifying what to present and what order to present logically. Therefore, reading good presentation document is not enough to learn how to compose the presentation documents. In addition, expert researchers are not always good teachers for composing the presentation documents. Of course, they could point out and fix inappropriateness of the presentation documents composed by the novices. But, it is not easy to teach the presentation composition skill directly to the novices. In traditional laboratory education, such skill could be heuristically acquired through daily research activities as cognitive apprenticeship [4].

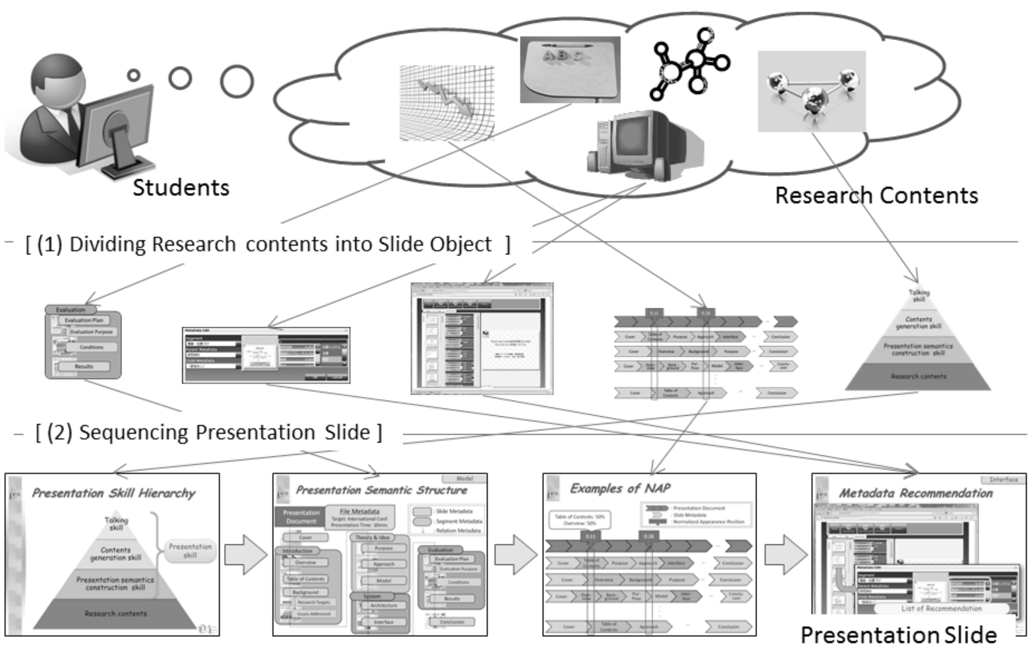


Figure 1. Overview of Presentation Document Composition Skill

1.2 Presentation Schema and Presentation Structure Model

In this paper, the presentation schema is represented as a typical presentation structure, which implies some heuristics for composing the documents in the laboratory manner. The presentation schema could provide the novices with how to divide the research contents into a number of the presentation slides and how to construct the presentation structure that expresses roles of and sequences among the slides. However, it would be difficult to specify such presentation schema since it is often embedded in the presentation documents accumulated in the laboratory. Our challenge is to extract the presentation schema from the repository of the presentation documents automatically.

In order to deal with the presentation structure and schema explicitly, we provide a presentation structure model which uses three types of metadata for presentation slides as shown in Figure 2 [2]. Slide metadata represent the role that each slide included in the presentation document plays in presenting the research contents. Such metadata does not necessarily correspond to the slide title. Some of them vary according to the presentation context. Others are nested since these slide metadata often appear as compound metadata in one slide. Segment metadata also represents the section of the presentation document that several slide metadata compose for presenting the research contents. We have defined four kinds of main segment metadata. Each segment metadata is associated with several slide metadata in advance. File metadata represent some attributes of the presentation context, which includes the presenter information and presentation contextual information.

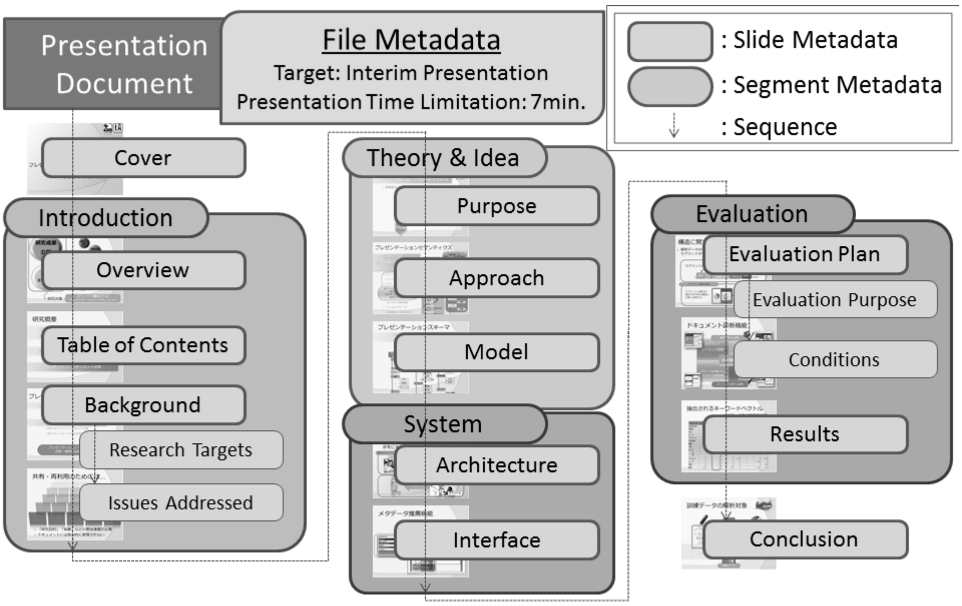


Figure 2. Presentation Structure Model

1.3 Related Work

Kohlhase [5] developed CPoint as a semantic PowerPoint extension that allows the authors to enrich PowerPoint documents by means of domain knowledge annotation and concept mapping. Hayama et al. [6] proposed an automatic approach for generating presentation slides from a technical paper. Li and Chang [7] developed the management model and tools that enable users to better exploit and transfer presentational knowledge assets for representing the domain knowledge.

In spite of the significance of the presentation schema and structure, each of these researches did not deal well with such information embedded in the presentation documents. In this paper, we accordingly address the issue of how to extract the presentation schema from the presentation documents accumulated from the laboratory members as training data, which are attached in advance with the metadata.

2. Extraction Technique for Presentation Schema

2.1 Basic Concepts of Association Rule Mining

Association rule mining is one of well-used techniques of data mining in various areas, which was first proposed by Agrawal et. al. [8]. It aims to extract frequent patters and casual structures among sets of items in the transaction databases [9]. In this section, we describe general definitions of the association rule mining to prepare our schema extraction algorithm.

Let $I = \{i_1, i_2, \cdots, i_n\}$ be a set of n distinct items, $T = \{t_1, t_2, \cdots, t_m\}$ be a set of m different transaction records, where each $t_m \subseteq I$. An association rule is indicated by the form of $X \Rightarrow Y$, where $X, Y \subset I$ are sets of items called item sets, and $X \cap Y = \phi$. X is called antecedent while Y is called consequent, the rule means X implies Y.

There are two important basic measures for association rules, a support described in $sup(X \Rightarrow Y)$ and a confidence described in $conf(X \Rightarrow Y)$. $sup(X \Rightarrow Y)$ is defined as the proportion of the number of transactions that contain $X \cup Y$ to the total number of transactions in *T*. $conf(X \Rightarrow Y)$ is also defined as the proportion of the number of transactions that contain $X \cup Y$ to the total number of transactions that contain *X*.

In regard to such $X \Rightarrow Y$, thresholds of support and confidence are usually predefined by users to extract those rules that are important. These thresholds are called a minimal support described in *min_sup* and a minimal confidence described in *min_conf* respectively. The association rules are finally extracted as the item sets that satisfy both *min_sup* and *min_conf*. However, there are several well-known problems in setting the thresholds [6]. The lower the thresholds are, the larger the numbers of rules are extracted, which are difficult to recognize. The higher the thresholds are, the smaller the numbers of just known rules are extracted.

2.2 Schema Extraction Algorithm

Our schema extraction algorithm aims to extract the typical semantic stricture as the presentation schema based on appearance order of the slide metadata of the presentation documents accumulated from the laboratory members, which are attached in advance with the metadata. Therefore, we adopt the association rule mining as shown in Figure 3, where *I* is a set of all kinds of slide metadata, *T* is a set of all presentation documents to be analyzed, *X* is an arbitrary slide metadata appeared in a certain presentation document as an antecedent, and *Y* is a slide metadata appeared next to *X* in the presentation document as a consequent. Suppose *conf(“Overview” ⇒ ”Background”) = 50%* and *sup(“Overview ⇒ ”Background”) = 33%*, it means that *50%* of *“Background”* slides are next to *“Overview”* slides and *33%* documents include such order relation.

In usual association rule mining, antecedent *X* and consequent *Y* are able to include multiple items but are not able to specify the order relation among these items. Therefore, our algorithm restricts the number of items to one per each *X* and *Y*, which means that *X* and *Y* only contain one slide metadata. This makes it possible to extract partial order relations, and to represent whole sequence of presentation schema by accumulating such partial relations as shown in lower right of Figure 3.

In addition, the infrequently appeared metadata are discarded in order to reduce amount of calculation. *freq*(*Z*), where $Z \subseteq I$, is defined as the proportion of the number of transactions that contain *Z* to the total number of transactions in *T*, and the threshold of frequency described in *min_freq* is also predefined. Such preprocessing would make it simple for the novices.

Based on the above assumptions, the algorithm contains the following steps:

Step 1. The algorithm extracts a set of frequently-appeared metadata that have *freq*(*Z*) larger than or equal to *min_freq*. Suppose *min_freq* is *40%* in Figure 3, then metadata *TOC, Approach*, and *SubCover* are discarded.

Step 2. It extracts partial order relations $X \Rightarrow Y \ (X, Y \subseteq Z)$ that satisfy both *min_conf* and *min_sup*. Suppose *min_conf = 40%* and *min_sup=20%*, the relation between *Cover* and *Concept* is discard.

Step 3. It composes a presentation schema diagram by combining the extracted metadata and relations. In the diagram as shown in Figure 3, the nodes are the slide metadata left in Step 1 and the links show the order relations left in Step 2. The loops mean dual-ordered relations such as *Background* and *Issue*, which have the links from node *Background* to *Issue* and from node *Issue* to *Background* at the same time.

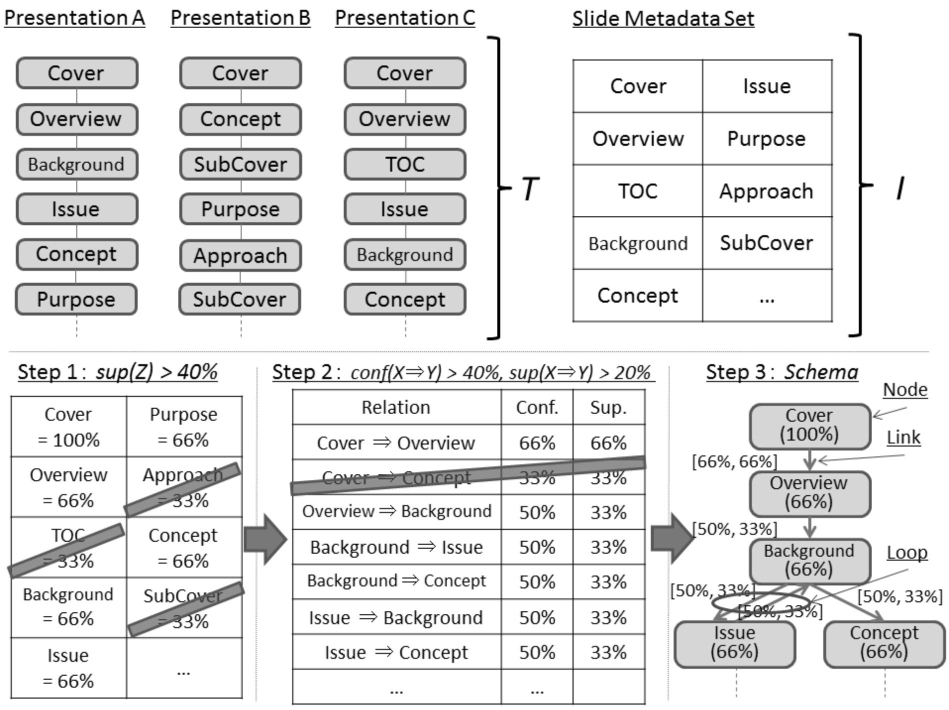


Figure 3. Overview of Schema Extraction Algorithm

3. Case Studies

This section describes case studies which investigated how to configure the thresholds of frequency, support and confidence in the association rule mining for extracting presentation schema, and compared the presentation schemas between different laboratories, audiences, and presentation time limitations since the presentation schema would vary according to such factors. The followings are detail information for sets of the presentation documents in these case studies.

The presentation documents accumulated in Laboratory A were final versions of the ones for graduation research of *30* undergraduate students belonged to the laboratory where they focused on development of support systems for self-directed learning, research activity, and experiential learning. The audiences of the presentations were faculties and students of their affiliation of the university, and the presentation time was *7* minutes. These documents were annotated in *30* kinds of the slide metadata to each slide by an experienced researcher in advance. The average and standard deviation for the number of the slides of the documents were *20.1* pages and *3.68*.

The presentation documents accumulated in Laboratory B were also final versions of the ones for domestic conferences of *15* graduate students or researchers belonged to the laboratory where they focused on development of web-based learning support systems and practice of distance learning systems. The audiences of the presentations were related filed researchers, and presentation time was *15 - 20* minutes. These documents were also annotated in *34* kinds of the slide metadata to each slide by the experienced researcher in advance. The average and standard deviation for the number of the slides of the documents were *22.9* pages and *3.93*.

3.1 Case Study 1: Analysis for thresholds

In order to consider the thresholds, we first investigated how the numbers of nodes, links, and loops included in a presentation schema diagram changed by values of *min_freq*, *min_sup* and *min_conf*. Figure 4 compares the numbers of them extracted by the thresholds on the abscissa in proposed schema extraction algorithm. From the results of our previous work [10], we have ascertained that *min_freq* could be set as the same value and *min_sup* could be set as the half value of *min_conf*. In this case study, we also followed this to set these thresholds.

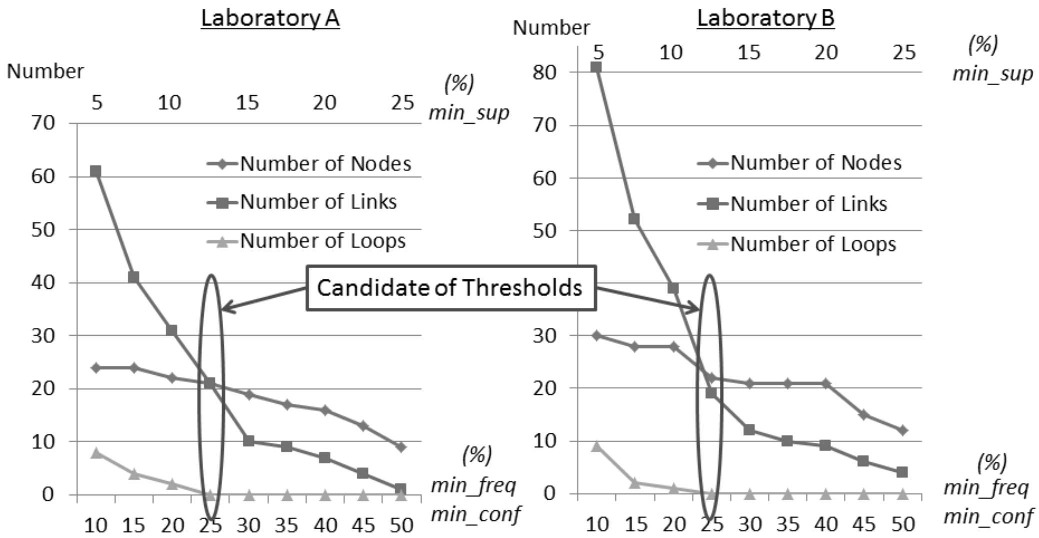


Figure 4. Numbers of Extraction by Changes in Thresholds

In case that the thresholds were sufficiently small, Figure 4 shows the numbers of links are larger than the number of nodes in both laboratories. The larger the thresholds were, the smaller the numbers of nodes, links and loops were. We can see the points (around *min_freq = min_conf = 25%* and *min_sup=12.5%*) at where the numbers of nodes and links were reversed and the numbers of loops were zero. In considering application of the presentation schema, too many links and loops may confuse the novices. Therefore, these points can be important candidates for setting the thresholds. In other words, suitable presentation schema could be obtained by finding out such points to set the thresholds.

3.2 Case Study 2: Assessment of validity for mining technique

The purpose of this case study was to assess the validity of the proposed mining technique by comparing the presentation schemas between Laboratory A and B. Figure 5 illustrates both presentation schema by setting *min_freq* = *min_conf* = *25%* and *min_sup* = *12.5%*. Values in round brackets are probabilities of appeared metadata *freq*(*Z*), and values in square brackets are probabilities of confidence $conf(X \Rightarrow Y)$. Comparing both schemas, for example, the schema regarding to *"Evaluation"* from Laboratory B was different from the one from Laboratory A. This showed a capacity of the presentation schema to represent importance of evaluation in presentation for the domestic conferences. In addition, the schema from Laboratory A tended to have a main path for making smooth presentations of graduate research. On the other hand, there were two paths found in the early segments of the schema from Laboratory B. One path was ordered by *Background*, *Issue*, *Purpose*, *Approach*, and *Technology*. Another was ordered by *Situation*, *Theory*, and *Model*. The reason would be that these presentations included not only researches for system development but also for classroom practice. Some slide metadata did not have any arrows to indicate transition. This showed there are no significant (over thresholds) transitions from the metadata because such metadata had different position in the presentation documents. Following the above consideration, we can say that the presentation schemas extracted satisfy specific conditions of the presentation contexts such as the research domain and/or philosophy each laboratory has. The proposed technique accordingly seems to be valid.

4. Conclusion

This paper has described the presentation structure model and proposed a fundamental technique to represent the presentation schema automatically by accumulating partial order relations extracted with the association rule mining. The diagram of roles of and sequence among the slides would enable the novices to be aware of the presentation schema in the laboratory's manner explicitly. Accordingly this is one of scaffolding ways for the novices to learn the presentation composition skill practically.

We have also discussed preliminary case studies. The results indicate a reasonable setting approach for the thresholds of the mining, and description capability of the schema which depends on the presentation context.

In the near future, it will be necessary to try out the proposed technique to different domain of laboratories. In addition, our research group proposed metadata recommendation, diagnosis, and learning services [2, 11] as the previous work. However, these previous services did not explicitly deal with the presentation schema, especially sequence of the presentation slides. Therefore, we will have to evaluate effectiveness of these services by adding the concept of presentation schema in a more detail.

Acknowledgements

This work is supported in part by Grant-in-Aid for Scientific Research (B) (No. 23300297), Grant-in-Aid for Scientific Research (C) (No. 23501141) from the Ministry of Education, Science, and Culture of Japan.

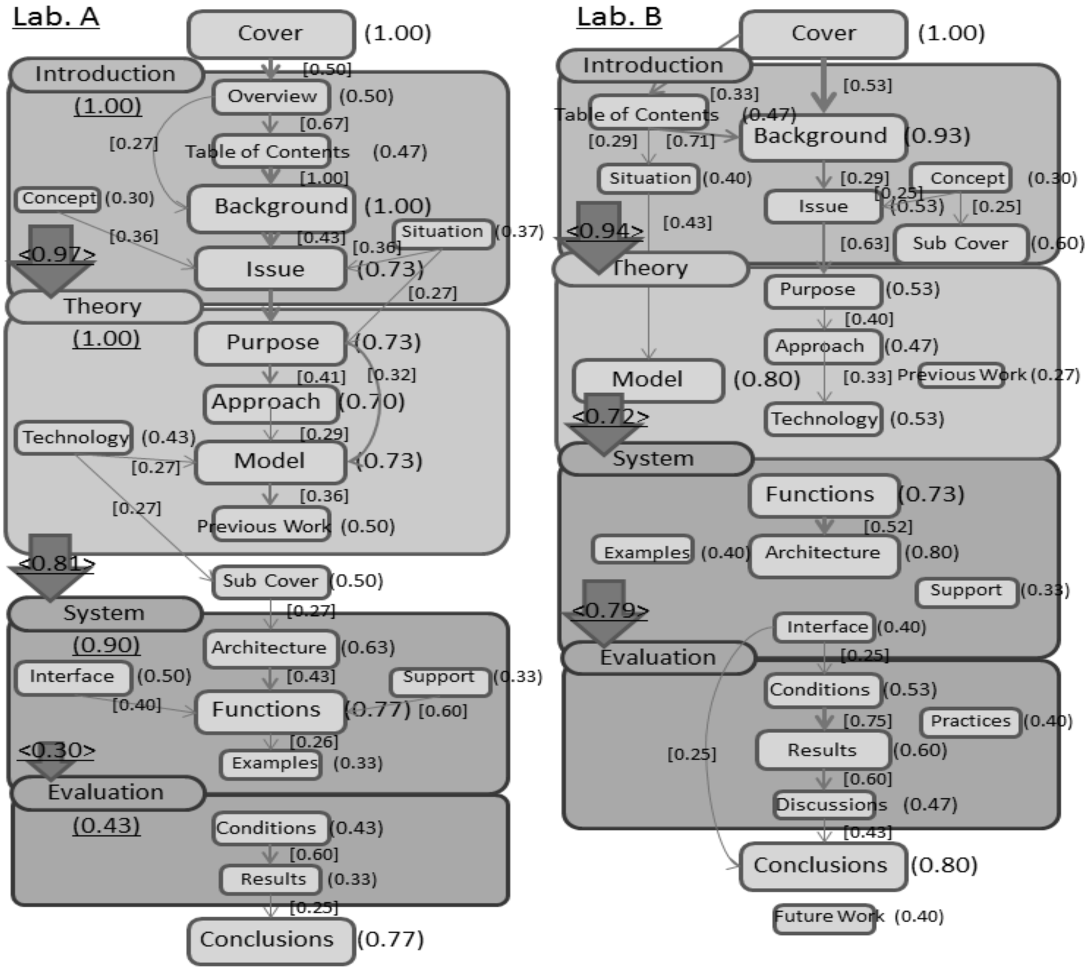


Figure 5. Results of Extracted Presentation Schemas

References

[1] Kirschner, P. A. (1988). The laboratory in Higher Science Education, Problems, Premises, and Objectives, Higher Education, 17, No 1, pp. 81-90.

[2] S. Hasegawa, and A. Kashihara. (2011). Recommendation and Diagnosis Services with Structure Analysis of Presentation Documents, Proceedings of KES2011, Part I, LNAI 6881, pp. 484-494.

[3] Convertino et al. (2004). A laboratory method for studying activity awareness, Proc. of the Third Nordic Conference on Human-Computer Interaction, pp. 313-322.

[4] Collins, A. (2006). Cognitive apprenticeship: The Cambridge Handbook of the Learning Sciences, R.Keith Sawyer (Ed.), Cambridge University Press, pp.47-60.

[5] Kohlhase. (2007). Semantic PowerPoint: Content and Semantic Technology for Educational Added-Value Services in MS PowerPoint, Proc. of World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA2007), 3576-3583.

[6] T. Hayama, H. Nanba, and S. Kunifuji (2005). Alignment between a technical paper and presenta-tion sheets using a hidden Markov model, Proc. Active Media Technology 2005, pp.102-106.

[7] S. T. Li, and W. C. Chang. (2009). Exploiting and transferring presentational knowledge assets in R&D organizations, Expert Systems with Applications 36, pp. 766-777.

[8] Agrawal, R., Imielinski, T., and Swami, A. N. (1993). Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 207-216.

[9] Kotsiantis, S., Kanellopoulos, D. (2006). Association Rules Mining: A Recent Overview, GESTS International Transactions on Computer Science and Engineering, Vol.32(1), pp.71-82.

[10] K. OTA and A. KASHIHARA. (2010). Mining Collective Knowledge for Reconstructing Learning Resource, Proc. of ICCE2010, pp.104-106.

[11] K. Saito, A. Tanida, A. Kashihara, and S. Hasegawa. (2010) An Interactive Learning Environment for Developing Presentation Skill with Presentation Schema, Proc. of E-Learn2010, pp. 2696-2703.

A novel approach to monitoring and creating significant learning experiences using social tag cloud navigation

Jun Ming Chen[a*], Meng Chang Chen[b], Yeali S. Sun[a]
[a]*Department of Information Management, National Taiwan University, Taiwan*
[b]*Institute of Information Science, Academia Sinica, Taiwan*
*d96725001@ntu.edu.tw

Abstract: Prior knowledge learning for increasing the efficiency of the learning processes is an important issue. Traditional studies on prior knowledge generation during language learning have focused on providing additional supplemental materials from reading materials that are manually generated by educators. This is time-consuming and hence personalized prior knowledge recommendation is difficult to perform and monitor. To cope with these problems, we purpose a novel tag-cloud visualization learning approach to automatically monitor running activities and student status. In addition, we incorporate cloud tags into a prior knowledge learning system (TAK), which provides students an engaging way to reinforce meaningful topics, identifies suitable supplementary materials through tag cloud navigation, and helps students reevaluate their reading process. Our experimental results demonstrate that our approach not only significantly improves the efficiency of prior knowledge learning but also helps teachers assist students in improving their reading ability.

Keywords: Social tagging, Reading comprehension, Information retrieval, Teaching and Tutoring

Introduction

In the past, effective reading comprehension requires the integrated interaction of derived text information and preexisting reader knowledge [6, 10], especially with learners of foreign languages such as English. Studies have found that students who are unable to link new knowledge with prior knowledge have problems understanding, recalling, and accessing the new knowledge later [3, 4, 9]. This further suggests it is important to assist students in obtaining relevant prior knowledge, as this can enable them to engage meaningfully with their learning material.

Additionally, keeping students focused is a concern for many teachers. Several studies have demonstrated that teachers, due to either a lack of administrative support or time constraints, have difficulty evaluating student literacy effectively [1, 2, 3, 4]. In some cases, it is very time-consuming to complete the evaluation process, and this leads to significant learning obstacles for students. Thus, identifying and understanding the level of a student's knowledge is important for teachers to help students learn efficiently [5].

To cope with this problem, this paper proposes a novel tag-cloud visualization learning approach. This novel approach uses a social tagging analysis mechanism and tag cloud visualization, which gives students an overview of the knowledge represented and helps them quickly grasp the structure and concepts of English articles. Moreover, it also provides a monitoring interface that enables teachers to navigate to potentially relevant information, and assists teachers in analyzing interactions from student tag behaviors and evaluating student learning performance.

1. Background and motivations

1.1 A typical Web 2.0 phenomenon: Tag cloud

One of the ways in which the Internet helps facilitate knowledge conversion and sharing is through the use of Web 2.0 social tagging. Tags usually represent words or phrases that enable users to easily add metadata to online content [1]. Tag clouds have emerged as an important new interface paradigm, quickly gaining popularity in social information sharing sites such as Flickr and Delicious, which need to find visually appealing ways to summarize vast amounts of information. Fig. 1 shows an example of a tag cloud on Delicious, where tags on the site are organized by popularity.

Figure 1: Example of a tag cloud (Extracted from Delicious)

Recently, researchers are increasingly turning to tag applications such as social networking sites to enhance classroom learning and develop a new generation of learning architecture [1, 2, 3, 4]. Thus, by aggregating tag annotations and corresponding resources from documents, the tag cloud provides navigational clues and helps people easily browse and discover on-line resources [11]. Meanwhile, a tag cloud can also provide learners a visually appealing picture of main themes by using tags to visually analyze the frequency of words in a text. An extension application to tag clouds involves using semantic web technologies to generate tag clouds of semantic concepts, which provides a new way to extract meaningful learning results and analyze learning behavior. This information is helpful to teachers for evaluating student learning achievements and determining learning status.

1.2 A tag-based prior knowledge recommendation system (TAK)

Several studies have been conducted to analyze the learning behaviors of students by using Web 2.0 social tagging techniques to collect implicit information for reading comprehension [1, 2, 3, 4, 7]. A tag-based learning environment called TAK [3, 4] has been developed to improve prior knowledge construction and learning. This system provides effective recommendations, especially through social tagging to spread prior knowledge generation. However, it is difficult for teachers to observe ongoing, running activities and student behavior without assistance.

To cope with the above problems, in the following sections, a novel tag-cloud visualization learning approach is proposed and implemented to extend the TAK system into an effective and efficient mechanism for assisting teachers in tracing and analyzing the information search behaviors of students. Moreover, this study also attempts to harnesses Web 2.0 principles to find new ways for using tag clouds to get a quicker grasp of the main focus of reading materials.

2. I-Cloud: A tag-cloud visualization learning approach

To improve English reading comprehension and assist teachers in monitoring student progress, this study proposes an integrated tag-cloud based learning and monitoring approach, based on fundamental skills of analysis, contextualization, conceptualization, and visualization, called I-cloud. Fig. 2 illustrates the framework of our approach, which consists of three parts: data preprocessing, tag cloud for social navigation, and tag cloud for learning alarm.

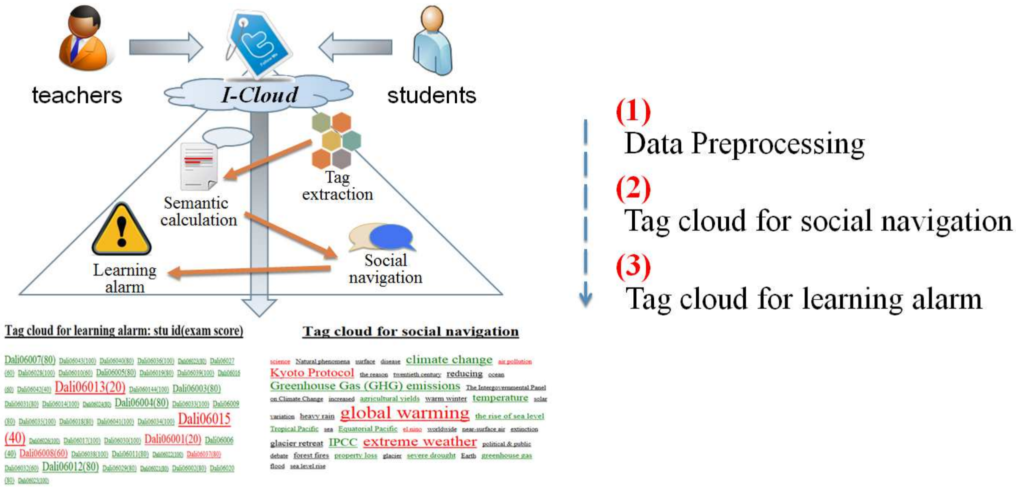


Figure 2: framework for the support approach

2.1 Data Preprocessing

The preprocessing of student tags and articles involves translating tags and reading text into a vector of terms, based on the Vector Space Model (VSM). Each article is preprocessed by 1) tokenizing, 2) removing stop words, 3) and Porter Stemming. In order to identify meaningful paragraphs and locate article topics, we use Latent Semantic Analysis (LSA) to measure the importance of each paragraph and the relationship between paragraphs. The tag is also considered a measure of the importance of a paragraph in the article. Meanwhile, we calculate a tag score to identify the importance of each subsequent paragraph by using the diagonal matrix elements with the LSA process and TF-IDF of the vector-based information retrieval scheme [4]. After this scoring process, we derive the tag score to identify the importance of each subsequent paragraph.

2.2 Tag cloud for social navigation

A tag cloud can be regarded as a collection of main topical terms mined from the articles. It enables the reader to identify the context of the retrieved text data and quickly determine if it is of interest. In this subsection, we describe the tag-cloud generation technique for student tags of well-tagged items. Here, we applied an information retrieval scheme and collaborative filtering (CF) to mine topical terms intelligently, which are then presented in the form of a cloud of tags with varying sizes emphasizing degrees of relevance. This process is summarized below:

(1) To identify the importance of each subsequent paragraph, the tag score ($W_t^{i,j}$) is calculated from the pre-processing process. $W_t^{i,j}$ represents the tag score of the t_{th} tag of i_{th} paragraph in the j_{th} article. It is regarded as the evaluation of the importance of a

paragraph and can help students to construct the article structure. Top K (user-given parameter) recommended tag sets $Top_K(\mathrm{t}) = \sum_{t\in\ Top_k} \sum_{i\in I} Top_K(\mathrm{W}_t^{i,j})$ are considered for recommending topic tags to students, and help students understand and identify key ideas within the reading (K set in our experiments to 5).

(2) Before identifying appropriate supplementary materials for a student, we recommend a tag cloud to each student. The basic idea of our recommendation approach is collaborative filtering (CF), information retrieval and tag usefulness [12]. The underlying assumptions of CF are that people with similar interests will prefer similar tags, and that those individuals who agreed in the pass tend to agree again in the future [13]. As in user-based CF [8], our approach predicts the tag score likelihood of neighborhoods of students to reflect meaningful tags to be recommended to the user. Tag usefulness is determined by powerful coverage of the most frequently used tags. Thus, the recommendation score of tag t for social navigation (SN) for student u is determined by:

$$\mathrm{SN(u,t)} = e^{-\gamma} \times [\alpha \times \sum_{v\in N(u)} \mathrm{w(u,v)} \times \mathrm{w(v,t)} + (1-\alpha) \times W_{Top_K}(\mathrm{t}) \times \sum_{t\in T(u)} \mathrm{w(u,t)}] + \sum_{u=1} \left(\frac{\log(d_{t,j})}{m^2} \right)$$

where γ is a propagation decay factor; α represents a parameter that controls the relative weight between people and tags (initial α is 0.5); w(u,v) and w(u,t) are the relationship strengths of u to student v and tag t, as given by the student profile; w(v,t) represents the relationship strengths between student v and tag t; $W_{Topk}(t)$ represents the tag weight for recommending topic tags; $d_{t,j}$ represents the frequency with which the t_{th} tag has been used to describe the j_{th} article; and lastly m is the number of different tags assigned to the j_{th} article.

(3) For each tag t, the normalized weight $SN_{norm}(u,t)$ is used to calculate its font size. The cloud accomplishes this by increasing and decreasing the font size of each tag according to its weight. Thus, when the student clicks on a given tag in the tag cloud the system selects suitable supplementary materials based on our previous study, the tag-based prior knowledge recommendation system (TAK) [3, 4], and presents them to the student. This function is an engaging way of visualizing meaningful information and helping rethink the reading process for students.

In brief, this stage adopted a social tagging approach, based on a tag cloud navigation schema, to help us examine and understand each aspect of student learning behavior. The social navigation interface is shown in Fig. 3(a). A tag is regarded as a single keyword or a phrase that describes the topic, theme, or idea of the article. Students can add as many tags as their wish, and use the tag cloud visualization tool to discover interesting clues and refine their thoughts or ideas of the reading. Thus, when students click on a given tag in the tag cloud, the system serves as a useful reference guide, as well as selects suitable supplementary materials for students by analyzing the characteristic of the tag cloud.

2.3 Tag cloud for learning alarm

To help teachers evaluate a student's learning situation and monitor the learning achievements of individual students more easily, the learning alarm is combined with the tag-cloud monitoring mechanism into a tag-based prior knowledge recommendation (TAK) system. A tag-cloud monitoring mechanism can be extremely useful to help teachers keep up with all relevant student activity and drill down into their detailed learning portfolio by tag cloud navigation. This process is summarized as below:

(1) Students read an article and take an exam, and the score is used to evaluate their comprehension. We assume the students' scores approximate their level of comprehension, and thus our tag score mechanism is a combination of a student's exam score with their tagging preferences for the same article.

(2) To estimate the error in interpreting an individual's exam score, the standard error of measurement (SEM) is calculated by using a reliability coefficient and the standard deviation of the exam, as follows:

$$\mathrm{SEM} = \mathrm{SD} \times \sqrt{1 - \frac{\mathrm{Q}}{\mathrm{Q}-1}\left(1 - \frac{\mathrm{M}(\mathrm{Q}-\mathrm{M})}{\mathrm{Q}\cdot\sigma^2}\right)}$$

where SD is the standard deviation of the exam, Q is the number of exam items, M is the average exam score of every student who takes the exam, and σ^2 is the variance of every student who takes the exam. The larger the SEM, the less reliable the test. Thus, we use the standard error of measurement (SEM) to recalculate the exam score.

(3) Not all tags carry the same amount of information, and so we use the entropy of a tag to measure its potential characteristics. The value is normalized into a range of 0 and 1.

$$Entropy(t) = -\sum_{u \in U} p_j(u,t) \times \log(p_j(u,t)) \quad ; \quad p_j(u,t) = \frac{TF_j(u,t)}{\sum_{t' \in T} TF_j(u,t')}$$

Here, $P_j(u,t)$ represents the probability of tag t appearing in the j_{th} article for student u, and $TF_j(u,t)$ denotes the number of times tag t appears in the j_{th} article for student u. Tags with higher entropy have lower tag scores. Tags with higher tag scores are important, because these tags help describe the topics or ideas they annotate in a more accurate way. The learning score in the j_{th} article that the u_{th} participant annotates can be represented as follows:

$$\mathrm{LA}_j(\mathrm{u}) = \lambda \times \frac{1}{\mathrm{SEM}_\mathrm{j}} \times \left(\frac{\mathrm{S}_\mathrm{u}^j}{\mathrm{Avg}(\mathrm{S}^j)}\right) + (1-\lambda) \times (1 - Entropy(t))$$

where λ represents a parameter that controls the relative weight between the exam score and a tags' importance, S_u^j is the exam score of the j_{th} article that the u_{th} student receives, and $\mathrm{Avg}(\mathrm{S}^j)$ is the average score of the j_{th} article for all students. Here, large values signify a strong alert status where a teacher should be aware of the early conditions of a learning problem.

(4) A threshold θ is used to determine the selection warning sign in order to discover an abnormal learning status from tagging behaviors and reading comprehension degrees. When a teacher logs into the proposed system, they can determine the threshold to warn of any potential abnormal learning status of individual students. When $\mathrm{LA}(u) < \theta$, a user's name in the tag cloud is shown with a green font color, and the larger font size represents better learning performance. In contrast, a user's name in red font attracts the teacher's attention and the larger font size represents lower learning performance.

To summarize, Fig. 3(b) shows the learning alarm interface presenting the current status of a student's learning performance and tagging behavior, recently forwarded to the alert monitor. When students who have a learning disability are diagnosed by analyzing the tag cloud for social navigation, this tool can assist teachers in providing focused feedback and questions to students as soon as possible. This helps activate prior knowledge, probe students' conceptual understanding, and leads to deeper understanding.

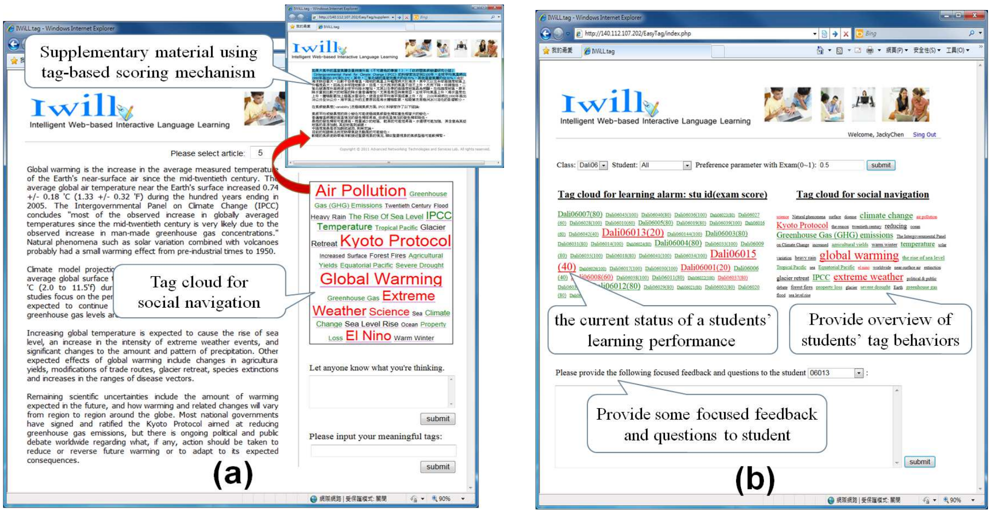


Figure 3: Interface of I-cloud (a) Tag cloud for social navigation (b) Tag cloud for learning alarm

3. Experiments

The goal of our experiment is to quantify the influence of I-cloud practices on an on-line learning environment. In this section, a series of investigations is conducted to evaluate the usefulness of I-cloud.

3.1 Student data and results

In order to evaluate the efficacy of tag cloud for social navigation, an experiment was conducted from April 2012 on reading activity at a senior high school in Taiwan. 86 participants were divided into two groups. A pre-test and post-test experimental design employed before-and-after surveys to demonstrate the usefulness of our approach among participants. Each group was given a pre-test evaluation and then a post-test after two weeks. The experimental students demonstrated significant improvement in reading comprehension ability by taking advantage of the tag-cloud for social navigation (Table 1; $p < .05$). Fig. 4 shows the learning curve of students that is representational of an average rate of knowledge gained over time. Students first start learning how to use an on-line reading system in the first 10 minutes of the reading activity. The control group on a flat learning curve demonstrates that the rate of knowledge gained is slowly spaced out over time. That implied that subjects with a flat curve are often very difficult to learn, so that students can't quickly grasp the concepts and use the knowledge. In the contrast, the experimental group gains knowledge quickly after 10 minutes and I-cloud helps them quickly grasp the concepts and knowledge of reading materials. In addition, students' reflections on the activity were collected using a questionnaire and personal interviews. The results of 43 questionnaires were conducted after post-test. Each question underwent a discriminate validity test by using factor analysis. The coefficients from the experimental results show that these factors were sufficiently reliable for representing student-tagging behaviors. The major findings are presented as follows:

(1) 74% of students agreed that the I-cloud interface was easy to use.
(2) 88% of students agreed that tag cloud offers a way to navigate through a structure of article based on tags, and it can help students quickly grasp the structure and concepts of English articles. Some students indicated that these tags help them easily realize new information from the article.
(3) 79% of students thought that tag cloud navigation reduced the amount of search time and enhanced their search experience.
(4) 93% of students agreed that tag cloud can enable students to select different views on the tag cloud, such as system tags or other student tags, and then help judge difference viewpoints and explicate their thinking processes.
(5) 86% of students agreed that the tag cloud provides navigational clues through corresponding supplementary materials, which help students understand prior knowledge in the article.

Table 1: Paired t-test of the pre-test and post-test results

Test	Group	N	Mean	Std. Deviation	Std. Error Mean	t-test
Pretest	Control Group	43	57.79	28.0794	4.2821	t = .588
	Experimental Group	43	54.65	25.5530	3.7443	p = .56
Posttest	Control Group	43	60.70	14.5807	2.2235	t = -2.487
	Experimental Group	43	66.86	11.8534	1.8076	**p = .017***

*p < .05

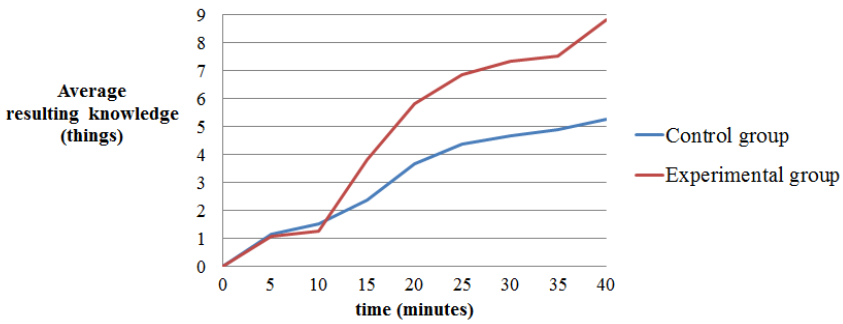


Figure 4: Learning curve simplified

3.2 Teacher data and Results

To evaluate the effectiveness of the tag cloud for learning alarm, five participating teachers from a senior high school in Taiwan were invited to experience the use of I-cloud. We designed a questionnaire that was used to collect feedback from the teachers. The major findings are presented as follows:
(1) All teachers agreed that the interface of I-cloud is helpful for teachers to realize the learning status of students. This interface can assist teachers in evaluating the reading ability of the students, such that constructive suggestions can be given to the students, and tutoring strategies can be improved accordingly. Moreover, I-cloud provides a new way to assist teachers that exploit tag information on the students' tag cloud to get their attention and bring them back into focus.
(2) 80% of the teachers agreed that the interface is intuitive and easy to use. Only one teacher indicated that the interface needed more guidance.
(3) 80% of the teachers believed that tag cloud for social navigation can enhance the ability of the students for prior knowledge learning.

4. Conclusions and Future Work

In this work, we propose a novel method for monitoring and creating significant learning experiences. The results of this study showed that experiences with applying social tagging and tag cloud helped students increase their reading comprehension and quickly grasp the structure and concepts of English articles. In addition, the experimental results demonstrated that this novel approach is also helpful in assisting teachers in evaluating student learning achievements by tag-cloud visualization. These results provide suggestions and references for the design of efficient web 2.0 supported collaborative learning activities in the future. Furthermore, we hope to examine the differences between the effectiveness of a more declarative approach, and make it easier to use and combine different tools and data sources for generating learning suggestions. These suggestions should further ensure that social tagging applications in collaborative learning environments improve reading and recommendation incrementally in future experiments.

Acknowledgements

This work was supported by the National Science Council of Taiwan, R.O.C., under contract numbers NSC100-2221-E-001-015-MY3.

References

[1] Chen, J. M., Chen, M. C., & Sun, Y. S. (2010). A novel approach for assisting teachers in assessment of student reading ability in web-based learning environment. International Conference on Computers in Education (ICCE), 260-264.

[2] Chen, J. M., Chen, M. C., & Sun, Y. S. (2010). A novel approach for enhancing student reading comprehension and assisting teacher assessment of literacy. Computers & Education, 55(3), 1637-1382.

[3] Chen, J. M., Chen, M. C., & Sun, Y. S. (2012). A hybrid tag-based recommendation mechanism to support prior knowledge construction. the 12th IEEE International Conference on Advanced Learning Technologies (ICALT).

[4] Chen, J. M., Chen, M. C., Sun, Y. S. & Chen, Y. Y. (2011). A novel approach for enhancing student reading comprehension by activating prior knowledge. International Conference on Computers in Education (ICCE), 279-286.

[5] Chen, L. H. (2011). Enhancement of student learning performance using personalized diagnosis and remedial learning system. Computers & Education, 56(1), 289-299.

[6] Corredor, J. (2006). General and domain-specific influence of prior knowledge on setting of goals and content use in museum websites. Computers & Education, 47(2), 207-221.

[7] Guan, Z., Wang, C., Bu, J., Chen, C., Yang, K., Cai, D., & He, X. (2010). Document Recommendation in Social Tagging Services. In Proceedings of the 19th international conference on World Wide Web, 391-400.

[8] Guy, I., Zwerdling, N., Ronen, I. Carmel, D., Uziel, E. (2010). Social media recommendation based on people and tags. SIGIR 2010: 194-201.

[9] Lin, Y. C., Lin, Y. T., & Huang, Y. M. (2011). Development of a diagnostic system using a testing-based approach for strengthening student prior knowledge. Computers & Education, 57(2), 1557-1570.

[10] Wang, T. H., Wang, K. H., Huang, S. C., & Chen, S. Y. (2004). Web-based Assessment and Test Analyses (WATA) system: Development and evaluation. Journal of Computer Assisted Learning, 20(1), 59–71.

[11] Sinclair, J. & Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? J. Inf. Sci., 34(1), 15-29.

[12] Yusef, H. M., & Víctor, H. S. (2006). Improving Tag-Clouds as Visual Information Retrieval Interfaces, International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006, 25-28.

[13] Zhen, Y., Li, W., & Yeung, D. (2009). TagiCoFi: Tag Informed Collaborative Filtering. In Proceedings of the 3rd ACM Conference on Recommender Systems (RecSys'09), 69-76.

An Interpretable Statistical Ability Estimation in Web-based Learning Environment

Yi-Ting Huang[a*], Meng Chang Chen[b], Yeali S. Sun[a]
[a] *Department of Information Management, National Taiwan University, Taiwan.*
[b] *Institute of Information Science, Academia Sinica, Taiwan.*
*d97008@im.ntu.edu.tw

Abstract: With growing interest in estimating true ability in contemporary learning, the demand for personalized learning and Web-based learning environments has become increasingly important. This paper develops a statistical and interpretable method of estimating ability. This method captures the succession of learning over time and provides an explainable interpretation of a statistical measurement, based on Item Response Theory and the quantiles of acquisition distributions. The results from the simulation and empirical study demonstrate that the estimated abilities can successfully recognize the actual abilities of students. The correlation values between the estimated abilities and the post-test score, which incorporate this testing history, are higher than values that only consider test responses at the time of testing. Furthermore, the pre-test and post-test administered to the experimental group show significant student improvement. These results suggest that this method serves as a successful alternative ability estimation and provides a better understanding of student competence.

Keywords: Ability estimation, Adaptive test, Item response theory

1. Introduction

Recently, theories on learning have focused increasing attention on understanding and measurement student ability. Vygotsky [12] states that a learner's ability in the Zone of Proximal Development (ZPD)—the difference between a learner's actual ability and his or her potential development—can progress well with external help. Instructional scaffolding [11], closely related to the concept of ZPD, suggests that an appropriate support during the learning process helps learners achieve their stated goals. Effective instructional support requires identifying a student's prior knowledge, tailoring an aid to meet their initial needs, and then removing this aid when he or she acquires sufficient knowledge.

Nowadays, estimations of ability offer extensive applications within e-learning systems in various domains. For example, Chen et al. [4] considered a learner's ability for recommending personalized learning paths in a Web-based programming learning system, while Chen and Chung [3] analyzed students' understanding by suggesting English vocabulary on mobile devices. Similarly, within Computerized Adaptive Testing (CAT), Barla et al. [1] calculated an examinee's ability to select suitable questions. All of these studies used Item Response Theory (IRT) to estimate a student's ability, and their results demonstrated improved student performance.

Item Response Theory is a modern theory of testing that examines the relationship between an examinee's responses and items related to abilities measured by the items in the test. Three well-known ability estimations proposed by IRT are maximum likelihood estimation, maximum a posteriori and expected a posteriori [6]. Examples of this research include [1], where researchers used expected a posteriori to score each examinee's ability at each time of a test. However, IRT has some disadvantages. First, every exercise performed by a student is recorded in most of the web-based learning environments listed above; however, the ability estimations of IRT only consider test responses at the time of testing, rather than incorporating testing history. Moreover, the interpretation of the result of estimating an examinee's ability is often defined in terms of the acquisition of a large portion of knowledge of the specific ability itself—through a test. Unfortunately, this definition is qualitative rather than quantitative.

In response to these issues, this paper proposes a statistical method and a novel interpretation of estimating ability with inherent randomness in the acquisition process. We conduct a simulation study to investigate the property of the proposed approach and an empirical study to evaluate practical performance. Our simulation results demonstrate the convergence between an examinee's current grade and his or her actual ability. We also implement this method on a Web-based learning environment. The empirical results find a strong correlation between the estimated ability and the post-test score that incorporates this testing record, and this correlation is higher than correlations between ability and values that only examine test responses at the time of testing. Moreover, the pre-test and post-test administered to the experimental group demonstrate significant student improvement.

The remainder of this paper is organized as follows. In Section 2, we present the proposed ability estimation. Section 3 reports a simulation and Section 4 contains the empirical procedure and results. Finally, Section 5 summarizes our conclusions.

2. Method

We propose the following interpretation of the quantitative definition: an examinee is said to have ability θ if s percent of items in a test $T = (t_1, \ldots, t_m)$ have been correctly answered each by r percent of the population.

We first consider that each item t_i in a test T has been correctly answered by r percent of the population. In general, there is a specific knowledge behind each tested item t_i. The level of the specific knowledge represents that most people have acquired knowledge of t_i. Most people understand some knowledge at an early age, whereas some understand this knowledge later in life. Here, we precisely denote the level the specific knowledge represents as the age at which r percent of the population has acquired knowledge of t_i, where age can refer to school grades or lifetime. When given a knowledge t_i and a population, the probability distribution of knowledge acquisition $p_t(\theta)$ can be calculated. Let the quantile function q_t of the cumulative distribution function correspond to the acquisition distribution p_t. In other words, $q_t(r)$ represents the age at which r percent of the population has acquired knowledge of t. This assumes a normal distribution,

$$q_t(r) = \mu_t + \Phi^{-1}(r)\sigma_t \quad (1)$$

where μ_t and σ_t represent the mean and standard deviation of the distribution p_t, and $\Phi^{-1}(r)$ is a quantile function representing the probability of exactly r to fall inside the interval of the distribution. When an examinee correctly responds to the item t_i, the examinee's ability is regarded as the age or grade level, etc. To investigate the distribution of the grade level of a test T, we collect the grade level values generated from each quantile function $q_t(r)$ as the distribution of knowledge acquisition within a single test f_Q.

In practice, this is time consuming and costly for each item t_i known in advance by the distribution p_t. Fortunately, under Item Response Theory [6], a response of an examinee to an item is modeled by a mathematical item response function, known as the item characteristic curve. The item characteristic curve is a mathematical family model that describes the probability of a correct response between an examinee's ability and the item parameters. These models employ one or more parameters, such as an item difficulty parameter and an item discrimination parameter, to define a particular cumulative form. When given the item parameters, the grade level at which *r* percent of the population correctly responds to item *t* can be inferred. Take one-parameter logistic model as an example,

$$q_t(r) = \ln(\frac{r}{1-r}) + b \qquad (2)$$

where variable *b* as item difficulty.

Estimating an examinee's ability through a test relies on the test responses of the test. We consider a percentage of correct responses in a test as variable *s* and define the *s*th quantile of the distribution of knowledge acquisition in a test f_Q as the examinee's ability. The distribution of the *s*th quantile of f_Q, where *s* percent of items in a test have been correctly answered by *r* percent of the population, can be performed using a standard formula for normal approximation of order statistics [5]:

$$q_T(r,s) \sim N(F_Q^{-1}(s), \frac{s(1-s)}{m[f_Q(F_Q^{-1}(s))]^2}) \qquad (3)$$

where F_Q is the cumulative distribution function and *m* is the number of items in a test. This result is more certain of the estimated grade level assigned to a large sample item size. In cases where an examinee correctly answered all items or no item, a smooth constant *c* is used (*c=0.01* in this study).

When given an examinee's responses in a test, the current examinee's ability θ_t can be described by the distribution (3) in which *r* percent of the population correctly answer *s* percent of items. We also consider an examinee's history record, and employ Exponential Moving Average (EMA) [2] to combine this history with the current ability, transformed by the following formula:

$$ability_t = \alpha \times \theta_t + (1-\alpha) \times ability_{t-1} \ (4)$$

where θ_t is the current ability in time *t* obtained from the mean of the equation (3), $ability_{t-1}$ is the past estimated ability in the time *t-1* as history records, and $ability_t$ is the final estimated ability in time *t* after the combination of the current ability and the past estimated ability with EMA. Additionally, $\alpha = 2/(n+1)$ is a smoothing constant represented as an exponential weight, and *n* represents the period as the length of the moving window.

3. Simulation

3.1 Settings

To understand the performance of the proposed method, we conducted a simulation. According to a one-parameter logistic model in Item Response Theory [6], the probability of correct response is *0.5* when an item difficulty is equal to an examinee's ability. In the simulation, we referred to this probability for setting the variable *r*. Moreover, the item response model also provides information in the estimation of the variable *s*. We used a one-parameter logistic model to predict the probability of a correct response when given the ability and an item, and then conditionally randomly sampled the variable *s*.

In each simulation, ten items were generated according to an examinee's ability at the time. The distribution of difficulty of these items acts as a normal distribution. For example, given an examinee's ability $\theta=3$, the difficulties of a test are {2, 2, 3, 3, 3, 3, 3, 3, 4, 4}. Ability and difficulty in this study range from one to six, corresponding to the school grades. In practice, an examinee's school grade is considered as their initial ability, and the ability is updated by responses in each test. Thus, the simulation starts with any grade ranging from one to six in order to simulate different grade students with various abilities, and then terminates 100 iterations after the convergence point. We found the convergence point and then counted the Root Mean Square Error (RMSE) during the 100 iterations. The definition of the convergence point is determined by computing the difference between the estimated ability and the ground truth, and the difference value is continuously four times smaller than a threshold (*thd = 0.25* in the simulation). Each simulation was processed 1000 times. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{k}\sum_i (\theta_i - \hat{\theta}_i)^2} \qquad (5)$$

where θ is the actual ability as the ground truth, $\hat{\theta}$ is the estimated ability, k is the number of the iterations. Here, k=1000. This metric represents the average distance between the ground truth and the generated results. The smaller RMSE value indicates that the estimated ability is close to the ground truth. In addition, we also discuss the parameter α in equation (4). The parameter is presented in terms of n time periods and represents the weight of the observation at the present time. The variable n was set from one to twelve.

3.2 Results

Table 1 shows the average convergence points in the number of variable n of parameter α in equation (4) over the degree of difference between the estimated ability and ground truth, and the results of RMSE during the 100 iterations after the convergence points. It is clear that the proposed method can successfully estimate abilities in the finite iterations. Specifically, an examinee's ability can be estimated more precisely when he or she continues to have more tests. Furthermore, the error distances between the estimated abilities and the ground truths are low enough to be acceptable after convergence. That is, an examinee's ability can be steadily measured during a long-term observation.

The parameter $\alpha = 2/(n+1)$ in the equation (4) is an exponential weight of the current ability, and n represents the number of time periods, such as times or days, taken into consideration. When $n=1$, it represents that an examinee's ability only considers the current estimated ability without the history record. In Table 1, the values in screentone present that the average convergence points are fewer than the points generated from $n=1$. This result shows that the estimated abilities are quickly found and the error distances decrease when considering the history record. In particular, it is apparent when the initial grade is equal to the ground truth. When n is small (e.g. $n=2$, $\alpha = 2/3$; $n=3$, $\alpha = 1/2$), the estimated ability is mainly decided by the current ability. The convergence points are smallest and the RMSE is slightly smaller than one generated from $n=1$. In contrast, when n increases, the estimated ability is principally composed of abilities from the past to now. If an examinee's initial ability is not close to his or her actual ability, it takes more information to accurately estimate. Although it takes time, the RMSE is clearly shrinking.

Table 1. The results of convergence point and RMSE (each row represents the degree of difference between the initial ability and the actual ability, and each column represents the number of time periods considered by the exponential weight of the current ability)

d \ n	1	2	3	4	5	6	7	8	9	10	11	12
0	20.61	13.88	11.72	11.53	10.98	10.90	10.26	10.52	10.16	10.35	10.18	10.04
1	21.96	16.17	15.74	16.31	17.40	19.07	20.43	22.29	23.98	25.45	26.92	28.42
2	22.91	18.08	18.54	19.91	21.90	24.18	26.64	29.06	31.50	33.53	35.62	38.58
3	23.86	19.67	19.91	21.91	24.59	27.62	30.33	32.90	35.74	38.43	41.52	44.13
4	24.30	20.73	21.52	23.51	26.71	29.68	32.96	36.00	40.19	42.83	45.45	48.65
5	24.50	21.41	22.66	25.22	29.10	31.92	35.97	38.22	42.62	46.40	49.18	53.12
RMSE	0.39	0.32	0.28	0.26	0.24	0.23	0.22	0.21	0.20	0.19	0.19	0.18

Consider a dramatic example to explain the properties of the proposed method. Assume that a first grade student, whose real ability is the sixth grade, learns and has a test in a web-based learning system once a day. Figure 1 illustrates the changes in the estimated ability computed from the proposed method in different weights. The black horizontal line at the sixth grade represents the student's actual ability as the ground truth. The other curves depict the estimated abilities under the different weights: a red dotted line, *n=1*; a green solid line, *n=3*; a purple solid line, *n=6*; and a blue solid line, *n=12*. The mark labels on each line are the convergence points (the value is continuously four times smaller than *thd* = *0.25*). It is clear that the estimated abilities are converging as *n* decreases in size. Although these estimated abilities are estimated using few iterations when *n=1*, the red-dotted line drastically fluctuates after the convergence point. In other words, if the ability estimation only takes the current responses into consideration, instead of past performance, the variance of every estimated ability may be large. In this situation, question selection in a test using inaccurate ability estimation could result in confusion by the examinee. In contrast, the estimated error gradually decreases when *n>1*, even though the estimated abilities when *n=1* take more time to estimate. In this situation, the students' abilities were gradually updated and the difficulties of items incrementally increased. This is thus a trade-off problem between time and precision.

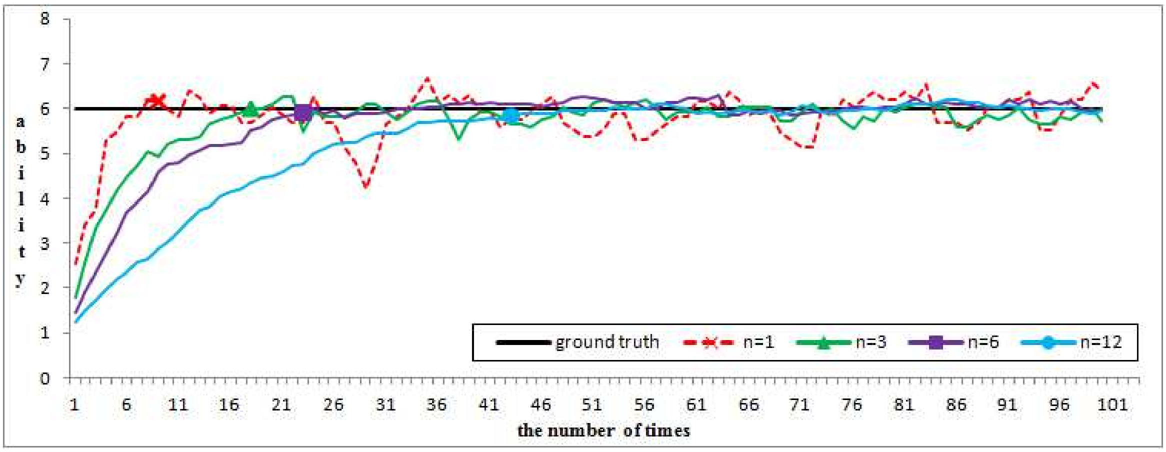


Figure 1. The changes in the estimated ability computed from the proposed method for the different weights (*n=1*, *n=3*, *n=6*, *n=12*)

4. Empirical Study

4.1 Materials

The measurement approach proposed in this study is implemented on a Web-based learning system developed by the AutoQuiz Project of the IWiLL learning platform [8]. It provides English language learners online English reading materials collected from up-to-date online news websites and multiple-choice tests and automatically generates related quiz material [9][10]. Each test was composed of ten vocabulary questions, five grammar questions, and three reading comprehension questions. A total of 2,425 items were automatically generated based on 72 reading materials. The grade level of the vocabulary and grammar questions are defined according to the semester of high school in which the correct answer is taught, while the difficulty of the reading comprehension questions are measured by a reading difficulty estimation [7]. In other words, the grade level in this experiment is defined from one to six, corresponding to the six semesters of senior high school.

4.2 Participants and Procedure

The participants in this study were high school students in Taiwan, divided into two groups: a control group where ability is estimated only based on current responses, and an experimental group that incorporates the history record into the current ability estimation. 30 students participated within the control group, while 47 students participated in the experimental group.

The experiment was held from January 30th to March 4th, 2012. During the experiment, the subjects were asked to participate in twelve activities, consisting of reading an article and then taking a test. In each activity, the subjects in both groups received an up-to-date article and a series of quizzes automatically generated based on their abilities. In addition, there was a pre-test and post-test for evaluating their abilities as the ground truth. The variable r was set as 0.5 based on IRT, and the variable s defined as the percentage of correctly answered items. Furthermore, the parameter $n=12$ in the exponential weight of the experimental group was equal to the period of activity, because all test records were taken into consideration.

4.3 Results

To validate the accuracy of the proposed ability estimation, the subjects' abilities in the two groups were estimated with twelve continuous activities. Table 2 reports the Pearson's correlation coefficient between the estimated abilities (the estimated grade is rounded by the estimated score) and the post-test scores among the three quiz types. All of the measures are significantly positively correlated. The results in the experimental group ranged from *0.44* to *0.69*, while ones in the control group ranged from *0.47* to *0.54*. Most of the correlation values in the experimental group are higher than the values in the control group; this suggests that estimating ability with the history record leads to a clearer relationship between the estimated ability and the ground truth.

Table 2. The correlation result between the estimated ability and the post-test in the control group and the experimental group

	vocabulary		**grammar**		**reading comprehension**	
	score	grade	score	grade	score	grade
Control group	0.47*	0.49**	0.54**	0.51**	0.54**	0.47*
Experimental group	0.51***	0.44**	0.55***	0.55***	0.69***	0.65***

$^*p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

Comparing the post-test score in each estimated ability (grade) is another way to assess the accuracy of the proposed ability estimation. If the estimated abilities are accurate, the subject performance of each ability will differ from that of other abilities. Table 3 presents the mean post-test score of the subjects of different estimated abilities between the control group and the experimental group. Intuitively, a subject estimated a higher ability should has higher post-test score than one estimated a lower ability. One-way Analysis of Variance revealed that there were differences in the estimated vocabulary ability ($F=5.75$, $p=0.001$), the estimated grammar ability ($F=4.71$, $p=0.003$) and the estimated reading comprehension ability ($F=5.98$, $p<0.001$) in the experimental group, while there were no statistical differences between the estimated vocabulary and grammar ability in the control group. Noticeably, although the estimated reading comprehension ability in the control group has a significant difference, the mean scores among every ability fluctuated. The bolded values in Table 3 are unreasonable, because the averaged scores of the higher estimated abilities (grade 2, grade 4 and grade 5) in the control group were lower than ones of the lower estimated abilities (grade 1 and grade 3). Though there was an unreasonable value for grade 6 of the estimated vocabulary ability in the experimental group, this is likely because only two students were assigned to grade 6. This sample size is likely unrepresentative. Moreover, in the experimental group, a Bonferroni post hoc test indicated that the performance of the estimated ability 1 and 2 were significantly different from the estimated ability 5 and 6. This indicates that the proposed ability estimation can effectively distinguish higher ability examinees from lower ones.

Table 3. The mean post-test score of the subjects in different estimated ability groups between both groups and the result of ANOVA

Estimated ability	**Control group**			**Experimental group**		
	vocabulary	grammar	reading	vocabulary	grammar	reading
1	-	37.50	46.80	-	-	37.67
2	48.33	47.00	**40.00**	23.00	34.33	46.63
3	**38.00**	51.40	52.57	52.86	52.80	53.50
4	54.40	41.40	**41.00**	62.33	54.94	64.50
5	61.22	62.83	**32.67**	69.71	66.81	66.90
6	65.83	65.56	70.18	**57.67**	72.00	78.00
F score	2.67	2.54	**6.12******	5.75***	4.71**	5.98***

**$p<0.01$, ***$p<0.001$

To further understand the impact of employing the proposed ability estimation on learners, we investigated the performance between the control group and the experimental group. In keeping with the previous results, the estimated subjects' abilities in the experimental group were more accurate than those in the control group. We assume that appropriate instructional scaffolding could help students advance their learning, when effectively identifying their abilities. Table 4 presents the descriptive statistic and results of a T-test between the pretest and post-test. The results of the independent T-test ($p=0.92$ in the pre-test and $p=0.51$ in the post-test) showed no significant effect on the post-test between the experimental group and the control group. It is noticeable that the average score of the experimental group in the pretest was lower than the control group, but that of the experimental group in the post-test made great progress and surpassed the control group. Additionally, the paired sample T-test showed a significant effect of the pre-test and the post-test in the experimental group ($p<0.001$), while the performance of the control group had no statistically significant effect ($p>0.05$). This indicates that the subjects in the experimental group with an appropriate support can exceed the past themselves when successfully recognizing their learning status.

Table 4. The results of the pretest and post-test between the control group and the experimental group

	Pretest		**Post-test**		**Paired sample t-test**
	mean	std.	mean	std.	
Control group	53.23	19.35	56.70	17.99	1.57
Experimental group	52.83	16.67	59.28	16.01	3.71***
independent t-test	0.20		0.66		

***$p<0.001$

5. Conclusion

This work develops a statistical and interpretable method of estimated ability that captures the succession of learning over time in a Web-based test environment. Moreover, it provides an explainable interpretation of the statistical measurement based on Item Response Theory and the quantiles of acquisition distributions. The result from the simulation demonstrated that the estimated abilities obtained from the proposed method could successfully approximate the actual abilities of students, and estimated abilities can be steadily measured during long-term observation. This proposed approach was also implemented on a Web-based learning environment. The empirical results show that the correlation values incorporating this testing history were higher than the values that only consider test responses at the time of testing. Additionally, the pretest and post-test administered to the experimental group demonstrated significant student improvement. This paper presents preliminary results of a pilot experiment; future research will be further expanded to include long-term evaluation of the effectiveness of the proposed approach under changes in student learning.

Acknowledgment

This work was partially supported by National Science Council, Taiwan, with Grant No. 100-2511-S-008-005-MY3 and NSC97-2221-E-001-014-MY3.

References

[1] Barla, M., Bielikova, M., Ezzeddinne, A. B., Kramar, T., Simko, M. and Vozar, O. (2010). On the impact of adaptive test question selection for learning efficiency. *Computer & Education*, 55(2), 846-857.

[2] Brown, R. G. (2004). *Smoothing, Forecasting and Prediction of Discrete Time Series*. Dover Publications.

[3] Chen, C.M. and 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.

[4] Chen, C.M. Lee, H.M. and Chen, Y.H. (2005). Personalized E-learning System Using Item Response Theory. *Computers & Education*, 44 (3), 237–255.

[5] David, H. A. and Nagaraja, H. N. (2003), *Order Statistics*. Marblehead, MA: Wiley.

[6] Embertson, S., & Resise, S. (2000). *Item response theory for psychologists*. New Jersey, USA: Lawrence Erlbaum.

[7] Huang, Y. T., Chang, H. S., Sun Y. and Chen M. C. (2011). A Robust Estimation Scheme of Reading Difficulty for Second Language Learners. *In Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies*, 58-62.

[8] Kuo, C. H., Wible, D., Chen, M. C., Sung, L. C., Tsao, N. L. and Chio, C. L. (2002). Design and Implementation of an Intelligent Web-based Interactive Language Learning System. *Journal of Educational Computing Research*, 27(3), 785 – 788.

[9] Lin, Y. C., Sung, L. C. and Chen, M. C. (2007). An Automatic Multiple-Choice Question Generation Scheme for English Adjective Understanding. *In Proceedings of the 15th International Conference on Computers in Education*, 137-142.

[10] Lin, Y. T., Chen, M. C. Sun, Y. S. (2009). Automatic Text-Coherence Question Generation Based on Coreference Resolution. *In Proceedings of the 17th International Conference on Computers in Education*.

[11] Wood, D., Bruner, J.S., and Ross, G. (1976). The Role of Tutoring and Problem Solving. *Journal of Child Psychology and Psychiatry*, 17, 89-100.

[12] Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Process*. Cambridge, MA.: Harvard University Press.

Deficiencies of scope statements in ITLET standardization

Tore HOEL[a] & Jon MASON[b]
[a] *Oslo and Akershus University College of Applied Sciences, Norway*
tore.hoel@hioa.no
[b] *Faculty of Education, Queensland University of Technology, Australia*
jon@intercog.net

Abstract: To execute good design one not only needs to know *what* to do and *how* to do it, but also *why* it should be done. For a standardization expert the rationale of a standardization project may be found in the proposal for a new work item or terms of reference, but rarely in the scope statement. However, it is also commonplace that the rationale of the project is not clearly stated in any of these parts. If the rationale is not surfaced in the early phases of a project, it is left to the design, sense-making and negotiation cycles of the design process to orient the project towards a goal. This paper explores how scope statements are used to position standardization projects in the IT for Learning, Education and Training (ITLET) domain, and how scope and rationale are understood in recent projects in European and international standardization. Based on two case-studies the paper suggests some actions for further research and improvement of the process.

Keywords: rationale, LET standardization, scope, purpose and justification of standards, process improvement

1. Introduction

Development of formal standards within international standard setting bodies (SSBs), like the International Organization for Standardization (ISO), progress through well-defined stages [1]. Any specification work, if required, takes place at the Preparatory and Committee stages, based on premises defined in the Proposal stage. In theory, when technical work starts, experts engaging in the process should understand why the project has been initiated, what can be realistically achieved, and where or how they should focus. In practice, especially with anticipatory standardization, participatory observations reveal that this is not always a straight-forward process and the rationale, scope and methodology of the project may be less than clear.

This paper focuses on the first stages of formal standardization and asks whether the process is sufficiently optimised to give participating experts a clear direction for technical development. To establish context, this paper looks at the Directives of ISO to get an overview of the formal requirements for establishing New Work Items (NWIs). Following this, more detailed research questions are developed.

1.1 The proposal process – justifying new work items

To present a NWI for ballot by national standards bodies the proposal needs to contain basic information structured by a template: a title, scope, a description of purpose and justification, program of work, and information on resources to be provided, etc. The section on purpose and justification should, according to the ISO Directives, explain *why* the standard is needed [1]. However, in the end, the acceptance of a new project is not measured by well-stated justifications. In practice, it is the weighting of votes cast by national bodies that is critical. Ideally, the scope statement is not supposed to change during the development from New Project to an International Standard.

How much of the justification, or initial rationale, is carried through to the later stages of the standardization process? Drafts of the standard will adhere to another template, laying out sections on Introduction, Scope, Normative References, Terms and Definitions, etc. There is no explicit placeholder for 'Purpose and Justification' in the template used by formal SSBs like ISO or CEN. Of course, the Introduction could give information about "reasons prompting its preparation", but the Directives warn, "it shall not contain requirements" [2].

Technical experts joining the project at the Committee stage will use the scope as their primary reference point in drawing up directions for their work. As formal standards development is a very document-driven process, the tug of war over wording is not about purpose of the work, but about scope and technical specifications. Thus, the question arises as to whether the scope allows any reasoning about why the standard is developed.

1.2 Defining scope of international standards

Statements of scope used by ISO serve a number of functions, including as a summary for bibliographic purposes. Therefore, it "should be succinct" [2], and it "shall precisely define the limits of the field of activity" [1]. Activity in this context can be defined from both an internal and an external point of view. The perspective of the ISO Directives [1, 2] is the latter. The scope is a statement precisely defining the limits of the work of a technical committee; "it assists those with queries and proposals relating to a field of work to locate the appropriate committee" [1].

Participants in standards projects may testify that scope statements also have internal functions, e.g., in focusing the work. However, the rules on how to draft the scope may constrain its usefulness in this respect. The Directives give clear rules stating that the scope should be "worded as a series of statements of fact" [2]. Scopes "shall not repeat general aims and principles governing the work of the organization but shall indicate the specific area concerned"; and "It shall not contain requirements" [2].

Scope statements are about setting constraints, boundaries and limits. As such, they can "help guard against moving outside the field of activities authorized by the parent committee" [1] – or, as commonly referred to as, "scope creep". If the boundaries are not defined clearly by scope, exclusions can be added. However, the Directives also warn against "self-evident exclusions" [1].

It is therefore evident from the ISO Directives that the question *why* a standard is to be developed plays a certain role in the Proposal stage of a project, but a more limited role in the Preparatory and Committee stages, where the *what* question is foregrounded. The question *how* a standard should be developed is also of great interest to project participants. As this question is very much dependent on context, the Directives give no guidelines. On the contrary, they explicitly warn against specifying design methods, referring to the Performance approach in the Rules for the structure and drafting of International Standards, to "provide for maximum freedom for further innovation" [2].

Formal standardization, at least as described in the guideline documents analysed above, seems to leave issues of purpose and justification behind when the actual technical work begins. This raises the following research questions:

- What role do conversations about rationale play in requirement elicitation and technical specification work of ITLET standard groups?

- How are the *Why*, *What* and *How* dimensions addressed in scope statements of ongoing and published projects of international ITLET standardization?
- Can an analysis of practices related to scope statements in ITLET give any clues to shortcomings in the processes or any input into practice improvements?

The first question is addressed through a literature review within the field of requirement practice and standards governance. The second question is addressed through two case-studies drawn from CEN Technical Committee 353 *Information and Communication Technologies for Learning Education and Training* (TC353) and ISO/IEC JTC 1/SC36 *Information Technology for Learning, Education, and Training* (SC36). Based on an analysis of these data, proposals are made for further research and consideration in developing new directions for development of standards in the field of ITLET.

2. Related work

It is common practice that the first phase of a project specifies clear goals and objectives, as these are identified as critical success factors (CSF), e.g., in implementing enterprise IT projects. Clear goals and objectives form a clear-cut CSF, but can also be rather problematic [3]. Goals are difficult to determine without consideration of the broader context of IT infrastructure. Akkermans and van Helden suggested that on the methodological level this viewpoint is consistent with the concept of IS development that considers “evolutionary complexity” [3].

The success or failure of projects has been linked to the quality and usefulness of the models representing the requirements [4, 5]. Elicitation of goals and objectives from such a complex domain as ITLET, especially in the field of anticipatory standardization, may require methods and skills that often go beyond the ‘toolbox’ found in many working groups in standardization [6].

Goals and objectives are not the only justification for enabling a project proposal to be accepted. Ideally, the objectives should inform the definition of what is to be specified in the standard. Therefore, the first step is to draft an informal specification of the Universe of Discourse (UoD), in order to inform the formal specification through a modelling and validation process [7]. Two types of experts, whom may be termed domain experts and system analysts, are needed for this work. “Roughly speaking, a domain expert can be characterized as someone with (1) superior detail-knowledge of the UoD but often (2) minor powers of abstraction from that same UoD. The characterisation of a system analyst is the direct opposite” [7]. Where the different areas of expertise meet, natural language is the basis for communication. The domain expert does not need to have any knowledge of formal modelling languages. However, the system analyst should have some abilities to communicate with the “owners” of the problem at hand. Following Frederiks and van der Weide, “the quality of the modeling process is bounded by the quality of concretizing into an informal description augmented with the quality of abstracting from this description” [7]. Of course, the better tool support (language, models, technologies, etc.) these concretisation and abstraction processes have, the better quality of the resulting requirements documents.

Conversations among domain experts play a critical role in the Committee phase of the process. The international and intercultural settings of such conversations make it even more important to identify the root constructs used to build coherency in a published standard. Sense-making is therefore an important process of standardization, together with design and negotiation [8, 9]. Mason [10] presents a model in which *why*-questioning is viewed as an important instigator of sense-making, particularly for knowledge creation. In this model, a set of primitive questions: *Who, What, When, Where, How,* and *If* work together [10]. A consequence is that without *why*-questions, and their answers explicit, understanding of function is necessarily constrained.

Schoechle has found the 'dramatistic pentad' of Burke useful in understanding standardization discourse, "particularly on the lower technical committee or working group level" [11]. Burke identified five key elements necessary to describe human drama, Act, Scene, Agent, Agency and Purpose [12]. These concepts correspond to the questions of what, where/when, who, how and why. For Burke the concepts in the pentad are strongly interrelated, and:

> "Men seek for vocabularies that will be faithful reflections of reality. To this end, they must develop vocabularies that are selections of reality. And any selection of reality must … function as deflection of reality. Insofar as the vocabulary meets the need of reflection, we can say that it has necessary scope." [15]

According to the ISO Directives, scope is only concerned with *what* questions. While this might be the outcome of a well-formed scope statement, however, in order to specify what to standardize, the development of the scope goes through a number of *why*-questioning cycles. The *how* question is informed by the scope, but does not contribute to the specification of the scope.

In addition to literature review this study is based on content analysis of documents in the TC353 and SC36 registries, analysed using a computer assisted qualitative data analysis tool (TAMSanalyzer) for tagging and analysing data in iterative steps.

3. Case studies of rationale and scope in two formal SSBs

3.1 CEN TC353

CEN TC353 was established 2007 and has since produced four European Norms and filed 129 documents in their registry (May 2012). An analysis of all TC353 documents shows that justification of the committee itself is a dominant theme that impacts the standards it produces. The Business Plan's scope is worded as a marketing statement: Developed standards "will have a well-defined European scope", and they are needed because the market is "sufficiently mature to require European specific standards". However, the first published TC353 standard was a rubber stamped ISO standard on quality, adding no European specific technical work, or even no European specific rationale.

CEN uses the same 'new project template' as ISO, but while TC353 may have utilised the template, it has also interpreted it. In a proposal for a new Curriculum Exchange Format the Scope section is missing. TC353 works closely with a CEN workshop, and therefore much of the technical work may be done when the TC launches a NWI. Therefore, the scope in a 'preparatory document' for a NWI on Learner Mobility reads like a business rationale: "There is a need for European metadata descriptions for LET offers (...) Europe has got a unique complexity of didactics...". When the NWI was presented for balloting, it had a nearly finished specification (from CEN WS-LT). Therefore, the scope on this occasion functions more as a summary of the standard, describing exactly what it is doing, whom should use it, what they get, and what they do not get.

The NWI descriptions of TC353 thus show that delivering a convincing rationale plays a more pivotal role than providing a well-defined scope. 'Scope' and 'Purpose and justification' are mixed up, e.g., as in the proposal for the fourth TC353 standard where the scope reads, "the harmonisation of these efforts is now necessary towards a European solution in order for providers to develop a new generation of technology-enhanced services for learners". Purpose and justification is furthermore spread out in several sections in the proposals, i.e., Outcome, Rationale, Policy relevance, and Market impact.

In summary, document analysis of the TC353 registry shows that argumentation on organisational and policy/market rationale is the focus of activities generating new work items. Formal procedures for structuring documents are therefore adapted. Scope sections are used for justification of the project, and there are few '*what'* statements that could instigate technical discussions at this stage. On the other hand, in some of the proposals an 'Approach' section is found (not part of the 'Form A' template), which directly invites discussions on '*how'* the project should be developed, also technically.

3.2 ISO/IEC JTC 1/SC36

SC36 has been operating since 1999 and has produced 30 standards, more than 2400 committee documents and many thousands of working group documents. Even with a longer practice and a stronger secretariat support, this committee adapts the procedures laid out in templates and Directives. For example, the scope of an e-Textbook standard was introduced by three paragraphs under the subheading of "Context", before stating the scope, which is arguably very general: "[to] specify an information model that will enable interoperability and exchange of learning resources contained within it".

This document analysis is limited to scope statements of all projects, either in process or published. The findings of the European study are replicated. Many justifying statements like the following are common: "this standard is intended to meet the need of learners…", "will support legal requirements", "equal access to education or information", "solve resource sharing and interoperability problem", "will integrate the most relative specifications mentioned above which apply to the field", "ensure that LET environments reflect the specific needs of mobile learners", etc. It can be claimed that these explanatory pieces of text are more directed to audiences outside the development team than to the experts who will design the specification.

Vagueness and internal use of delimitation statements. Scope is supposed to be "a series of statements of fact" [2]. It could be argued that the statement "a quality standard 'complements' another standard, 'which does not have detailed processes'..." is a kind of fact. However, the role of the statement is not delimitation to focus work, but to argue about rationale. Describing what should be done by characterising what others are not doing is not a particularly effective way to circumference one's project. In this case, the project should produce "a quality framework for both creation and delivery of e-Tests", consisting of "a quality model and a Process reference model". With the term 'e-Test' not defined, the use of computer science constructs like 'information model', 'reference model', and 'framework' do not give the domain expert much direction for their work.

Understanding of the scope hinges upon the use of commonly used or well-defined key concepts describing the subject of the standard. Some SC36 scope statements, however, lack specificity concerning the subject, as if it is internationally common knowledge what is meant by terminology such as "e-Tests", "Proficiency Level Information", "e-Portfolio", "e-Textbook", "e-Schoolbag", etc. If the subjects as such are not defined, one may expect them to be defined indirectly by defining "the aspects covered, thereby indicating the limits of applicability of the document or particular parts of it" as the Directives prescribes [2]. However, stating that the subject (e.g., the e-Textbook), will be described by a "concept model" or "guidance regarding usage" does not actually explicate the aspects covered by the standard.

The scope template accommodates "Exclusions (if any)"; however, it advises against specifying exclusions (which are an endless domain). Nevertheless, in some SC36 scope statements the exclusions are more specific than what is defined in scope. It seems that the role of these statements is not to point to “other existing ISO or IEC technical committee(s)” to help the users of standards browsing the catalogue. Analysis of the documents and participatory observations make it clear that the exclusions in these documents are used to build a case for further standardization work within the same community, often organised as a “project split” (within SC36) but within the Directives described as “subdivision of the subject matter” [2]. It is arguable that the exclusion parts of scope statements in SC36 are used for internal purposes (to keep control of new work within the committee itself), given that no *published* SC36 standards define exclusions.

Scope creep. Evidence shows that SC36 scope statements are sometimes changed during the technical work. The boundaries of what is standardized are moved, both as a result of editorial work and of the comments resolution processes.

How scope statements are developed may provide an indication of the quality of the development process of the standard. Some of the data of this study could be interpreted this way. To expert participants of SC36 some of the projects are known to be severely delayed, causing conflicts in the working groups, etc. A closer look at the scope description of these projects reveals that some scope statements:

- are often verbose (often covering many pages);
- do not clearly describe what the standard does (e.g., “this (multipart) standard identifies and summarizes principles governing (…) requirements which are generic and applies them to the field of learning, education and training”;
- include a number of references to standards and policy documents in other domains; and,
- include a number of sections on exclusions, and what is defined out of scope is defined more in detail than what is in scope.

In reviewing the document history of these draft standards it is sometimes evident that the scope is defined at the end of the specification process, which explains the use of a great number of editor’s notes arguing for the different alternative wordings.

4. Discussion

These preliminary case-studies of TC353 and SC36 scope statements show that sense-making about scope goes beyond discussing the circumference of the object of standardization. Scope plays a major role in the Design-Sense-making-Negotiation (DSN) cycles [8, 9] of standardization. Applying an Actor-Network theoretical perspective [13], scope serves as an *actant* that could be mobilised in support of one’s view. The more diffuse the scope is the more possibilities there are to claim that a position is within or out of scope. It is said that changing scope during a project “wreaks havoc with project timelines and lowers the morale of the development team” [14]. Nevertheless, minutes from Comments Resolution Meetings (CRM) of SC36 show that scope statements often receive comments and are updated on a regular basis throughout projects.

The role scope plays in standardization does not necessarily reflect the role it is given in the Directives. The SSB takes a committee and document management position, where the scope statement plays the dual role of providing a separation of work and an abstract of published work. In the staged development process, there is a time for asking why work should be done, and another time for doing the work. This is why the rationale, the Purpose and Justification section is left behind in the NWI proposal, and not included in the standard document.

4.1 The narrative role of scope

The SC36 case-study shows that statements answering *why* questions do find their way into scope descriptions. To explore why this is happening, interpretive communication studies theory, in particular the Dramatism developed by Burke, have been studied [12]. The dynamics in the amendment of scope statements demonstrated could be interpreted as a strong need to construct a *narrative* that represents the project and is able to enlist support, what Burke calls a "representative anecdote" [12]. In the case of a standard project the representative anecdote is the business case for the standard, which is negotiated and renegotiated, often in discussions on scope. Burke came up with a dramatistic pentad as an instrument to discover how the speaker persuades the audience to accept his view of reality. Positioning the definition of scope in the narrative or a dramatistic perspective brings to the fore the interrelationships between the different questions that constitute standardization discourse. The act, what is to be done, has to take place on a scene, in time and places (when and where). Reduction or expansion of the scope, of what is to be done, can only be discussed with reference to a purpose. If the drama is going to take place, agents have to move with agency, i.e., using means, organisation, commitment and methods.

4.2 The role of explanations in scope statements

In DSN activities, the core processes of standardization, scope is a 'common starting point', a statement that ideally all parties commit to at the outset of a project. The commitment has nothing to do with the acceptance of the encircled entities identified in the scope, but with the presumption that a partially shared understanding is within reach as part of the normal process. Our data show that in pre-structured documents, narratives make their way into sections reserved for only certain kind of questions. Also, scope statements are found worded as rich explanations. "An explanation is defined as a complex kind of speech act put forward by a proponent in a dialogue to meet a certain type of request made by a respondent" [15]. In explanations why-questions play an important role as *why* is the question that provokes the widest response [10] and is the question that often initiates the dialogue [16]. And, as Hoel and Mason [16] have pointed out, *why*-questioning spans a broad variety of semantics.

If the original intention of the scope statement in a standard is to be maintained, *why* questions related to justification of a claim are 'out of scope".

It is of course debatable whether scope statements are the proper place for the inclusion of explanatory content. Scope, after all, is about stating with precision and clarity what the boundaries of the specification are. But it is also very much the case that explanatory content performs an important role within the standardization life-cycle. While the foregoing discussion has been principally focused on the early stages of standards development there also exist numerous examples within multi-part SC36 standards of "afterthought" parts that are entirely focused on explanations of "how to use Standard ##" [17].

5. Conclusions

The conversational and rhetorical aspects of scope definitions cited in this study (and also implicit in the DSN theory of standardization) pose questions about how standardization processes are managed. It is not possible to exclude *why* questions from the specification process, even if these questions have little "support" from the formal structure of the process, i.e., document templates, Directives, staging, etc. If the value of explanatory content is to be preserved, however, then a placeholder for *rationale* may improve the quality of ITLET standards. It is also arguable that good management is more than sticking to the formalities. Knowing how important discourse on purpose and justification is, a working group chair may make space for discussions of this nature in special seminar like sessions, presentations, national body document contributions, etc. There is a need for further research on processes in the early phases of standardization projects, especially related to requirement elicitation and how requirements will inform the technical work in ITLET standardization.

References

[1] ISO (2011) ISO/IEC Directives, Part 1, Procedures for the technical work

[2] ISO (2004) ISO/IEC Directives, Part 2, Rules for the structure and drafting of International Standard

[3] Akkermans, H., & van Helden, K. (2002). Vicious and virtuous cycles in ERP implementation: a case study of interrelations between critical success factors. European Journal of Information Systems, 11(1), 35–46. doi:10.1057/palgrave/ejis/3000418

[4] Rittgen, P. (2010). Quality and perceived usefulness of process models. Proceedings of the 2010 ACM Symposium on Applied Computing, 65–72.

[5] Moody, D. (2005). Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. Data & Knowledge Engineering.

[7] Hoel, T. and Holtkamp, P. (2012), Requirements modelling in International Information Systems Design – what competencies are needed and how to manage them?, under publication

[7] Frederiks, P. J. M., & van der Weide, T. P. (2006). Information modeling: The process and the required competencies of its participants. Data & Knowledge Engineering, 58(1), 4–20. doi:10.1016/j.datak.2005.05.007

[8] Fomin, V., Keil, T., & Lyytinen, K. (2003). Theorizing about Standardization: Integrating Fragments of Process Theory. in Light of Telecommunication Standardization Wars" Sprouts: Working Papers on Information Environments. Systems and Organizations 3 (Winter 2003).

[9] Hoel, T., Mason, J., & Pawlowski, J. (2012). Making sense of sense-making in standardization processes: focusing on the "why" to improve ICT standards (submitted for publication).

[10] Mason, J. (2008). A model for sense-making: exploring why in the context of learning and knowing. Proceedings of the 16th International Conference on Computers in Education, (pp. 545-549), Taipei, Taiwan: Asia-Pacific Society for Computers in Education. Available online http://www.apsce.net/ICCE2008/papers/ICCE2008-paper286.pdf

[11] Schoechle, T. (2009) Standardization and Digital Enclosure: The Privatization of Standards, Knowledge, and Policy in the Age of Global Information Technology, ISBN: 978-1-60566-334-0

[12] Burke, K. J. (1945). A grammar of motives (pp. XXIII, 530 s.). New York: Prentice-Hall.

[13] Latour, B. (1999) On Recalling ANT. In Actor Network Theory and After, edited by J. Law and J. Hassard. Malden, MA: Blackwell.

[14] Ashmore, D.C. (2003). The J2EEE™ Architect's Handbook, DVT Press

[15] Walton, D. (2004). A new dialectical theory of explanation. Philosophical Explorations, 7(1), 71–89. doi:10.1080/1386979032000186863

[16] Hoel, T., & Mason, J. (2011). Expanding the Scope of Metadata and the Issue of Quality, in T. Hirashima et al. (Eds.) (2011). Proceedings of the 19th International Conference on Computers in Education. Chiang Mai, Thailand

[17] ISO/IEC 19796-5 Information Technology for Learning, Education and Training – ITLET – Quality Management, Assurance, and Metrics – Part 5: Guide "How to use ISO/IEC 19796-1"

Moving toward active reading: constructing a scaffolding tool for inquiry based reading on e-books

Li-Hsiang HSIAO*, Hsin LEE, Shing-Tsaan HUANG, Liang-Yi LI, and Gwo-Dong CHEN
Department of Computer Science and Information Engineering, National Central University, Taiwan
**shair78227@gmail.com*

Abstract: Reading actively helps students become better readers. Unfortunately, students in modern education system have to acquire so plentiful knowledge in textbook that they absorb them all passively like sponge absorbing water [1] without active participation and thinking. Therefore, strategies have been proposed to teach students active reading. Questioning is a very important strategy for successful active reading. In this paper, we synthesized some reading strategies such as SQ3R, self-questioning, and QARs into "inquiry based reading". The main spirit of inquiry based reading is "asking questions" and "finding answers". However, it is difficult for most students to implement while reading. The reason can be generalized as two major problems: (1) students have no capability to ask questions and (2) they have no motivation to ask questions while reading. As a result, we bring up a training process and a scaffolding tool on an e-book to solve these two problems. A validation of the instrument is processed containing 32 undergraduate students. The results show that students who used the scaffolding tool have learned inquire based reading better after training.

Keywords: active reading, inquiry based reading, self-questioning, QARs, e-book.

Introduction

Active Reading [2] enables students to think, comprehend, and maintain their interest in what they are learning instead of passively reading and receiving all the information in the book. Active reading tends to involve writing, such as highlighting, note-taking and annotation [3].On the other hand, another factor of success in active reading may be asking questions and finding answers [4]. Researches have already developed some appliances to assist readers in writing. In this study, we focus on the issue of questioning and answering while reading. There were no previous research has done this through a digital approach. We try to construct a scaffolding tool to help readers moving toward active reading through asking and answering.

A good reader should ask questions and find the answers while reading [5]. Many reading strategies have been developed by requiring the reader to ask question while reading. Self-questioning is an effective strategy for improving reading comprehension [6]. It requires student to ask question before, during and after reading. The Question Answer Relationships (QARs) [7] strategy is devised as a way for students to understand that the answer to a question is directly related to the type of the question asked. QARs assist students in differentiating among questions based on where the answer can be found: either In the Book or In My Head [8]. The SQ3R strategy [9] divides reading process into five steps:

survey, question, read, recite, and review. The question step requires a reader to formulate questions about the content of the reading, for example, by converting headings and sub-headings into questions, and then looking for answers in the content of the text. Because these methods can lead the reader to inquire the information inside the book by actively questioning and answering, we summarized this kind of reading style "inquiry based reading".

Asking questions and finding answers are the main methods to implement inquiry based reading, but actually, it is not easy to make students accustomed to asking questions while reading a traditional paper book. The two difficulties are: First, students do not have capability to do inquiry based reading (active reading). Second, students have no motivation to learn, try and use this inconvenient, complex reading behavior. This study is in order to solve these two problems.

For the first problem, students have to learn how to ask questions through not only a short-term lecture but a training process. Megan D' Ambrosio [10] has developed a training process in 2004. In this study, we developed a four-step training process referred from Megan to make students be familiar and skilled with inquiry based reading gradually.

For the second problem, we deduce there are two reasons. The first reason, it is inconvenient for students to record questions if they have to write down their questions and present answers they find [11]; these time-consuming works may cause high cognitive load and reading interruption. The second reason, it is not easy to see better reading achievement when they just start to try questioning and answering while reading. These two reasons reduce their motivation to take inquiry based reading. The scaffolding tool was designed to solve the second problem by facilitating student's task. It can support student recording questions and answers completely. In addition, the tool has to make students feel fulfilled while doing inquiry based reading.

E-book is a suitable approach to implement active reading [12]. The reason we do the research is that e-book bring some new possibilities of implementing scaffolding tools for a better operation, display method and data storing way than traditional book and paper sheet. All the data reader made can be add like a layer on the textbook which can be hid or shown or change by easily operations. Also, these data was stored in the database which can be accessed for any purpose.

1. Study Description

1.1 Teaching student "How to do inquiry based reading"

A teaching material was edited to inform students of "what is inquiry-based reading", "why it is beneficial to do inquiry based reading" and "how to ask questions and find answers", and this material lists different perspectives of questioning skill.

The perspective of reading text content introduces the "question" step from SQ3R. Here students can ask questions from "table of content" to capture the "range of knowledge" and "structure of knowledge." They then can convert title, sub-title and keywords into question to get important points and summary of a paragraph. They can transform tables into related questions to get knowledge by organizing and structuring into tables. They ask questions related to figures to know the illustrated information such as "time", "space" or "categories".

The perspective of reading timing introduces self-questioning strategy that divided the questioning timing into before reading, during reading, and after reading. Questions asked before reading include "What is this about?", "Why am I reading this?" to prepare a process of active inquiry. Questions asked during reading are to check their understanding. Even after reading, students should ask questions to recheck and reflect on what they learned.

The perspective of the character of question and answer we summarized from King's "three levels of questions" [13], Call & Patricia E's "asking questions on three levels" [14], and Raphael's QARs (Question Answer Relationship) classifies questions and corresponding answers into three levels. The first level, factual questions have corresponding answer right there in the text. These kinds of questions can begin with "who, where, what, and when." Asking the first level factual questions may help readers memorizing the facts in the book. The second level, comprehension questions require readers to integrate, interpret, and infer the information. These questions could begin with "why and how". In this level, readers have to think and search in the book to find out and then compare and summarize the answers. The third level is critical questions where readers have to connect the knowledge in the book with their prior knowledge and bring up their own opinion to ask this kind of questions. Expected answers of the third level questions may be unable to find out in the book. Questions at this level may begin with "agree or disagree," asking to explain why or why not, or to criticize and ask "what if."

1.2 Training process for inquiry-based reading

Referring to the approach of Megan, the training process designed in this research contains four steps:

The first step, students read the teaching material "How to do inquiry based reading" to know the goal of inquiry based reading and increase their motivation to do inquiry based reading. Students can keep the teaching material for following steps.

The second step, students read an assigned textbook scope and guided by some demonstrated questions corresponding to the context scope that was previously wrote by a teacher. Each demonstrated question has an explanatory information telling students the reason for asking this question, the level this question belong to, benefits of answering this question and pick out the irrelevant questions. Students look at demonstrated questions and explanations while reading the textbook in this step.

The third step, students read another context scope and get fewer demonstrated questions. In this step, students try to ask questions and find answer by themselves. After they list their questions and find answers, a teacher will examine the questions and answers students make. Then, teacher gives them feedback such as the quality of questions, criticisms, and suggestions.

The fourth step, students read the comment given by teachers, self-reflection questions and answers they make in the last step. Then, students have to do inquiry-based reading independently (Figure 1).

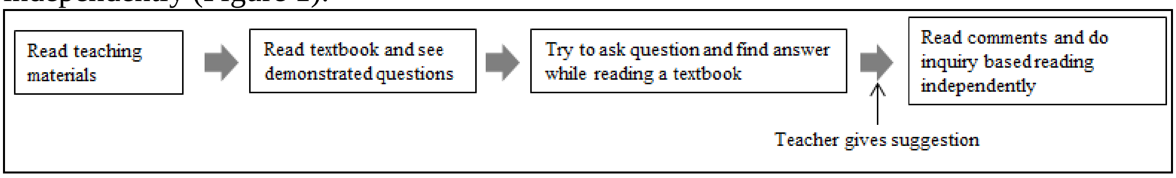


Figure 1. Steps of inquiry based reading training process

1.3 A scaffolding tool on e-book

The scaffolding tool is constructed on a web-based e-book so that users can read e-book on any device (Figure 2) that support web browser and is able to connect to the Internet. All of the data reader make will be saved in the database. The system contains three main functions: questioning function, answering function, and a question-answer tree structure.

When readers use the questioning function, the e-book will show a draggable post-it note that they can type their question in the post-it note and then the post-it note can be reduced into a small, translucent icon. The icon can be expanded back to the post-it note by clicking it, and questions can be modified in the post-it note. The icon is also draggable that the readers can drag it to the most suitable position related to the context of the question (Figure 3).

Figure 2. Web-based e-book

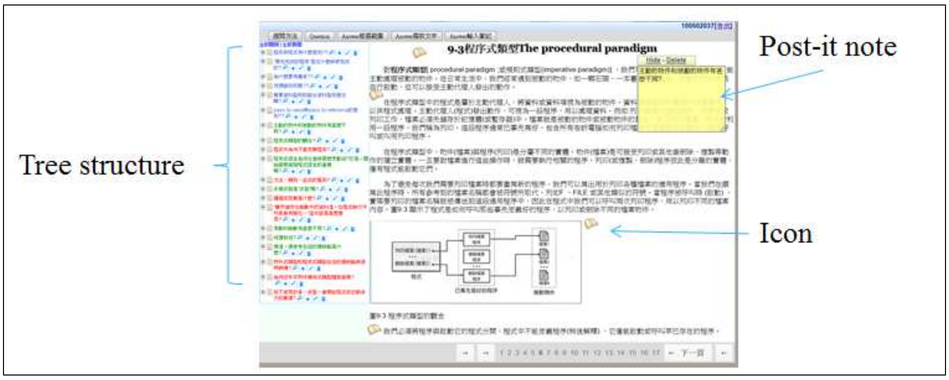


Figure 3. Questioning function and Question-answer tree structure

The answering function has three methods for the readers to answer their questions. The first, they can draw a rectangle area on the textbook by dragging. The second, they can select the text in the page for an answer. The third, they can type their answer with a keyboard (Figure 4). The first and second function would produce marks in the textbook content, and users can “hide” or “show” these marks on their own.

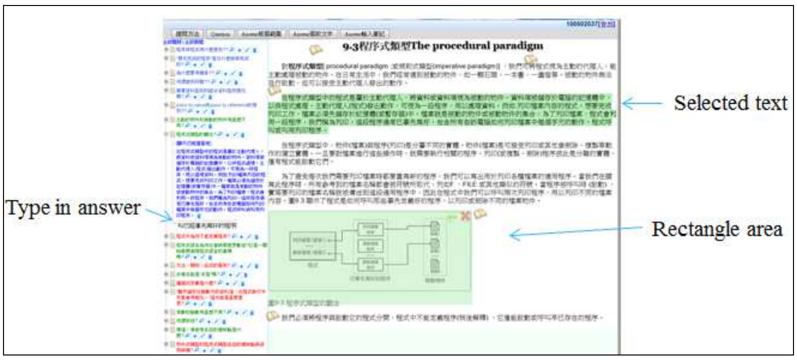


Figure 4. Three ways to present answer

Every question and answer would be added in the tree structure (Figure 3) automatically, and the tree structure was fixed on the left side of textbook content. Each node represents a question, and expansion of the node shows answers or explanations of its parent node question. Question text displayed in three different colors represents three different meanings: blue means this question was initially provided by a teacher in the system, red means the node question has not been answered yet, and green means the node question have been answered. Every questions and answers are hyperlinks that the readers can easily go to the corresponding textbook page and show the post-it note of the question or the marked answer.

2. Methods

An experiment was constructed to evaluate the effect of training process and the difference between using scaffolding tool and traditional approach. The textbook we selected to use in the training is “Foundations of Computer Science” chapter nine: “Programming Language” and chapter ten: “Software Engineering”.

2.1 Participants

Participants were 32 undergraduate students from Department of Computer Science and Department of Electrical Engineering, National Central University, who have some prior knowledge of Computer Science. Each student was paid 400 NTD (13 USD) for participating in the experiments. Students were distributed randomly into experimental group and control group (16 participants in both groups). The experimental group used the scaffolding tool on an e-book while taking the training process of inquiry based reading. On the other hand, the control group read a traditional paper book and used paper sheets to write down their question and answer. In the second and third step, demonstrated questions and corresponding explanation were initially displayed in the system for the experimental group and printed on a paper for the control group.

2.2 Procedure

Before training, both group participants take a simple questionnaire (five-point Likert scale) to know student’s attitude about inquiry based learning. Also, students take a pretest to know their prior ability of questioning.

The training process time distribution of each steps are: 30 minutes for the first step to read the teaching material of “How to do inquiry based reading,” 30 minutes for the second step to read a five-page textbook content related to demonstration questions and corresponding explanation, 60 minutes for the third step to read an eleven-page textbook and related questions, and 60 minutes for the fourth step to do inquiry based reading independently by reading a thirteen-page textbook. The training process was sliced up into time-slices in two weeks.

After training, students took a five-point Likert scale questionnaire designed based on Kirkpatrick’s learning and training evaluation theory and Kirkpatrick’s four level of evaluation model [15] (7 questions for reaction level,7 for learning level ,11 for behavior level and 5 for results level). In addition, a simple posttest exam was taken to evaluate the learning effect of textbook content students read in the fourth step.

2.3 Evaluation value

In addition to evaluating the training effect by posttest questionnaire, we calculated three values for assisted verification results. Two values for the ability of questioning are student's "total question quantity" and "the sum number of level 2, 3 questions". One value for the ability of answering is "the average points of each question's answer completeness" (0 to 5 for unanswered to complete). Two teachers assisted in this study to classify students' question level (three level mentioned in 1.1) and rate the answer completeness of each questions, they discussed with each other if they had different judgments to a question.

3. Results

The average question number of pretest shows no significant difference ($p=0.58$, $p>0.05$) between the control group ($M=2.27$, $SD=0.9$) and the experimental group ($M=2.5$, $SD=1.0$). The pretest questionnaire is shown in Table 1. Most students agree that questioning can help their reading (Question 2, 3). But they did not know how to do this very well (Question 1, 4).

Table1. Average points of pretest questionnaire

	Question statement	Points
1	I often take an attitude of skeptical to the information in the book while reading	2.93
2	I think it is helpful to ask questions while reading	3.90
3	If I had a question while reading, I would write it down.	3.46
4	I often don't really know whether I have a question while reading.	2.83

Table 2 shows the result point of pretest questionnaire. The average point scores of the experimental group are higher than the control group, especially the behavior level. The three evaluation values from third and fourth step are shown in Table 3. Total question quantity in the third step and the forth step: Experimental group are significantly higher than the control group ($p<0.05$). The sum number of level 2, 3 questions shows the experimental group is higher than control but the standard deviation is high. The average points of each question's answer completeness shown the experimental group is higher than the control group. The result score of learning effect posttest are 82.28 for the control group and 81.75 for the experimental group.

Table2. Average points of posttest questionnaire

	Control Group		Experimental Group	
	Average points	Cronbach's alpha	Average points	Cronbach's alpha
Reaction level	3.84	0.711	4.09	0.752
Learning level	4.08	0.773	4.24	0.764
Behavior level	3.41	0.742	4.47	0.894
Results level	4.22	0.628	4.24	0.906

Table3. Three values of evaluation

	Step	Control Group	Experimental Group	p
Total question quantity	3	M=8.36 SD=3.69	M=12.75 SD=5.40	0.014
The sum number of level 2, 3 questions	3	M=4.85 SD=2.87	M=7.93 SD=4.18	0.025
The completeness of answer (0 to 5)	3	M=1.05 SD=1.22	M=1.94 SD=0.85	0.025
Total question quantity	4	M=10.64 SD=6.66	M=20.18 SD=7.56	0.001
The sum number of level 2, 3 questions	4	M=5.85 SD=4.09	M=6.93 SD=4.37	0.488
The completeness of answer (0 to 5)	4	M=1.46 SD=1.74	M=2.43 SD=0.75	0.051

4. Discussion

According to the posttest questionnaire, the training process certainly has made students realize the importance of inquiry based learning and learn how to ask questions and find answers. In general, students do inquiry based reading better using the scaffolding tool than using pen and paper sheet especially at the behavior level.

The two group student's reading time is the same, but the students who used scaffolding tool generated more questions than the other group of students. Therefore, this indicated students promoted higher motivation in inquiry based reading by the high-efficiency operations.

The sum number of level 2, 3 questions is higher in the experimental group meaning these students were willing to spend time on thinking deeper questions. However, each student’s question styles are very different. According to the observation, with the increase of the proficiency to the tool (compared with the third step) some students in the experimental group asked so many first-level questions and answered them “right here” in the textbook content easily that they forgot to think deeply and ask some higher-level question. Nevertheless, it can be improved by reminding.

As to the completeness of answer, students in the control group can be classified into two types: First type students concentrated on asking questions and they indeed generated immense number of questions, but they answered very little of them, some students even answering no question. The other type of students paid most attention to answering their questions. As a result of spending long time on writing (sometimes also drawing) answers lead to the lack of time, this type of students asked very few questions. Compared to the control group, all of the students in the experimental group answered more than 80 percent questions they asked. Furthermore, some students of the control group mentioned sometimes they found other questions while answering and led deviate from the content.

Finally, the exam score shows both groups of students learned well through inquiry based reading, and asking questions during reading did not have adverse effect on reading comprehension.

5. Conclusion

Helping students to become active readers (learners) is always an important topic in education. Since inquiry based reading is one of solutions to implement active reading, it is worth to research how to improve students' capability and motivation of questioning and answering. In this study, the scaffolding tool we developed on e-book successfully increased students' performance in inquiry based learning. Even if there are more effective training methods were invented, this scaffolding tool can still make its result even better.

In addition, because of questions and answers were presented clearly on the tree structure, students felt interested and curious of the questions which are unanswered (displayed in red). The tool could be enhanced by combining with other tools such as search engine or discussion forum to support students to figure out their unanswered questions and construct their knowledge more completely after they finished reading.

Acknowledgements

This work is partially supported by the National Science Council, Taiwan, under the grant number NSC-100-2511-S-008-003-MY3

References

[1] M. Neil Browne, Stuart M. Keeley (2005) *Asking the right questions: a guide to critical thinking* 8th ed.

[2] Adler, M. J., Van Doren (1972), C. L. *How to read a book*

[3] Hartmut Obendorf (2003). Simplifying annotation support for real-world-settings: a comparative study of active reading *HYPERTEXT '03 Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*. 120 - 121

[4] Ernest Balajthy (1984). Using Student-Constructed Questions to Encourage Active Reading *Journal of Reading* Vol. 27, No. 5 408-411

[5] Paris, Scott G, Meyer (1981) . Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behavior* Vol 13 , 5-22.

[6] Wong, B.Y.L. (1985). Self-questioning instructional research: A review. Review of Educational Research, 55, 227-268.

[7] Raphael, T.E. (1986). Teaching question-answer relationships, revisited. *The Reading Teacher*, 39, 516-522.

[8] Caldwell, J.S., & Leslie, L. (2005). Intervention strategies to follow informal reading inventory assessment: So what do I do now? Boston: Pearson Education, Inc.

[9] Robinson, F. P. (1961). *Effective study*, Rev. edition, New York, NY: Harper.

[10] Megan D'Ambrosio (2004) . *Asking Questions As You Read* http://classes.seattleu.edu/masters_in_teaching/teed521/professor/Literacy/questions.pdf

[11] Lucia Lumbelli, Gisella Paoletti, Tiziana Frausin (1999). Improving the ability to detect comprehension problems: from revising to writing Learning and Instruction 9 143–166

[12] Landoni Monica (2008). *The active reading task: e-books and their readers*, New-York : ACM ,New York.

[13] Alison King (1994). Guiding Knowledge Construction in the Classroom: Effects of Teaching Children How to Question and How to Explain. *American Educational Research Journal*, Vol. 31, No. 2, pp. 338-368

[14] Call, Patricia E. Reflective (2000). questioning: A strategy to review notes. *Journal of Adolescent & Adult Literacy*, Vol. 43 Issue 5, p487, 2p

[15] Kirkpatrick, D. L. andKirkpatrick, J. D. (2006). Evaluating Training Programs: The Four Levels. San Francisco, CA : Berrett-Koehler

Proposal and Evaluation of a Method of Estimating the Difficulty of Items Based on Item Types and Similarity of Choices

Shinichi IKEDA[a], Teruhiko TAKAGI[b], Masanori TAKAGI[c], Yoshimi TESHIGAWARA[a]

[a] *Graduate School of Engineering, Soka University, Japan*

[b] *Graduate School of Information Systems, University of Electro-Communications, Japan*

[c] *Faculty of Software and Information Science, Iwate Prefectural University, Japan*

{e11m5202, teshiga}@soka.ca.jp

Abstract: In recent years, on the idea of supporting the composition of the tests by using statistical data, such as the difficulty level of the items that constitute the tests, has been investigated. In general, item response theory (IRT) is used to quantify the difficulty level of items. However, this approach requires that the items are answered by many learners in advance and it is difficult to ensure that all items in the bank are answered. We propose a method of estimating the difficulty level of unanswered items. In our method, the level of new items is estimated from the level of similar existing items based on the differences between item types and the similarity between choices. A simulation experiment shows that the difficulty levels calculated by IRT and by using the proposed method can have a reasonable correlation. However the results obtained using the new estimation method can be very different from the IRT results if incorrect answers to an item are similar to the correct answer.

Keywords: IRT, E-testing, Difficulty Level, Similar Item, Item Types

Introduction

In recent years, Web-based testing, commonly referred to as "e-testing", has been attracting much attention [1][2]. In e-testing, more reliable tests can be conducted by preparing an item bank with managed statistical data [3] that includes information on the difficulty level of items and correct answer rate. In addition, a number of studies in the literature have shown support for composing tests through the use of such statistical data [4][5][6]. In these studies, item response theory (IRT) [7] is used to quantify the difficulty level of test items. In order to estimate the difficulty level, the items need to be answered by many test takers (subjects) in advance. Furthermore, new items are added periodically to replace items in the item bank. However, it is hard to ensure subjects answer all items in the item bank to gather complete data, and estimating the difficulty of new items when they are added to the item bank takes time and resources.

Therefore, the objective of this study is to estimate the difficulty level of unanswered items. The difficulty of items can change depending on how the question is asked (the item type) and the similarity of answer choices [8][9] and, in this paper, we focus on such differences. We also restrict our considerations to multiple-choice items. We propose a method of estimating the difficulty level of items by comparison with existing "similar" items. Similar items are defined as being those where the knowledge questioned and the knowledge needed for the solution are similar. Items are then classified according to a measure of similarity [10].

1. Item Response Theory

This section examines a method for estimating the difficulty level of similar items using item response theory (IRT). A statistical model, called the IRT model, is used to reveal the statistical properties of test items. The properties of items are given by the item characteristic curve (ICC), where the vertical axis is the correct answer rate and the horizontal axis is latent ability (θ), representing the learning ability of the candidate, which does not depend on the candidate population. In this study, a commonly used two-parameter logistic model (2PLM) is applied. The probability of subject i with learning ability θ_i answering item j correctly is defined as

$$P_j(\theta_i \mid a,b) = \frac{1}{1+\exp\{-Da_j(\theta_i - b_j)\}} \tag{1}$$

where a_j is the discrimination level showing the degree to which item j discriminates between subjects, and b_j is the difficulty level of item j (typically, $-3 \le b_j \le +3$) [7]. Figure 1 shows three ICCs on the same graph. All have difficult combinations of discrimination level and difficulty level. When the curve moves to the right, the difficulty level of the item increases because the probability of a correct answer is low at the lowest ability level. When the curve becomes steep, the discrimination level of an item is high. In the 2PLM, the slope of the curve is maximized when the probability of a correct answer is 0.5, and this value of the slope is the discrimination level. In addition, when the answers of the n items of subject i are given by $\mathbf{u}_i=\{u_{i1}, u_{i2},\cdots, u_{ij},\cdots, u_{in}\}$, where $\mathbf{u}_i$ is 1 in the case of a correct answer and 0 in the case of an incorrect answer, the probability of the vector $\mathbf{u}_i$ is given by

$$P_j(u_i \mid \theta_i) = \prod_{j=1}^{n} p_j(\theta_i)^{u_{ij}} q_j(\theta_i)^{1-u_{ij}} \tag{2}$$

where $p_j(\theta_i)$ is the correct answer rate of subject i to item j, and $q_j(\theta_i) = 1- p_j(\theta_i)$.

By using such a model, it is possible to estimate the learning ability θ of a subject, the discrimination level (a), and the difficulty level (b) from the test answers of all the subjects. However, the subjects must answer items in advance for these parameters to be estimated. In general, the number of answers required in order to estimate the difficulty level of items using IRT is 1,000 in 3PLM and 500 to 1,000 in 2PLM [11].Therefore, in this paper, a method for estimating the difficulty level of unanswered items is studied.

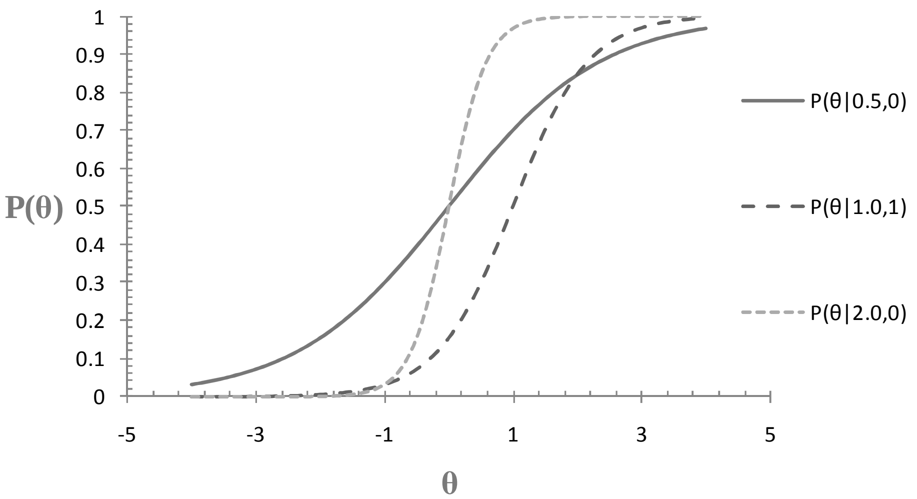


Figure 1 Item Characteristic Curve of 2PLM

Table 1 Item Types

Type ID	Item Type (Example)
Pa+	Select a correct example or explanation of a technical term. (Select the correct description of Morse code.)
Pa-	Select an incorrect example or explanation of a technical term. (Which of the following is not a Real-time Distributed System?)
Pb+	Select a technical term having the same type or attributes as a given technical term. (Which of the following is a type of visual communication?)
Pb-	Select a technical term having a different type or attribute from a given technical term. (Which of the following is not a type of visual communication?)
Pc+	Select a correct example or explanation of something relating to a technical term. (Which of the following is a problem affecting data management in a distributed environment?)
Pc-	Select an incorrect example or explanation of something relating to a technical term. (Which of the following does not have an impact on the structure of a computer network?)
Pd+	Select a correct technical word relating to a given technical term. (Which of the following devices is suitable for telephone communication?)
Pd-	Select an incorrect technical word relating to a given technical term. (Which of the following practical applications of a computer network does not appear in banks or convenience stores?)
Pe+	Select a correct combination of a technical term and an explanation of it. (Select a correct description of 4 layers in the OSI reference model.)
Pe-	Select an incorrect combination of a technical term and an explanation of it. (Which of the following is not a correct description of 4 layers in the OSI reference model?)
Pf	Select a correct technical term based on an example or explanation of it. (What is the host-centralized system which uses a single host computer and multiple terminals?)
	Others.

2. Method of Estimating the Difficulty Level

2.1 Item Type and Difficulty Level

In a preceding study, items were classified into 11 types according to how knowledge is tested [12]. We classified items based on the basis of their content and the answer choices. Also, when classifying items, we took into account whether the item requested the subject to select a correct or an incorrect answer. Table 1 shows the 11 item types and provides examples. In Table 1, the “Other” category includes computational items, fill-in-the-blank items, and flawed items.

The difficulty level of items can change depending on the phrasing used, such as whether the item seeks knowledge of a technical term (Pa and Pf in Table 1) or the item requires the subject to apply knowledge and use it for the answer (Pc and Pd in Table 1). Thus, for similar items, the difficulty level of unanswered items may be estimated by focusing on differences in item type.

In our proposed method, the difference of difficulty level between similar items i and j which arises from the differences in the item types is defined as

$$bp_{ij} = (D_{\max} - D_{\min}) \cdot sw_{ij} \tag{3}$$

where D_{max} is the maximum difficulty level in the Item bank, D_{min} is the corresponding minimum, and *sw* is the rate of changes for the range of difficulty levels (D_{max} - D_{min}). Thus, when the difficulty level of an item is known, the difficulty level of similar items can be estimated by adding the difference of difficulty level calculated by formula (3).

2.2 similarity of answer choices and difficulty level

In the case of multiple-choice items, the difficulty level may change according to the similarity of answer choices [9]. For example, the difficulty level of items is increased when the choices include a "confounding answer." On the other hand, the difficulty level of items is decreased when the choices contain an "obviously correct or incorrect answer." One possible measure of the difficulty level of items is the probability that each answer choice is selected (the selection probability). Thus, we propose a method of estimating the selection probability from the similarity of answer choices [13]. In this method, in order to estimate the selection probability, the similarity of each answer is calculated by representing the documents as a weighted collection of terms in a vector space. However, it may not be possible to calculate the similarity if there are few terms contained in the question and answer choices. So, terms that are related to the question or answer choices (related terms) are extracted from the item bank. Then, the similarity of answer choices is calculated using the related terms.

Therefore, in proposed method, the difference of the difficulty level which arises from the difference in the similarity of answer choices between similar items *i* and *j* is defined as

$$bc_{ij} = (D_{\max} - D_{\min})(cv_i - cv_j)bp_{\max} \tag{4}$$

where D_{max} and D_{min} are the same values as in formula (3), v_i and v_j are the unbiased variances of the selection probability for items *i* and *j*, *c* is the number of answer choices and bp_{max} is the maximum value of the difficulty level difference calculated using formula (3). Thus, when the difficulty level of an item is known, the difficulty level of similar items can be estimated by adding the difficulty level difference calculated using formula (4).

2.3 Calculation Procedure for Difficulty Level

In this study, the difficulty level of item *i* (the estimation item) is estimated using the formula

$$b_i = \frac{1}{n}\sum_{j=1}^{n}(bs_j + bp_{ij} + bc_{ij})\,. \tag{5}$$

Here bs_j is the difficulty level of one of *n* similar items (the comparison items), The changes in the difficulty level arising due the difference in item types bp_{ij} are estimated by IRT. The changes in the difficulty level between similar items bc_{ij} are based on the differences of selection probability for each answer choice. The difficulty level of estimation item *i* is calculated by adding bp_{ij} and bc_{ij} to the difficulty level of similar item bs_j. The average value over the *n* comparison items is used as the difficulty level of estimation item *i*.

Figure 2 shows the method for estimating the difficulty level proposed in this study. First, the difficulty level of items used in the test are estimated using IRT (estimated items - bs_j), as shown in Figure 2-(1). Then, from the same item bank, the target items for which the difficulty level must be checked are selected as the estimation items (*i*). Next, the estimation items are used to select similar items from the already estimated similar items (comparison items) as shown in Figure 2-(2). Then, the changes in difficulty level (bp_{ij}, bc_{ij}) are calculated based on the differences in the item type and the selection probability of each answer choice between estimation item *i* and comparison item *j*, as shown in Figures 2-(3)

and 2-(4). After that, Equation (5) is used to calculate the level of difficultly for the estimation items (b_i) using the results from steps 3 and 4. Finally, the calculated result (b_i) is registered as shown in Figure 2-(5).

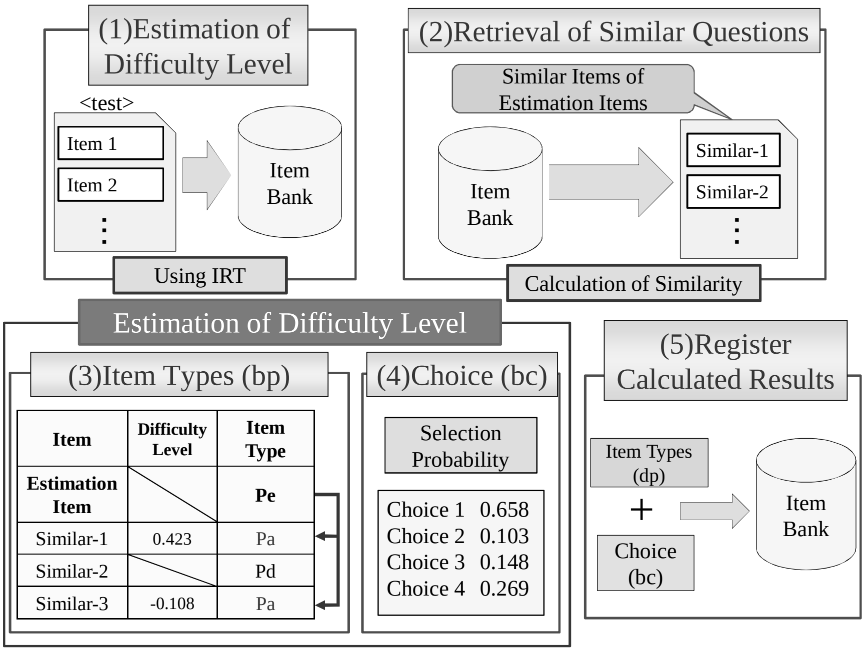


Figure 2 Procedure for calculating level of difficulty

3. Experiment

3.1 Experiment Outline

In this section we describe a comparative experiment that was conducted to verify the relevance of the difficulty level estimated by the proposed method. In this experiment, 1000 items given in previous "System Administrator" and the "Fundamental Information Technology Engineer" examinations are accumulated in the item bank. The differences and the correlation coefficient of difficulty level estimated by the proposed method and IRT are calculated. In the proposed method, the difficulty level is estimated in three ways: using only item types (Dp), using only similarity of answer choices (Dc), and using both item types and similarity of answer choices (Dp+Dc). The 1000 items are used to extract the related terms for estimating the selection probability of answer choices, and then bc_{ij} is calculated from the selection probability.

First, a test consisting of 20 items (Test 1) was conducted. Then, the difficulty levels of the items set in Test 1 were estimated using IRT with a 1PLM. These items were then used as comparison items. In this experiment, Test 1 was given to 82 students in three universities: Soka University, the University of Electro-Communications and Iwate Prefectural University. Second, 15 items similar to those in Test 1 were retrieved from the item bank. Furthermore, 5 items contained in Test 1 are used for items of Test 2 and those items are used for equating of Test1 and Test2. Third, a test consisting of these similar items (Test 2) was conducted and the difficulty levels of the items were estimated by IRT and using the proposed method in three ways (Dp, Dc, Dp+Dc). Finally, the differences and correlation coefficient of difficulty levels estimated by the proposed method and IRT were calculated.

Table 2 Estimation results for difficulty levels.

Item	1	2	3	4	5	6	7
Dp	0.50	0.78	-0.11	1.37	0.45	0.85	-1.95
Dc	-0.25	0.80	-0.60	1.26	0.78	-0.52	-1.88
Dp+Dc	0.48	1.29	-0.11	1.26	0.18	0.22	-1.85
IRT	-0.66	1.40	-0.26	-1.90	-2.36	4.35	-1.90

8	9	10	11	12	13	14	15
-0.57	-1.30	0.76	0.64	0.34	-0.73	-1.50	0.27
0.65	-0.72	0.65	0.86	0.26	-0.96	-1.60	0.30
0.05	-0.94	0.65	0.83	0.49	-1.18	-1.60	0.27
-0.26	-4.10	0.96	-2.36	-1.47	-1.90	-1.90	-1.06

Table 3 The differences in the difficulty levels for each method.

Item	1	2	3	4	5	6	7
Dp	1.16	0.61	0.15	3.27	2.81	3.50	0.05
Dc	0.41	0.60	0.33	3.16	3.14	4.87	0.02
Dp+Dc	1.15	0.11	0.15	3.16	2.54	4.13	0.05

8	9	10	11	12	13	14	15
0.31	2.80	0.20	3.00	1.82	1.17	0.41	1.33
0.91	3.39	0.31	3.22	1.73	0.95	0.30	1.36
0.31	3.16	0.31	3.19	1.96	0.72	0.30	1.33

3.2 Experimental Results

Table 2 shows the difficulty levels of items estimated by the proposed method and IRT. Table 3 shows the differences of difficulty levels between the proposed method and IRT. The correlation coefficients between the difficulty levels estimated by the proposed method and those found by IRT are 0.46 (Dp), 0.12 (Dc), and 0.37 (Dp+Dc). On the other hand, the correlation coefficients are 0.72 (Dp), 0.76 (Dc), and 0.80 (Dp+Dc) when items which have a large difference of difficulty level (items 4, 5, 6, 9, and 11) are removed. Therefore, the difficulty levels of items which have a small difference of, difficulty level estimated by IRT could be predicted quite well using the proposed method.

Since the correlation coefficient for method Dc is the lowest, the estimation of difficulty levels could be affected according to selection probability. Table 4 shows the selection probabilities of large difference items estimated using the proposed method (estimated selection probability) and calculated using answer data (calculated selection probability). The correct answer rates of items 4 and 5 of Test 1 and item 6 of Test 2 are 30% or less. In particular, the calculated selection probability of an incorrect answer choice is the highest in items 4 and 6. In the proposed method, the difficulty level of items is estimated from the difference of the variances of selection probability. However, the variances of selection probability may become equal even if the difficulty level of items is different. Thus, in the case of items for which selection probability of an incorrect answer choice is the highest, the difference between the difficulty level estimated by the proposed method and IRT becomes large because the difference of the difficulty levels is not calculated correctly from the selection probability. For item 5, the variance of selection probability is increased

Table 4 Selection probabilities of items with large differences.

Calculated Selection Probability		Choice 1	Choice 2	Choice 3	Choice 4
Item 4	Test 1	0.27	0.43	0.17	0.13
Item 5	Test 1	0.30	0.21	0.22	0.27
Item 6	Test 2	0.21	0.13	0.57	0.09
Item 9	Test 2	0.88	0.04	0.04	0.04
Item11	Test 1	0.21	0.17	0.24	0.38

Estimated Selection Probability		Choice 1	Choice 2	Choice 3	Choice 4
Item 4	Test 1	0.65	0.35	0	0
Item 5	Test 1	0.39	0.25	0.31	0.05
Item 6	Test 2	0	1	0	0
Item 9	Test 2	0.50	0.17	0	0.33
Item 11	Test 1	0	0.26	0	0.76

because the selection probability of answer choice 4 is estimated to be low. Item 9 of Test 2 is an easy item for which the correct answer rate is about 90%. However, if the selection probability of an incorrect answer rate is estimated to be high, the difference of difficulty levels becomes larger.

On the other hand, items for which the estimated selection probability is 0 are constrained because there are no related terms in the proposed method. In item 6 of test 2, the related terms for incorrect answer choices do not exist, and the selection probability of correct answer choice becomes 1. For item 11 of Test 1, for which the correct answer is "router", the selection probability of incorrect answers "gateway" and "repeater" are 0 because the related terms do not exist. However, these incorrect answers are similar to correct answer and the calculated selection probability is 0.2 for both incorrect answers. Thus, the difference of the difficulty levels becomes larger. The results of the experiment show that the estimation of the difficulty level is strongly influenced by the estimation of the selection probability. Therefore, it is necessary to develop a more accurate method for estimating the selection probability.

4. Conclusions

In order to estimate the difficulty level of unanswered items, the items were first classified into 11 types according to how knowledge is tested. We then estimated the difficulty level of items based on item types and similarity of answer choices. In our proposed method, the difficulty level of similar items is estimated by comparing the item types and the selection probability of answer choices with those of some similar items for which the difficulty levels have already estimated. In addition, a method for estimating the selection probability of answer choices based on representing the documents as weighted collections of terms in a vector space is proposed.

The results of an experiment show that the proposed method could provide estimates of difficulty levels which are close to those estimated by IRT. Therefore, the difficulty level of unanswered items could be estimated when new items are added to an item bank, thus reducing the costs and time when constructing an item bank. However, for some items the selection probabilities of answer choices are not estimated correctly, so the difference between the difficulty levels estimated by the proposed method and by IRT is large. In the future, we plan to develop an extended method to estimate the selection probability more accurately, with a focus on the method of weighting related terms and the deletion of unnecessary terms.

References

[1] Ueno, M., (2005). Web based computerized testing system for distance education. *Educational Technology Research*, 28(1), 59–69.

[2] Ueno, M. & Okamoto, T., (2008). System for online detection of aberrant responses in e-testing. in *Proceedings of the2008 Eighth IEEE International Conference on Advanced Learning Technologies*, ser. ICALT '08. Washington, DC, USA: IEEE Computer Society, 824–828.

[3] Japan Society for Educational Technology, (2000). Educational Technology Dictionary: Jikkyo Press. (in Japanese)

[4] Songmuang, P., Ueno, M. (2008). Development of Prediction System of Score and Time in e-Testing. *Journal of IEICE* J91-D(9), 2225-2235. (in Japanese)

[5] Otomo, K. (2009). Principles and Selected Applications of Item Response Theory. *Journal of IEICE* 92(12), 1008-1012. (in Japanese)

[6] Takahashi, N., Nakamura, T. (2009). Development and Evaluation of the Adaptive Tests for Language Abilities (ATLAN). *The Japanese Journal of Educational Psychology* 57(2), 201-211. (in Japanese)

[7] Baker, F. B. (2001). *The Basics of Item Response Theory. 2nd ed*. (USA, ERIC Clearinghouse on Assessment and Evaluation).

[8] Ikeda, S., Takagi, T., Takagi, M. & Teshigawara, Y. (2011). A Study on a Method of Estimating the Difficulty of Quizzes Focused on Quiz Types, Proceeding from ICCE2011, *The 19th International Conference on Computers in Education*, Chiang Mai, Thailand, 312-316.

[9] Tsumori, S. Kaijiri, K. (2009). A method for automatic generation of multiple-choice questions adapted to student's understanding. *The Journal of Information and Systems in Education* 26(3), 240-251. (in Japanese)

[10] Takagi, T., Takagi, M. & Teshigawara, Y. (2009). A Proposal and Evaluation of a Method of Calculating Similarity between Quizzes Created by Students. *Journal of IPSJ*, 50(10), 2426-2439. (in Japanese)

[11] Ayala, R. J., (2009). *The Theory and Practice of Item Response Theory*. (New York, Guilford).

[12] Takagi, T., Takagi, M. & Teshigawara, Y. (2008). A Proposal and Evaluation of a Similar Level Calculation Method by Type of Quizzes Created by Students. *Proc. IPSJ Summer Symposium in Setouchi*, 95-102. (in Japanese)

[13] Ikeda, S., Takagi, T., Takagi, M. & Teshigawara, Y. (2011). A Study on a Method of Estimating the Difficulty of Items Focused on Item Types and Similarity of Choices. *Proc. IPSJ Summer Symposium in Setouchi*, 133-139. (in Japanese)

A Systematic Review of Methods for Developing Open Educational Resources

Maurício Massaru ARIMOTO[a*], Ellen Francine BARBOSA[a*]
[a]*University of São Paulo (ICMC/USP}, Brazil*
*{mauricio, francine}@icmc.usp.br

Abstract: Educational resources, such as learning objects, have been more and more used in education and training context. The free and open distribution of these resources contributes to the dissemination of knowledge and facilitates access to information. Following this trend, Open Educational Resources (OER's) have emerged to assist in the teaching and learning processes in general. Motivated by this scenario, the purpose of this paper is to characterize the state-of-the-art regarding the development, delivery and reuse of OER's. A systematic literature review was conducted and some initiatives were identified and investigated. Additionally, a preliminary set of characteristics to be considered in the development of OER's was also established. In general, we noticed a lack of systematic methods for the appropriate creation and adoption of OER's. In this sense, this work provides guidance for new research and development in the area.

Keywords: Open educational resources, development methods, systematic review.

Introduction

The advent of the Internet and advances in Information and Communication Technologies (ICT's) has changed the educational context, both in traditional as well as in blended and distance learning. As a result, there has been a change in the way that educational content is designed, developed and delivered to learners. Faced with these transformations, in recent years education and training issues have been attracting more and more interest from researchers around the world.

Learning objects (LO's) have emerged as interesting alternatives in this context. In short, a LO can be characterized as a reusable digital content used as educational support. The main idea is to allow the educational content be "broken into small pieces", which can be reused in different contexts and scenarios of education and training [33].

In many aspects of development, the production of LO's is similar to software development. In the case of software, methods, procedures and tools have been established aiming at contributing to the development of quality products [24]. Similarly, the use of appropriate mechanisms to ensure the productivity of the development process and the quality of the resultant products are also critical with respect to LO's.

Agile methods [10] fit in this context, addressing a new approach to development, focusing on agility, flexibility, skills to communication and the ability to deliver new product and services with added values to the market and in defined time [2],[3].

In a different but related perspective, the advent of Free Libre and Open Source Software (FLOSS) [15] has also motivated research and development in the education area. In 1998, Wiley [33] introduced the concept of "open content" to refer to all content available to its users in an open way, creating the Open Content License and later the

Open Publication License for the publication of these contents. The idea was to encourage debate and the availability of open educational content by institutions of higher education.

Recently, there arose the term *Open Educational Resources* (OER’s) in an attempt to standardize the educational content available in a free and open way through the Internet. The OER's are characterized as digital materials available in a free and open way to the general community, with the purpose of teaching, learning and research. Basically, an OER encompasses: (1) *learning resources*, such as LO's, full courses and educational modules; (2) *tools*, such as supporting systems to the development, (re)use and delivery of learning content; and (3) *implementation resources*, such as intellectual property licenses to promote the publication, reuse and dissemination of the educational content [16].

Motivated by this scenario, this paper aims at characterizing the state-of-the-art of methods for developing OER’s. A systematic review was conducted and, in order to compare and evaluate the studies found through this process, a set of characteristics for the development of OER’s was proposed. The results presented herein will underlie the establishment of systematic methods for the development of quality and reliable OER’s.

The remainder of this paper is organized as follows. Section 1 outlines the systematic review planning, defining the review protocol for the study; the review execution is also described. The results obtained are presented and analyzed in Section 2. Finally, in Section 3, our conclusions and perspectives for further work are summarized.

1. Systematic Review Planning and Execution

A Systematic Literature Review can be seen as the process of identification, evaluation, and interpretation of available and relevant researches for a research question, thematic area or interest phenomenon. The aim is to provide a fair assessment (not biased) for a research topic, by an auditable, reliable and accurate approach [21].

The systematic review process begins with the planning phase through a pre-defined review protocol, which includes: (1) the research objectives; (2) the research questions to be answered; (3) the search strategy for conducting the review; and (4) the criteria and procedures for the studies selection.

1.1 Research Objectives

Our systematic review aimed at characterizing the methods that support the learning objects development, focusing on OER’s. Additionally, agile methods were also considered. To clarify, the objectives were subdivided into: (1) investigate and analyze methods that support the development of LO's; (2) investigate and analyze the applicability of agile methods in the development of LO's; and (3) investigate and analyze methods that support the development and delivery of OER's.

1.2 Research Questions

From the aforementioned objectives, the following research questions were defined: (1) *Primary question*: Which methods have been used in the development of LO's? (2) *Secondary questions*: Which methods are specifically related to agile methods? and which methods are specifically related to the development and delivery of OER's?

The research questions were grounded in different perspectives, being included within the following scope and specificities: (1) *intervention*: development of OER’s; (2) *population*: researches that cover the development of LO’s (and OER’s); and (3) *results*: methods to support the development of LO's, especially OER’s.

1.3 Search Strategy

The search strategy consists in defining: (1) the search sources for selecting the studies; (2) the language of the studies; and (3) the terms and synonyms for constructing the search string. The sources were selected taking into account criteria such as: the importance and relevance of search sources, the availability of search query through the web, the number of indexed researches, and the reliability of results. The selected sources are shown in Table 1, including electronic and manual databases (conference proceedings, periodicals, technical reports, etc.) as well as consulting specialists and researchers related to the area.

Table 1: Search sources

Source	**Location**
ACM Digital Library	http://portal.acm.org/dl.cfm
IEEE Xplore Digital Library	http://ieeexplore.ieee.org
Scopus	http://www.scopus.com
Web of Knowledge (Web of Science)	http://apps.webofknowledge.com
CBIE – Brazilian Congress on Computer Science Education	http://www.sbie.org.br
RBIE – Journal of Computing in Education	http://www.rbie.org.br

A general search string was built from the combination of terms and synonyms associated to educational resources in general, using boolean operators (AND/OR). When necessary, the search string was analyzed and refined according to the needs and characteristics of each database. This procedure was aimed at evaluating the relevance of the terms used and the relevance of the studies returned.

1.4 Studies Selection

Inclusion and exclusion criteria had also to be explicitly defined in order to return coherent and consistent results for the research. Inclusion criteria were: (1) *Primary Question*: studies describing methods to support the development of LO's. (2) *Secondary Questions*: (2a) studies applying agile methods in the development of LO's; and (2b) studies addressing the development and delivery of OER's.

Exclusion criteria were: (1) *Primary Question*: studies describing approaches applied in other areas. (2) *Secondary Questions*: (2a) studies applying agile methods in other contexts; (2b) studies addressing open content outside the educational context; and (2c) studies that were not fully available for reading.

The studies selection was performed in two phases. In first phase, *preliminary selection*, the emphasis was on reading the abstracts of the studies returned by the search string. In the second phase, *final selection*, the studies were read in full by the reviewer.

The systematic review was conducted during February – April/2012. While executing the searches, the results were documented for further analysis. All the data extracted was summarized and documented. To facilitate the management of the references, we used the JabRef tool [20]; the tool supports the prioritization of the readings, the retrieval of studies for reading, and the identification of duplicate studies. We also used ReVis [31], a supporting tool for the visual exploration of the studies collected.

2. Results and Discussion

The number of the studies selected in each search source, both in the preliminary and in the final phases, is shown in Table 3. The systematic review returned 240 papers, discarding duplicated studies. In the preliminary stage, 45 papers were pre-selected for full reading. In the end, 14 papers were considered relevant to the aims of the review.

Table 2. Studies selected in each database

Database	**Result**	**Preliminary Selection**		**Final Selection**	
		Included	Excluded	Included	Excluded
ACM Digital Library	45	8	37	2	6
IEEE Xplorer Digital Library	64	12	52	6	6
Scopus	78	10	68	1	9
Web of Knowledge (Web of Science)	46	10	36	1	9
CBIE and RBIE	4	3	1	2	1
Specialists	3	2	1	2	-
Total	240	45	195	14	31

Figure 1 illustrates the bi-dimensional maps of the studies collected, constructed from the ReVis tool [31]. The bi-dimensional maps distribute the studies collected over the window space, grouping the most similar studies and isolating the most distinguished. This distribution is made according to the contents of the study, i.e., a combination of titles, keywords and abstracts. Figure 1(a) shows the studies selected in the preliminary phase, represented by small circles colored in green; the red circles represent the studies that were excluded. Figure 1(b) highlights the studies selected in the final phase.

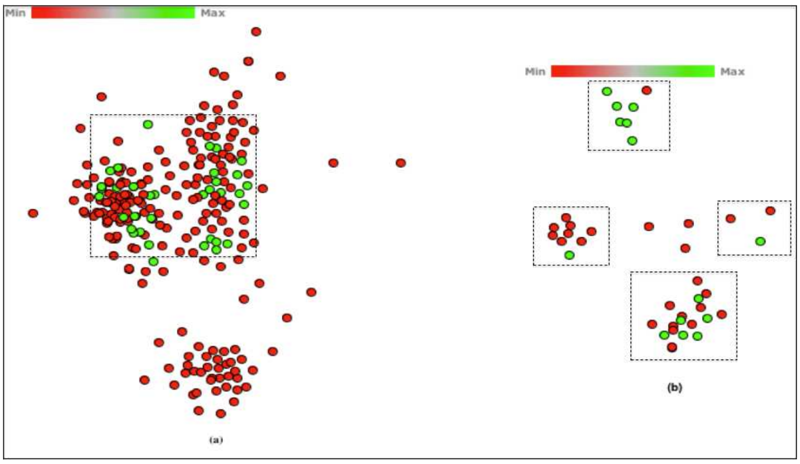


Figure 1. Visual maps of collected studies

A synthesis of the selected studies, along with the inclusion criteria used, is presented in Table 4.

In order to compare and evaluate the studies found through the systematic review, we have also defined a preliminary set of characteristics considered relevant from the perspective of the development of OER's (Table 5).

The results obtained with the evaluation of the studies are summarized in Table 6. From the perspective of methodology (foundation), Boyle et al. [13] propose a method for the development of LO's based on agile practices, aiming at providing a robust and flexible structure that contributes to the creation of quality educational content. Lapolli et. al. [23] propose a model of instructional design for LO's grounded in assumptions of agile methods, specifically in the technique of Behavior-Driven Development (BDD) [9] and in eXtreme Programming (XP).

Barbosa and Maldonado [6] have proposed a Standard Process for Developing Educational Modules (SP-DEM) based on the ISO/IEC 12207 standard [19] and on the ADDIE model [32]. The SP-DEM aims at establishing guidelines for development and systematic evolution of educational modules. Patricia et. al. [28] also uses the ADDIE model as the basis for a life cycle for OER's; activities to incorporate social networking and semantic technologies were also considered. Finally, Oliveira et al. [27] have proposed a life cycle for the elaboration of LO's based on the spiral model [30] of software development.

Table 4. Focus of selected studies

Author	Main Goal
[4]	Proposition of a process model for Web-based educational modules.
[22]	Development of a LO for teaching network technologies. Design and development principles of the LO are discussed.
[6]	Proposition of a standard process for the elaboration of educational modules based on ISO / IEC 12207.
[5]	Development of an educational module for teaching mutation testing according to a content modeling approach.
[27]	Proposition of a life cycle for the elaboration of LO's based in the spiral model.
[7]	Development of an education module for teaching inspection and testing techniques.
[8]	Development of an educational module for teaching fundamentals of programming and testing. Proposition of an automated tool for evaluating programming assignments based on testing activities.
[29]	Proposition of a process for the development of LO's.
[11]	Elaboration of LO for Learning Virtual Community (COMVIA).
[13]	Proposition of a method for developing LO's based on agile methods.
[23]	Proposition of an instructional model of LO's based on interaction design and agile practices.
[28]	Proposition of a life cycle for OER's based on social tools and web semantics.
[14]	Proposition of a model of LO's for online learning based in the European Higher Education Area (LOMOLEHEA).
[12]	Proposition of a supporting tool for modeling educational content.

Table 5: Set of characteristic

Characteristic	Description
Methodology	To ensure that OER's effectively contribute to the teaching and learning process, it is necessary to use appropriate approaches and methodologies that support design and developmental tasks and activities.
Standards	The use of standards for metadata (such as Learning Object Metadata (LOM) [17] and for packaging (such IMS Content Packaging [18]) ensures the availability of OER's in different Learning Repositories (LR). Besides that, standards facilitate the search, retrieval and reuse of the educational content.
Learning repositories	OER's should be available through LCMS or LR in order to provide reuse.
Collaborative and distributed development	The elaboration of OER's may involve developers from diverse areas of knowledge, working in multidisciplinary teams and heterogeneous, geographically dispersed or not. In this case, it is necessary to consider aspects of collaborative and distributed development as the systematic monitoring of activities and modules that compose the resources, and the support for communication among the teams.
Web 2.0 and semantic web technologies	The integration of social tools encourages the active participation of developers and users in the construction of OER's, being also important in distributed and collaborative development of such resources. Moreover, semantic web technology improves searches of educational resources by using their semantic meaning.
Environments and tools	The use of environments and tools to support the construction of OER's tend to facilitate the development tasks, besides contributing to the quality of the final product.
Licenses	To preserve the authorship rights and intellectual properties, the free and open distribution of an educational resource must take place under the appropriate license.

The IMA-CID approach (Integrated Modeling Approach – Conceptual, Instructional and Didactic) is used as a basis for developing educational resources in [5], [8] and [12]. Furthermore, in [7] and [8] the authors also use the SP-DEM process.

Table 6. Comparison of the methods considered

Author	*Methodology (Foundation)*	*Standard*		*LR*	*Collaborative /Distributed Development*	*Web 2.0 and Semantic Web*	*Environment and Supporting Tool*	*License*
		Metadata	*Packaging*					
[27]	Spiral Model	LOM	SCORM					
[29]		LOM	SCORM					
[6]	ISO/IEC 12207, ADDIE Model	LOM						CCL
[13]	Agile Methods	LOM/IMS Metadata		√	√			
[23]	Agile Methods		SCORM					
[28]	ADDIE Model	LOM/IMS Metadata	SCORM	√		√		CCL
[10]				√	√			
[14]		LOM	SCORM					
[22]					√			
[4]								
[5]	IMA-CID	LOM						
[7]	SP-DEM	LOM						
[8]	SP-DEM, IMA-CID	LOM						
[12]	IMA-CID	LOM						

The other studies ([4], [10], [14], [29]) did not specify any method for the development of educational resources.

Considering the use of standards, most of the studies adopted LOM for metadata and SCORM for content packaging and associated metadata. On the other hand, [4] and [22] did not use any standard for metadata and packaging, while [5], [6], [7], [8] and [12] did not adopt any standard only for content packaging.

Regarding learning repositories (LR), in general they are rarely explored by in the storage and retrieval of educational resources, which hinders the dissemination and reuse of such resources. Among the methods found, only a few ([10], [13], [28], [29]) addresses educational resources for institutionalized repositories. Likewise, aspects related to the collaborative and distributed development are almost unexplored, being investigated only by [6] [10], [13], [22]. In the case of OER's, the use of collaborative technologies (such as e-mail, mailing lists, web, social tools, versions control systems, information repositories, etc.) is essential not only in the development process, but also during the delivery and use of the content itself in order to facilitate the conducting the activities and assessments proposed to learners.

Little attention was also given to the adoption of emerging technologies, such as Web 2.0 and semantic web. Despite the great potential to effectively contribute to the development, delivery and search of OER's, such technologies were considered only in the method proposed by Patricia et al. [28].

Another limitation observed refers to the lack of licenses to establish the authoring rights and intellectual properties on the content developed as well as the details for its use, modification and distribution. Only the methods proposed by [5] and [6] specify the use of licenses; in both cases, the Creative Commons License (CCL).

Finally (and interestingly), none of the methods analyzed specify the adoption of computational tools and environments in order to support the automation of the activities and tasks associated.

4. Conclusions and Further Work

In this paper we presented a systematic literature review aiming at identifying the state-of-the-art regarding the development, delivery and reuse of open educational resources. Additionally to the review conducted, we also highlight as a contribution of this paper the establishment of a set of characteristics to be considered in the development of OER's.

In general, we have identified different methods being used, ranging from the modeling phase and instructional design to the development of associated educational content. However, despite OER's have emerged as a viable technology to the social-economic development, we highlight the lack of systematic methods for their appropriate creation and adoption. In this sense, this work intends to provide guidance for new research and development in the area.

The limitations and weaknesses observed motivate the study and proposition of systematic methods for the development, delivery and reuse of flexible, quality and reliable OER's. Characteristics such as openness, collaboration, cooperation and distribution should also be considered in this perspective. At the same time, there is a need for the establishment and adoption of automated tools and environments to support the related activities.

As further work, based on the results obtained, we intend to work on the establishment of systematic methods for developing and providing quality educational resources, capable of motivating learners and contributing to their process of knowledge construction. Aspects related to the establishment of supporting environments for the OER's creation must also be defined and incorporated into the proposed method. Among the expected results, we emphasize the development of a pilot environment for the construction, storage and retrieval of open educational resources its validation in the creation of OER's for different knowledge areas.

Acknowledgements

The authors would like to thank Brazilian funding agencies (FAPESP, CNPq and CAPES) for their financial support.

References

[1] Advanced Distributed Learning. (2004). Sharable Content Object Reference Model – SCORM. 1. ed.

[2] Arimoto, M. M., Murakami, E., Camargo, V. V., & Cagnin, M. I. (2009). Adherence Analysis of Agile Methods According to the MR-MPS Reference MOdel. In: VIII Simpósio Brasileiro de Qualidade de Software (SBQS'09) (pp. 249-263), Ouro Preto, MG.

[3] Arimoto, M. M., Murakami, E., Camargo, V. V., & Cagnin, M. I. (2010). ADAMA: a Pattern Language for Adherence Analysis and Adaptation of Agile Methods to MR-MPS Reference Model. In: VIII Latin-American Conference on Pattern Language of Programming (p. 14), Salvador, BA.

[4] Ammar, H. H., Mohammed, A. A., & Abdalla, O. (1997). Process-centered Environments for Web-based Instructional Modules. In: 27th Annual Frontiers in Education Conference (327-332).

[5] Barbosa, E., & Maldonado, J. (2006a). Establishing a mutation testing educational module based on IMA-CID. In: II Workshop on Mutation Analysis (p. 14).

[6] Barbosa, E. F., & Maldonado, J. C. (2006c). Towards the Establishment of a Standard Process for Developing Educational Modules. In: 36th Annual Frontiers in Education Conference (FIE'06) (pp. 1–6), San Diego, CA.
[7] Barbosa, E. F., & Maldonado, J. C. (2008). Specialization and Instantiation Aspects of a Standard Process for Developing Educational Modules. Proceedings of ISOLA'08 (1-16), Kassandra, Greece.
[8] Barbosa, E. F., Souza, S. R. S., & Maldonado, J. C. (2008). An Experience on Applying Learning Mechanisms for Teaching Inspection and Software Testing. Proceedings of CSEET'08 (189-196), Charleston, SC.
[9] Beck, K. (2003). Test Drive Development: by Example. Addison-Wesley Professional.
[10] Beck, K.; Andres, M. Extreme Programming Explained: embrace change. 2 ed. Addison-Wesley, 2004.
[11] Behar, P. A., Frozi, C. S. A. P., & Bernardi, M. (2008). Virtual Learning Communities: a learning object integrated into e-learning plataform. IN: Kendall, M. and Samways, B., ed. (pp. 163-166).
[12] Borges, V., Maldonado, J. C., & Barbosa, E. F. Towards the Establishment of Supporting Mechanisms for Modeling and Generating Educational Content. (2011). *Proceedings of the ACM SAC'11* (1202-12-7), NY, USA.
[13] Boyle, T., Cook, J., Windle, R., Wharrad, H., Jeeder, D., Alton, R. (2006). An Agile Method for Developing Learning Objects, *23rd Annual Ascilite Conference: Who's Learning? Whose Technology* (pp. 189-196). Sydney, Australian.
[14] Cocón, F., & Fernández, E. (2010). LOMOLEHEA: (Learning object model for online learning based on the European higher education area. *Proceedings of ITiCSE'10* (pp. 78-82), New York, USA.
[15] Free Software Foundation. (2012). http://www.fsf.org [03/2012].
[16] Hylén, J. (2006). Open Educational Resources: opportunities and challenges. In: Open Education 2006: Community, Culture, and Content. Utah State University, Logan UT: Centre for Educational Research and Innovation (pp. 49-63).
[17] IEEE Learning Technology Standards Committee. (2002). Learning Object Metadata (LOM). *Draft Standard 1484.121.1*. IEEE Press.
[18] IMS Global Learning Consortium. (2007). http://www.imsglobal.org/content/packaging/ [04/2012].
[19] International Organization for Standardization Systems and Software Engineering – Software Life Cycle Processes. (2008). ISO/IEC 12207.
[20] JabRef Reference Manager. (2012). JabRef version 2.8. http://jabref.sourceforge.net/ [02/2012].
[21] Kitchenham, B. (2007). Guidelines for Performing Systematic Literature Reviews in Software Engineering. Technical Report EBSE-2007-01, Keele University and University of Durham.
[22] Klobas, J., Giordano, S., Renzi, S., Sementina, C. (2004). Multidisciplinary Learning Objects: technology and pedagogy. *Proceedings of ICALT'04* (470-474), Joensuu, Finland.
[23] Lapolli, F., Motta, C. L. R., Cruz, C. M., Oliveira, C. E. T. (2010). Modelo de Desenvolvimento de Objetos de Aprendizagem Baseado em Metodologias Ágeis e Scaffoldings. Revista Brasileira de Informática na Educação, 18(2), 41-53. In Portuguese.
[24] Myers, G. J. (2004). The Art of Software Testing. 2 ed. New York: John Wiley & Sons.
[25] National Information Standards Organization. The Dublin Core Metadata Element Set. NISO PRESS. n. ANSI/NISO Z39.85-2007.
[26] OECD. (2007). Giving Knowledge for Free. The Emergence of Open Educational Resouces. Centre for Educational Research and Innovation (CERI).
[27] Oliveira, E. R., Nelson, M. A. V., Ishitani, L. (2007). Ciclo de Vida de Objetos de Aprendizagem baseado no Padrão SCORM, XVIII Simpósio Brasileiro de Informática na Educação (SBIE'07), Mackenzie, São Paulo. In Portuguese.
[28] Patricia, C. C. S., Rocio, R. M. G. D., Elizabeth, R. P. A. (2010). OER's Production Cycle with Social Authorship and Semantic Tools, *IEEE Education Engineering Conference* (pp. 121-128).
[29] Pessoa, M. C., & Benitti, F. B. V. (2008). Proposta de um Processo para Produção de Objetos de Aprendizagem, *Hífen* 32(62), 172-180. In Portuguese.
[30] Pressman, R. S. (2010). Software Engineering: a practitioner's approach. NY: McGraw-Hill Higher Education.
[31] Revis Tool. (2010). Systematic Review Supported by Visual Analytics. University of São Paulo, ICMC-USP.
[32] Strickland, A. W. (2006). ADDIE. Idaho State University College of Education, Science. Math & Technology Education (Retrieved 2006-06-29).
[33] Wiley, D. A. (2002). *Connecting Learning Objects to Instructional Design Theory: a definition, a metaphor, and a taxonomy.* Bloomington: Association for Educational Communications and Technology.
[34] Wiley, D. A. Openness as Catalyst for an Educational Reformation. EDUCAUSE, v. 45, n. 4, p. 14–20, july/august, 2010.

Designing a Preliminary E-Textbook with the Information Model

Zhe XU, Haibo LI & Xiaoqing GU
Shool of Education Science, East China Normal University, China
xuzhe_xz@163.com

Abstract: Recently, ISO/JTC1 SC36 has initiated a new standard project on e-Textbook, which was proposed by the authors' team. In this project, the information model of e-Textbook is the most fundamental part and further studies all will be based on it. The purpose of this paper is to manifest the information model of the e-Textbook by describing a practical case in our ongoing work. The process of designing an English e-Textbook for grade 6 is presented to explain the main steps of element structure design and function structure design. The information model and its practice reflect what kind of e-Textbook we need and how to design a standard e-Textbook for discussing the disorder of e-Textbook's development.

Keywords: e-Textbook, information model, element structure, function structure, case

Introduction

Recent several years witness the fast adoption of e-Textbook around the world. In China, Shanghai, Beijing,Guangzhou,Shenzhen, and many other cities start e-Textbook programs and practice in middle and primary schools. The similar actions have been taken in other countries. However, as a new research area, the development of e-Textbook seems a bit disorder in technical aspect and educational aspect (Gu & Fu, 2012).So in previous work, we have worked on the e-Textbook information model in order to provide solution. In this paper, we introduce the on-going standard project of e-Textbook informational model by presenting a case that is being developed based on this model.

1. Textbook design based on e-Textbook information model

1.1 Description of e-Textbook information model

The information model defines content structure and function structure of an e-Textbook. The content structure defines content elements which constitute content of an e-Textbook. Content categories are units and aggregations; units and aggregations can be assembled into learning content according to specific learning needs (Figure 1).The function structure defines the operation and its corresponding results thus cause. There are three kinds of operations: basic operations, connecting operations, and comprehensive operations (Figure 2).

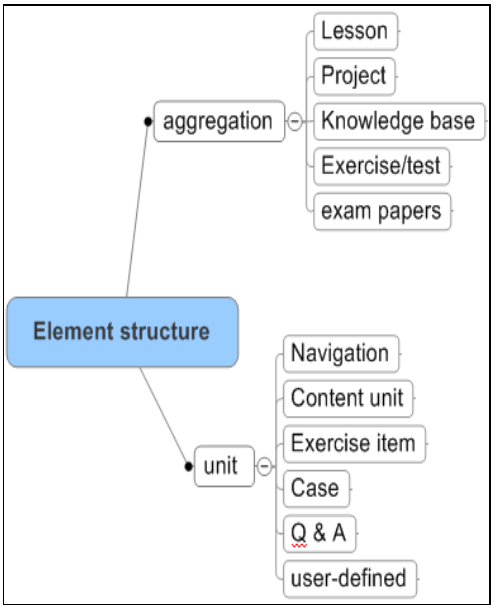


Figure1 element structure

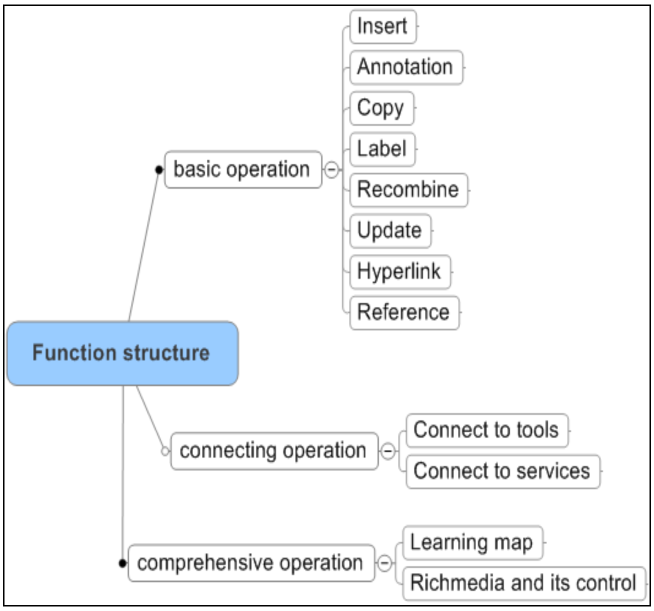


Figure2 function structure

1.2 A preliminary e-Textbook case

In this part, the process of developing a preliminary e-Textbook was introduced by a case undergoing. The case e-Textbook is Oxford English 6A published by Shanghai Education Press for 6 grades.

1.2.1 The design of element structure

Based on the information model, the first step in designing an e-Textbook is to structure the content of traditional textbook, into units and aggregations, which can be reused and shared across the curriculum. The case of structured content can be manifested from Unit4 of the said textbook (Figure 3).

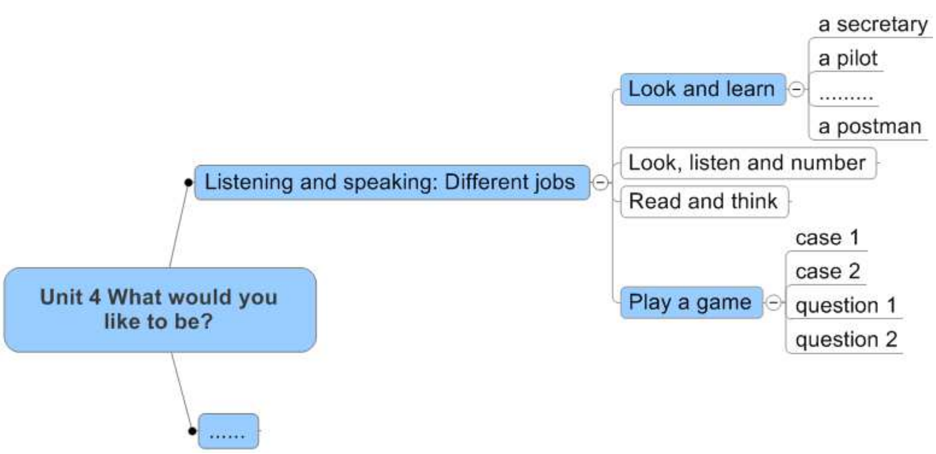


Figure 3 content design

These units and content aggregations can be assembled just like building blocks. Based on content design, the corresponding content structure is presented (Figure 4).

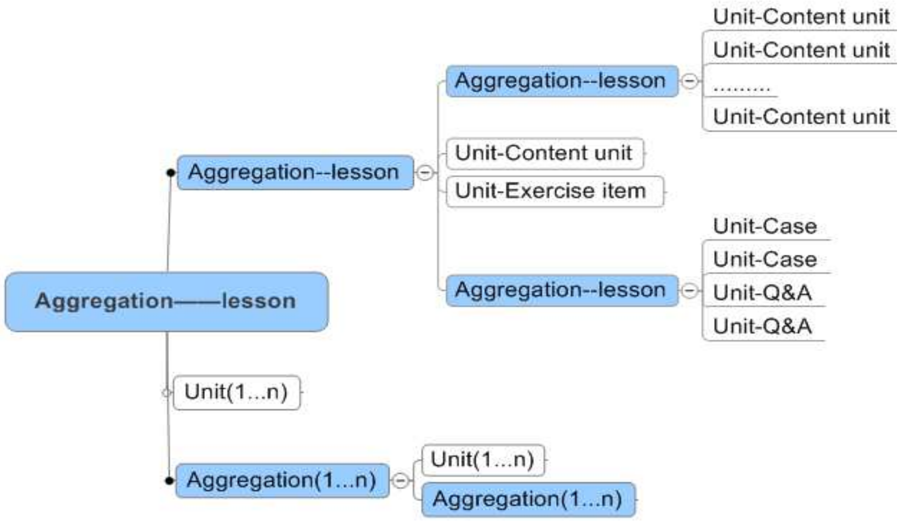


Figure 4 content structure

1.2.2 The design of function structure

Based on the function structure (Figure 2) of the information model, the functions of the preliminary e-Textbook have been designed as listed in Table 1.

Table 1 English e-Textbook's function structure

	functions	Specific items of functions for English e-Textbook
Basic operations	Insert, Annotation, Copy;Label;Update;Recombine;Hyperlink;Reference	Not all the unit/ aggregation must have all the kinds of basic operation. Other user-defined operations are allowed too.
Connecting operations	Connect to tools Connect to services	Notepad(a local record tool) Recorder(a local record tool) Google(for online search) LINO(an online sticker) Midomo(an online mind map) YouDao Cloud Note(an online notepad) YouDao Online Dict (an online dictionary) Online Exam Library(for online test) Other tools and services(user-configured)
Comprehensive operations	Learning map Richmedia and its control	recording users' personal and dynamic learning activities(in developing) English situational dialogues using richmedia Other operations

2. Analysis of the e-Textbook with information model

2.1 Analysis of the case's element structure

As showed in Figure 5, section of "Look and learn" is a content aggregation composed of content units. Each word in this aggregation is a content unit with independent knowledge point, consisting of word interpretation, phonetic symbol, pronunciation, example sentences and other knowledge. And a collection of words' list forms an aggregation for words learning.

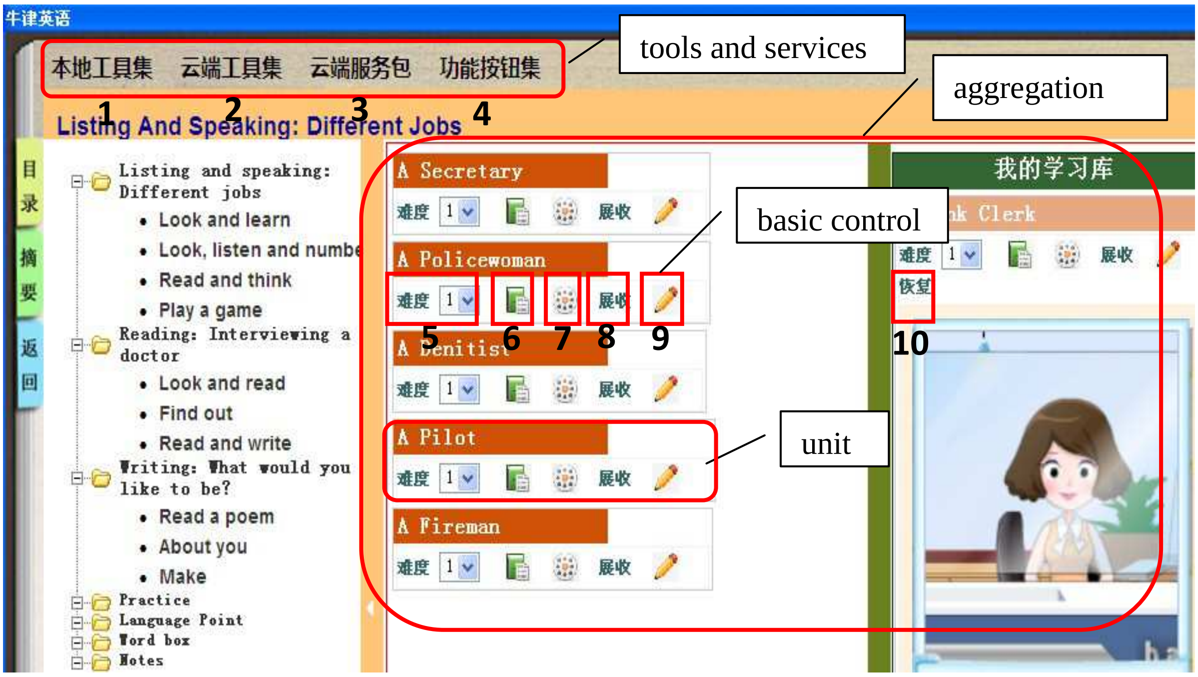


Figure 5 e-textbook based on element structure and function structure

2.2 Analysis of the case's function structure

According to English e-Textbook's function structure and the characteristic of operations, four function modules (marked by numbers in Figure 5)are developed in function area, referring "1 tools set in local", "2 tools set in Cloud", "3 services package in Cloud" and "4 Other functions package". Such tools and services can be applied for all kinds of learning contents. And some basic operations are imposed on learning contents. In the learning process, learners can use these operations to meet different needs in different context. The specific tools and services are deployed in the four modules, such as Recorder, YouDao Cloud Note, etc.

Moreover, some basic operations are designed for a unit to supporting in-depth learning. In this case, six kinds of operations are designed to meet different learning needs for each word, and a Personal Learning Library is designed to support personal learning. We analyze them with function structure (Table 2).Students can mark levels of difficulty and make notes and can also remove some words to Personal Learning Library according to their learning. Each learner can have own word library.

Table2 operations for specific content

Operations in interface	Operations with function structure	Description
5 难度 1 difficulty set		Students set the difficulty level for a word.
6 unfold /fold details of word	Copy	Click it and a new window with more information about a word will be open. Students can copy information.
7 open /close media	Richmedia and its control	Click it ,a new window with rich media will be play, and students can control it
8 展收 display/hide picture		Click it and picture corresponding to a word will be display
9 notes	Insert;Annotation;Copy; Label; Update	Click it and a new editable window will be open. Some basic operations are allowed.
10 恢复 move into Personal Learning Library /put back	Recombine	Word can be move into Personal Learning Library to construct own word list for simple recombination.

3. Conclusion and future work

The information model and its practice reflect what kind of e-Textbook we need and how to design a standard e-Textbook for discussing the disorder of e-Textbook's development. This paper focuses on how to design a standardized e-Textbook with information model and two important steps are elaborated by a practical case. We believe e-Textbook based on information model can manifest the significant features of e-Textbook, such as relevance, interactivity, openness, richmedia, content interoperability (Zhu & Yu, 2011). This is essential for the e-Textbook to play an important role in education, which has not yet at present. With the progress of ISO/JTC1 SC36 e-Textbook standard, more work needs to be done either to finish the preliminary case introduce here, or to conduct experiments with e-Textbooks in schools.

References

[1] GU xiaoqing,&Fu wei. (2012).Connecting Reading and Learning: Design the Information Model of E-Textbook. *JOURNAL OF CHINA NORMAL UNIVERSITY (NATURAL SCIENCES)*, 02: 81-90.

[2] ZHU zhiting, &YU xiaohua.(2011).E-Schoolbook System and Its Function Modeling. *E-education Research*, 4:27-34

An Application of Eye-tracking Technology to Fashion Design Education: A Pilot Study on How Women Gaze at Handbags

Hong-Fa HO[a*], Chi-Wu HUANG[b], Yen-Lin HUANG[a] & Wei-Fan KAO[a]
[a]*Department of Applied Electronics Technology,*
National Taiwan Normal University, Taiwan, R.O.C.
[b]*Department of Industrial Education,*
National Taiwan Normal University, Taiwan, R.O.C.
[*]*jackho@ntnu.edu.tw*

Abstract: Questionnaire is a conventional approach for fashion designers and educators to survey consumers' general preferences of products' appearances; however the precision of survey results remains questionable. The authors proposed an alternative in this pilot study in which five eye trackers were used to record 10 female participants' eye movement data of gazing at 20 randomly displayed handbag pictures. Six types of Regions of Interest (ROI) of handbags were assigned to understand which regions were more attractive to the participants. The times and duration of fixations showed the degree of ROIs' attractiveness to the participants. The findings provide persuasive eye-movement evidence of predicting potential female consumers' gazing behaviors. Textbook editors of fashion design can also add eye-tracking approach to future books. And eye-tracking technology can also help fashion design educators to enhance teaching and students' learning. Moreover gazing behavior can be broadly considered a form of reading, and the findings of this study further extend the existing knowledge of human reading process.

Keywords: Fashion design education, eye tracking technology, women's handbags, gaze pattern

1. Introduction

Fashion design education aims to foster future designers, and understanding consumers' visual preferences is important to teaching and learning of this field. Still it is challenging for educators and designers to precisely predict consumers' visual preferences by using conventional questionnaire survey. Eye-tracking approach proposed in this paper can help them can develop better understanding of consumers' general gazing behaviors so that they can enhance their teaching, students' learning and designing.

Eye tracker is an instrument designed to record eye movement data to be used to analyze process of reading a written text or behavior of gazing at an object, for instance. Researchers can identify viewers' degrees of visual attractiveness to a material or the crux that causes readers reading difficulty based on their eye movement data such as frequency of gazing at particular Regions of Interest (ROIs) of certain materials and the duration of fixations on the regions. Eye-tracking technology has been applied to various fields such as language teaching and learning, neurosciences, psychology, engineering, marketing, computer sciences, and others [3], [6], [7], [8].

In literature survey, some researchers used eye trackers to study consumers' gazing behavior of fashion products [4], [5]. However, by far, no literature has specifically documented how women gaze at women's handbags by taking an eye-tracking approach. This approach is applicable and time-efficient, and it is able to provide valuable references to both fashion design educators and designers.

This pilot study aims to look for women's general gaze patterns of handbags instead of studying the differences and correlations between various styles and degrees of attractiveness. In the experiment the authors recorded and analyzed 10 female participants' eye movement data of gazing at 20 randomly displayed pictures of women's handbags. The authors defined six types of ROIs of a handbag on the basis the structure of a handbag: handle (denoted as R_1), main body (R_2), hanging ornament (R_3), strap (R_4), featured area (e.g. a larger-sized metal part) (R_5), and textual information (e.g. brand's name, logo figure or trademark) (R_6). Figure 1 is an example of a handbag's ROIs.

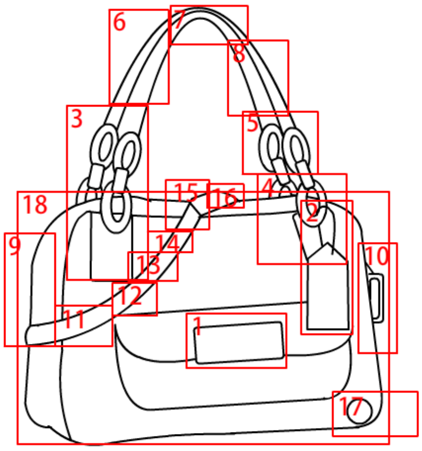


Figure 1. An example of a women's handbag with different ROIs

The research questions are list below.

Q_1: What is the gaze sequence when the participants gaze at the images of handbags?

Q_2: What is the sequence of the capacities of ROIs to adhere attention?

Q_3: What is the sequence of degrees of attention to ROIs?

Before this pilot study, the results of a preliminary experiment were used to propose the following hypotheses.

H_1: The gaze sequence is R_2, R_3, R_6, R_5, R_1, and R_4.

H_2: The sequence of the capacities of ROIs to adhere attention is R_5, R_1, R_4, R_3, R_2, and R_6.

H_3: The sequence of degrees of attraction to ROIs is R_5, R_2, R_3, R_1, R_6, and R_4.

2. Method

2.1 Participants and Procedure

The participants were 10 voluntary female undergraduate students of National Taiwan Normal University, Taipei, Taiwan (20.75 years old on average, SD=1.7 years, range 19-22 years), and they all signed consents of participating in this experiment. The participants were not informed the purpose of this research before the experiment in order to observe their natural gazing behaviors. Instead, they were told that the aim of this experiment was to measure the expansion of pupillary responses. A self-developed eye-tracking instrument named EyeNTNU-180 was used to collect the participants' eye movement data from their left eyes. When the experiment was completed, the authors told them the true purpose of this experiment and asked them if they had ever questioned the purpose of the experiment. None of them questioned the purpose.

The 20 selected pictures had six types of ROIs. The distance from the participants' eyes to the screen was 60 centimeters. The authors used self-developed ROI Tool to define ROIs of the 20 pictures and another self-developed Fixation Calculator was used to process the participants' eye movement data.

2.2 Stimulus Material

The authors collected 80 pictures of women's handbags from online stores' webpages. Each picture only had one handbag without background colors or patterns. Some pictures had brands on it; some did not. To remove pictures that include any distracting elements, the authors carefully selected 20 pictures with good image quality (10 pictures with textual information, brands, or trademarks on the handbags). The pictures were placed in the same orientation and image sizes of the pictures were very similar.

The definitions of six types of ROIs are illustrated below. The bag's main body was defined as a common ROI of 20 pictures; 10 handbags' handles were defined as a ROI type; 10 handbags' textual information was defined as a ROI type; 10 handbags' hanging ornaments were defined as a ROI type; 10 handbags' straps were defined as a ROI type; and 10 handbags' featured areas were defined as a ROI type. The 20 pictures were randomly displayed and they were all centered on the computer screen. The participants were allowed to decide how much time they would like to spend on gazing at each picture. They were asked to start from the centers of the pictures. They were also asked to calibrate the eye trackers before they view every picture to ensure the accuracy of their eye movement data. Finally the authors asked them to fill out questionnaires in order to collect their basic background information, age or majors, for instance.

2.3 Eye Tracking: Recording and Quantification

Five EyeNTNU-180s were simultaneously used to record the participants' eye movement data. The sampling rate was 180 Hz. Two infrared LEDs provided lighting that reflected on dark pupil to measure the participants' eye movements. The outputting energy of the LEDs was 3.5 mW/cm^2 at a working distance of four centimeters. The participants were asked to do nine-point calibration by tracing nine white moving points on the screen (the background color was black) before the experiment. The participants' eye movement data was collected every 5.56ms from their left eyes as they gazed at the working area on the screen. No data would be recorded when the participants blinked.

The ROI Tool was used to define ROIs of every picture. Since the sizes of ROIs were not the same, the sizes of pictures were adjusted to 80%-100% in proportion to the height of the screen. When the participants gazed at certain spots for at least 80ms, one time of fixation would be counted into the total number of fixations. Fixation Calculator was used to calculate the following parameters of each participant's eye movement data:

(a) the dwell time: total contact time of all fixations of gazing at each ROI,
(b) the number of fixations on each ROI,
(c) the latency of first fixation on each ROI, and
(d) the duration of the first fixation on each ROI.

The authors considered ROI as a factor and used descriptive statistics for data analysis.

3. Results

3.1 Eye movement Data

The authors interpreted the eye movement data through three perspectives. First, the dwell time shows the participants' the degree of attention to each region. Second, latency of the first fixation shows that the attractive regions which participants can quickly position. And finally, the duration of the first fixation and the number of fixation shows the participants' adhere attention to the ROI.

3.2 Latency of the First Fixation and Duration of the First Fixation

The latency of first fixation (Table 1) indicates that the participants' gaze sequence is R_2, R_3, R_6, R_5, R_1, and R_4. It also shows that which ROIs can quickly attract the participants' attention. This result confirms H_1.

The duration of first fixation in Table 1 shows the sequence of the duration of the first fixation is R_5, R_1, R_4, R_3, R_2, and R_6. The longer duration shows that the capacity of attention of the ROI is stronger than the others; H_2 is confirmed.

3.3 Number of Fixations and Dwell Time

The number of fixations shows the participants' times of fixation on each ROI and the sequence is R_5, R_2, R_3, R_1, R_6, and R_4. The dwell time shows the total duration of fixations on each ROI, and the sequence is R_5, R_2, R_3, R_1, R_6, and R_4.

The number of fixations and dwell time show that the two sequences are the same; therefore H_3 is supported. A possible explanation is that when the participants have faster speed of saccade, the dwell time will be longer if the number of fixation times is larger.

Table 1 Eye movement data of each ROI

	Number of fixations		Dwell time (ms)		Latency of first fixation (ms)		Duration of first fixation (ms)	
ROI	Mean	SE	Mean	SE	Mean	SE	Mean	SE
R_1	3.01	.95	440.36	195.89	4567.45	2042.16	212.78	50.53
R_2	4.35	.95	648.30	205.64	1437.58	213.93	128.35	12.78
R_3	3.75	1.34	511.72	208.13	1805.79	418.13	143.96	17.97
R_4	2.79	.60	365.18	97.30	5702.36	2067.85	193.40	29.24
R_5	4.62	1.69	742.60	329.73	4515.88	1990.06	226.98	49.31
R_6	2.88	.56	397.21	106.49	1900.41	400.09	120.80	10.65

Note: Mean and standard error (SE)

4. Discussion and Conclusion

According to the results, R_2 can quickly attract the participants' visual attention, and it also has more times of fixations although its duration of fixation is shorter. A possible answer is that the area of R_2 is largest among the six types of ROIs. The participants' fixations do not immediately fall on R_5, but it has the longest duration at the first fixation, the largest number of fixations, and the longest duration of total fixations. In this regard, R_5 can be considered the most attractive region of a handbag, and it might easily arouses the participants' curiosity and a sense of novelty [1], [2], [9].

The concrete eye movement evidence demonstrates that eye-tracking approach is an effective and efficient approach of predicting potential female consumers' gazing behaviors and visual attention as they look at images of fashion products. The findings of this research have practical contributions to fashion design education. For instance, if featured area (R_5) is the most attractive region to female consumers, educators can advise students to emphasize on the design of this region. The proposed approach can help educators to improve their teaching, and textbooks editors can also add eye-tracking approach to future textbooks. Moreover, gaze can be defined as a form of reading. The findings of this research also extend the existing knowledge of human reading process.

The insufficient participants and the number of handbag pictures, and few variety handbags' styles are the limitations of this pilot study. Nevertheless the eye-tracking approach can also be applied to other fields of education and researches that related to reading process.

Acknowledgements

This research was supported by the "Aim for the Top University Project" from National Taiwan Normal University, the Ministry of Education and National Science Council (NSC 101-2511-S-003-046).

References

[1] Berlyne, D. E. (1958). The influence of complexity and novelty in visual figures on orienting responses. *Journal of Experimental Psychology*, *55*(3), 289-296. doi: 10.1037/h0043555

[2] Daffner, K. R., Scinto, L. F. M., Weintraub, S., Guinessey, J., & Mesulam, M. M. (1994). The impact of aging on curiosity as measured by exploratory eye movements. *Achieves of Neurology*, *51*(4), 368-376. doi:10.1001/archneur

[3] Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods*, *34*(4), 455-470, doi: 10.3758/BF03195475

[4] Ju, H. W., & Johnson, K. K. P. (2012). Fashion advertisements and young women: Determining visual attention using eye tracking. *Clothing and Textiles Research Journal*, *28*(3), 159-173. doi: 10.1177/0887302X09359935

[5] Park, J., Delong, M., & Woods, E. (2012). Exploring product communication between the designer and the user through eye-tracking technology. *International Journal of Fashion Design, Technology and Education*, *5*(1), 67-78. doi: 10.1080/17543266.2011.633566

[6] Pietinen, S., Bednarik, R., & Tukiainen, M. (2010). Shared visual attention in collaborative programming: A descriptive analysis. *Proceedings of the 2010 ICSE Workshop on Cooperative and Human Aspects of Software Engineering*, 21-24. doi:10.1145/1833310.1833314

[7] Radach, R., Inhoff, A., & Heller, D. (2002). The role of attention and spatial selection in fluent reading. In E. Witruk, A. D. Friederici, & T. Lachmann (Eds.), *Basic functions of language, reading, and reading disability* (pp. 137-153). Dordrecht, the Netherlands: Kluwer Academic.

[8] Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

[9] Tarbi, E. C., Sun, X., Holcomb, F. J., & Daffner, K. R. (2011). Surprise? Early visual novelty processing is not modulated by attention. *Psychophysiology*, *48*(5), 624–632. doi: 10.1111/j.1469-8986.2010.01129.x

Open Source Energy Simulation for Elementary School

Sze Yee Lye [a*], Loo Kang Wee[a] & Yong Cheng Ong[a]
[a]*Educational Technolgy Division, Ministry of Education, Singapore*
*lye_sze_yee@moe.gov.sg

Abstract: With the interactivity and multiple representation features, computer simulations lend itself to the guided inquiry learning. However, these simulations are usually designed for post-elementary students. Thus, the aim of this study is to investigate how the use of guided inquiry approach with customized energy simulation can improve students' understanding of this topic. In this ongoing research, the case study is adopted. In the first phase of the study, we have modified open source energy simulation based on principles for reducing extraneous processing, existing energy simulation and guided inquiry approach. The modified simulation is sent to teachers for evaluation and the feedback is encouraging. In the next phase of the study, the guided inquiry lesson package involving the energy simulation would be designed and deployed in an elementary classroom. Multiple data sources would be collected to seek a deeper understanding on how this learning package can possibly impact students' understanding of the physics concepts.

Keywords: simulation, elementary school, physics, science, open source,

Introduction

With the proliferation of computing devices in the classrooms, science teachers are increasingly using technology (e.g., data loggers and computer simulations) to create meaningful learning experiences [1] for the students. Computer simulations, in particular, have gained popularity among the science educators as there is a wealth of easily available of online free and realistic computer [2]. Some examples of computer simulations include PhET's interactive Science Simulations at http://phet.colorado.edu/ , Open Source Physics Simulations at http://www.opensourcephysics.org/ and NTNUJAVA Virtual Physics Laboratory at http://www.phy.ntnu.edu.tw/ntnujava/. A computer simulation is "a program that contains a model of a system (natural or artificial; e.g., equipment) or a process" (p. 180)[3]. Such simulation can also accept inputs from the users and present the computational results in multiple representations like graphs or tables [4]. By providing guidance in inquiry-based activities, computer simulations can be adopted in guided inquiry approach.[5]

Despite the potential of the simulations for use in guided inquiry learning, there is paucity of such research conducted in elementary school . Searches done in the three databases, *Academic Search Premier*, *PsycARTICLES* and *PsycINFO*, using descriptors "science simulations" and "elementary school" only yields 2 results [6,7]. This might be due to the lack of online science simulations created especially for elementary school. Moreover, little research has been done to investigate how other factors (i.e., teachers' facilitation and classroom settings) may impact the students' learning using such simulations[2]. To address this research gap, this paper examines the use of customized energy simulation in elementary classroom. The case study approach is adopted and is guided by the following questions : (1) What are the inquiry learning principles (e.g., level of teacher facilitation) that can improve students' understanding of energy concepts? (2) What are the design features of the computer simulations that help the students to understand energy concepts? In this study, the energy simulation is modified based on one of the many physics simulations found in Open Source Physics Simulations[8]. This is made possible as open source applications are characterized by the access to the source code and free distribution of the application[9]. This simulation is customized in such a way that it can : (1) provide more support for the inquiry-based activities and (2) reduce cognitive load. Such freeing up of cognitive resources is likely to enhance the guided learning experience [5].

1. Computer Simulations in Guided Inquiry

Inquiry based approach has always been popular in science education [10]. Learners are situated in an inductive learning mode, in which students are the "active agent in the process of knowledge acquisition" [2]. Usually, guided inquiry is preferred as the absence or little scaffolding may hinder the students' learning [2]. During such learning, the students are "posing and responding to questions, designing investigations, and evaluating and communicating their learning". [11]

Guided inquiry can be implemented with computer simulations supporting inquiry-based activities. Firstly , as cognitive tool, simulations will help processing of the data (e.g., the representation of data in table or graphic form)[12]. With this ability to share cognitive load, the simulations enable them to focus on higher order thinking skills (e.g., evaluating findings and designing investigation). Moreover, the simulation offers multiple representations (e.g., word, pictures, diagrams, graphs and table of values) of the same or related concepts which help the learners in responding to questions, evaluating and communicating their conclusion[13]. Such representations can foster deeper understanding of the science concepts as the learners can "integrate information from the various representations to achieve insights that otherwise would be difficult to achieve with only a single representation" [14]. Lastly, computer simulations offer interactivity, in which experimental variables can be manipulated [15]. Such affordance allows the learner to design investigation and evaluate their conclusion.

2. Open Source Energy Physics simulation

The topic on energy (e.g. kinetic energy and potential energy) is an important and difficult topic for the students [16]. The use of guided inquiry with energy simulation has the potential to solve this problem. Hence, we choose to modify the existing open source roller coaster simulation [8] so that elementary school students can investigate the energy concepts in a virtual setting. The original roller coaster simulation is designed with equations that model closely the energy concept. This makes the simulation fairly realistic and is not oversimplified such that students will have misconceptions [17]. There are, of course, other good energy-related simulations on the web . As such simulations are designed with older students in mind, the simulations may contain extra information which makes them unsuitable for use in elementary school.

In this study, we have successfully customized an open source roller coaster simulation using Easy Java Simulation (EJS). EJS at http://fem.um.es/Ejs/, free authoring toolkit for creating physics simulation, is part of the Open Source Physics project which aims to spread the use of open source code for physics simulations. Continuing the spirit of open source, this remixed simulation is shared online so that others can further refine or benefit from it [18] . Such spirit resonate with us, the science educators, as both open source applications and scientific knowledge are “built on the concept of creating shared knowledge and the desire to have one’s work adopted by the scientific or computer-using community” [19].

3. Customization the Open Source Energy Simulation

As this simulation is meant for students in middle and high school, there is a need to customize the roller coaster simulation [8]. In the customization process, we are guided by: (1) the principles for reducing extraneous processing, (2) the existing good energy simulations and (3) inquiry learning principles. We used some of the principles(See Table 1) for reducing such extraneous processing outlined by Mayer, who had made significant contribution in multimedia learning [20] . Such reduction in cognitive load is vital as excessive cognitive load may impede guided learning process [5].

Table 1 : Principles for Reducing Extraneous Processing

Principle	**Description**
Coherence	Removal of extra information which is not necessary for learning
Signaling	Inclusion of cues to facilitate the organization of information
Spatial Contiguity	Placing of related representation (i.e., words, diagrams) close to one another

In addition to these principles, we sourced for potential enhancements to be included in the remixed simulation by reviewing the existing energy simulation [19]. Lastly, we also included features that support the inquiry-based activities. See Figure 1 and Table 2 for the features of the remixed simulation. Some of the design features outlined in Table 2 can be found in Figure 1 as indicated by the corresponding number.

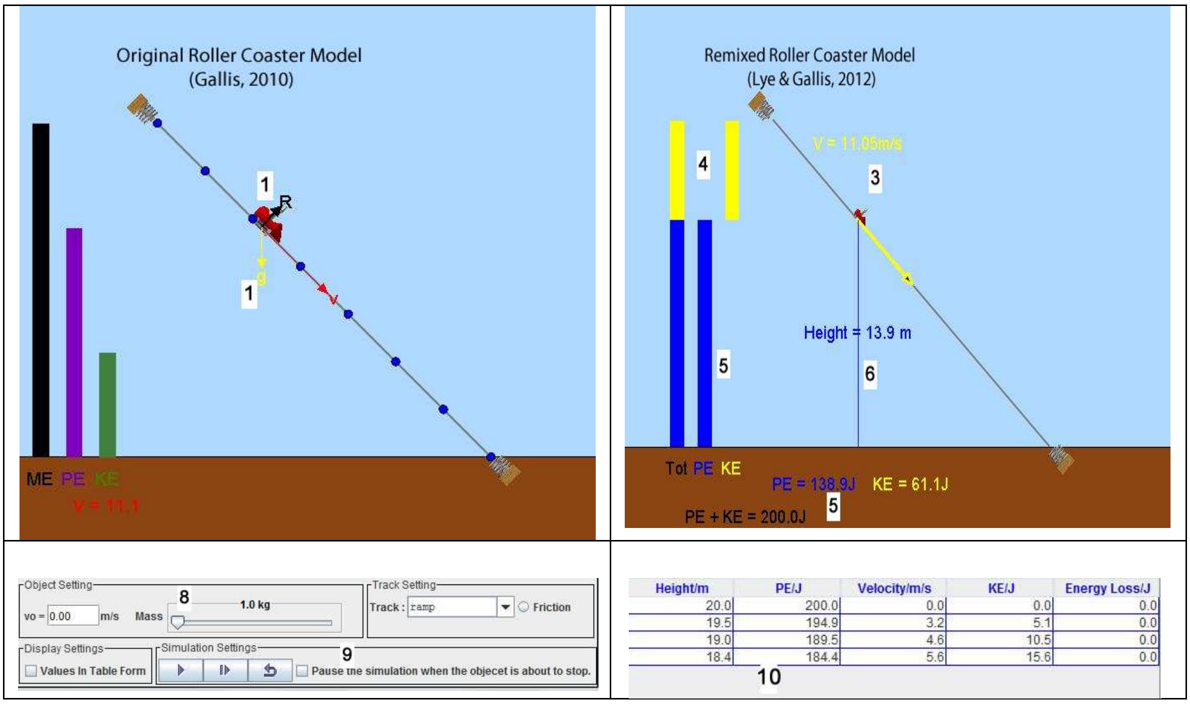


Figure 1 : Design Features of Roller Coaster

Table 2 : Design Features of the remixed simulation

No	Design Features	Rationale
1	Removal of extra information like g (acceleration due to gravity) and R (normal force)	Coherence Principle.
2	Consistent colour scheme for concepts related to potential energy (PE) and kinetic energy (KE).	Signaling Principle. Make the relationship between the related concepts more explicit .
3	Display of velocity information in symbolic (length of arrow) and numerical form near the object.	Spatial Contiguity Effect Better visualization how the velocity changes as the object moves.
4	Modification of Total Energy bar such it is made up of KE and PE energy bars Change the orientation of KE bar such that the .	Spatial Contiguity Effect Better visualization of the total energy concept as the total energy bar is now dynamically made up of KE and PE energy bars.
5	Representation of kinetic energy, potential energy in bars and the total energy in numerical form.	Multiple representations to guide the students in inquiry-based activities like evaluating and communication of findings, responding to questions .
6	Inclusion of the height to show the relationship between height and PE	
7	Arrangement of control settings by grouping similar settings control together.	Spatial Contiguity Effect and Signaling Principle
8	Allowing the learner to pause and step through the simulation when the object is about to stop	Signaling Principle. Focus the learner's attention on what happens when the object is about to stop.
9	Allowing the learner to change the mass	Enhancement made after reviewing [19] Allows the students to design investigation to investigate how mass can affect the energy
10	Inclusion of the table of values of variables like KE	Allows the students to design experiment to investigate the relationship between these variables.

4. Conclusion and Further Work

This is an ongoing case study research which is guided by the following research questions : (1) What are the inquiry learning principles (e.g., level of teacher facilitation) that can improve students' understanding of energy concepts? (2) What are the design features of the computer simulations that help the students to understand energy concepts? The first phase of the research has been completed with the energy simulation being remixed based on Mayer's principles of reducing extraneous processing, existing simulations and inquiry learning principles. The energy simulation is sent to some teachers (N=6) for reviewing. Response is largely positive with all of them agreeing that the added features are useful.

> *I like the correlation between the P.E and K.E as the roller coaster gains K.E and loses P.E and vice versa on the different sets of platforms. It clearly shows the loss and gain of energy.*
>
> *The value of speed with the roller coaster (in yellow) is great because the pupils can see the motion, and see the values*

In the next phase, we will work closely with the teacher to co-design the guided inquiry lesson package. Data from multiple data sources are collected. The data collection instruments include (a) teachers' feedback on the simulation (b) field notes based on classroom observations (c) student-produced artifacts (d) students' feedback on simulation , (e) students' results and (f) teacher's reflection on lesson.

Acknowledgements

We wish to express our deepest gratitude to the following group of people or organizations for making this research possible:

(1) Francisco Esquembre, Fu-Kwun Hwang and Wolfgang Christian for their contribution to the open source physics simulation.
(2) Michael Gallis for creating and sharing his original simulation
(3) Teachers for providing us with feedback on the simulation
(4) Ministry of Education, Educational Technology Division (ETD) and National Research Funding (NRF) for supporting and funding the project: *eduLab-003 : Java Simulation Design for Teaching and Learning* for which this research is spin off from.

References

[1] Linn, M. (2003). Technology and science education: starting points, research programs, and trends. *International Journal Of Science Education*, *25*(6), 727-758.

[2] Rutten, N., Joolingen van, W. R., & Veen van der, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education*, *58*(1), 136-153.

[3] Jong de, T., & Joolingen van, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179–201.

[4] Liu, H., & Su, I. (2011). Learning residential electrical wiring through computer simulation: The impact of computer-based learning environments on student achievement and cognitive load. *British Journal Of Educational Technology*, *42*(4), 598-607.

[5] Jong de, T. (2005) *The guided discovery principle in multimedia learning*. In: Cambridge handbook of multimedia learning. Cambridge University Press, Cambridge, pp. 215-229.

[6] Jaakkola, T., & Nurmi, S. (2008). Fostering elementary school students' understanding of simple electricity by combining simulation and laboratory activities. *Journal Of Computer Assisted Learning*, *24*(4), 271-283.[7] Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. *Computers & Education*, *57*(4), 2281-2290.

[8] Gallis, M. (2010). Roller Coaster Model [Software]. Available from http://www.opensourcephysics.org/items/detail.cfm?ID=8228;

[9] Open Source Initiative. (n.d). The Open Source Definition. Retrieved April 21, 2012, from http://opensource.org/docs/osd

[10] Veermans, K., Joolingen van, W.R, & Jong de, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education*, *28*(4), 341–361.

[11] MOE (2008). Singapore Science Syllabus. Retrieved from April 21, 2012 from http://www.moe.gov.sg/education/syllabuses/sciences/files/science-primary-2008.pdf

[12] Jonassen, D., Howland, J., Marra, R., & Crismond, D. (2008). Meaningful learning with technology. New Jersey, USA: Pearson.

[13] Rosengrant, D., Etkin, E. & van Heuvelen , A. (2006). In L.McCullough, P. Heron, P. &. L. Hsu (Ed An overview of recent research on multiple representations. *2006 Physics Education Research Conference. AIP Conference Proceeding* (pp. 149-152).

[14] Wong, D., Sng, P. P., Ng, E. H., & Wee, L. K. (2011). Learning with multiple representations: an example of a revision lesson in mechanics. *Physics Education*, *46*(2), 178-186

[15] 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.

[16] Stylianidou, F. (2002). Analysis of science textbook pictures about energy and pupils' readings of them. *International Journal Of Science Education*, *24*(3), 257-283

[17] Blake, C., & Scanlon, E. (2007). Reconsidering simulations in science education at a distance: features of effective use. *Journal Of Computer Assisted Learning*, *23*(6), 491-502.

[18] Lye, S.Y., & Gallis, M. (2012). Roller Coaster Remix for use in Elementary School [Software]. Available from http://www.phy.ntnu.edu.tw/ntnujava/index.php?topic=2383.msg8955

[19] Pfaffman, J. (2008). Transforming High School Classrooms with Free/Open Source Software: It's Time for an Open Source Software Revolution. *High School Journal*, *91*(3), 25-33

[20] Mayer, R. E. (2005). *Principles for reducing extraneous processing in multimedia learning: Coherence, Signalling, Redundancy, Spatial Contiguity and Temporal Contiguity Principles*. In: Cambridge handbook of multimedia learning. Cambridge University Press, Cambridge, pp. 183-200.

Development and Application of the Learning System of Basic Subjects of Information Technology Implementing the Scaffolding Teaching Strategy

Sun Hoi KIM[a*], Dong Soo NAM[b] & Tae Wuk LEE[c]
[abc]*Dept. of computer education, Korea national University of education, South Korea*
*sun5252@hanmail.net

Abstract: This study aims to improve academic achievement and learning satisfaction by applying the learning system that employs the scaffolding teaching strategy in the courses for 'basic of information technology' of the technical high school students. To this end, first a web-based scaffolding teaching-learning model was designed, and a teaching-learning lesson plan was devised to form a basis to develop a learning system. This learning system was applied technical high school freshman students. As a result, the experimental group taught by the scaffolding teaching strategy had more positive changes in academic achievement and learning satisfaction than the control group. In addition, a high correlation between academic achievement and learning satisfaction was observed; and therefore, it was concluded that the group with higher academic achievement is more satisfied with what they have learned.

Keywords: Scaffolding teaching strategy, web-courseware, academic achievement, learning satisfaction

Introduction

These days, development of computer and information communication has changed all areas of society, the computer is not a machine that simply performs operation. It's the most important tool that manage audio, data, video multimedia data to sustain data society and make new creations possible. Hence, in 21st century information society, information education using computer is an inevitable problem. The technical high school adopted 'Basic of Information Technology' to make students understand and use the computer and information communication, which can make progressive outstanding individual [1].

So in these circumstances, systematic learning system to make technical high school students is desperate. we are trying to adopt Scaffolding teaching strategy in teaching-learning process to make students to discover their inner potential and use it. Based on this, the purpose of this study is to implement the system that can educate 'Principle of computer' chapter in 'Basic of Information technology' system using the web that maximize teacher-learner and student-student's active interaction, increase academic achievement by increasing individual's problem solving ability and improve the learning satisfaction by increasing academic achievement.

1. Background

1.1 Web-based instruction and scaffolding

Web based instruction can be used by diverse learners with various learning objectives, learning backgrounds, knowledge levels and learning abilities. Unlike original class lesson that was done with learning group at particular level, web based instruction environment should be supported so individuals can perform effective learning activities. Also, web based instruction is operated under the circumstances where it is hard to get help from teachers or peers at the right time and so, it has to be developed so one can learn oneself [2]. By providing scaffolding that reflects content knowledge and project's characteristic in web based instruction, it is able to support learner's individual learning. Scaffolding is teaching strategy that is essentially individualized and engineering approach where scaffolding is adopted in computer learning environment where individualized learning is possible is being tried [6][7][8][10].

Based on the Gagne's 9 events and the scaffolding developed by You Jin JANG(2004), teaching strategy is developed for this study as shown in the <Table 1>[4][5].
Learning system using scaffolding teaching strategy is developed based on the <Table 1>

<Table 1> The development of teaching strategy based on the web using scaffolding

Scaffolding (You Jin JANG, 2004)	Web-Based Instruction (Gagne's 9 events)	Student	Teacher
Activation of background	Gain attention / Inform learners of objectives	Gain attention	Motivation / confirm of learning objectives
	Stimulate recall of prior learning		
	Present the content / Provide learning guidance	Understanding learning content for model lesson	Model lesson
Activities for problem-solving	Elicit performance (practice)	Find a problem-solving method/ Problem-solving for scaffolding	Monitoring / Verbal coaching
	Provide feedback / Assess performance		
Application and wrap up	Enhance retention and transfer to the job	Wrap up for problem solving/ self-assurance	Provide feedback / wrap up

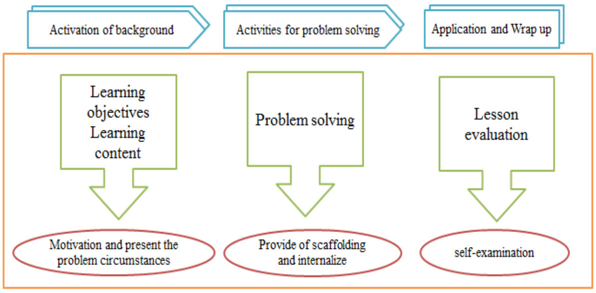
2. System Design and Implement

2.1 Analysis of the learning content

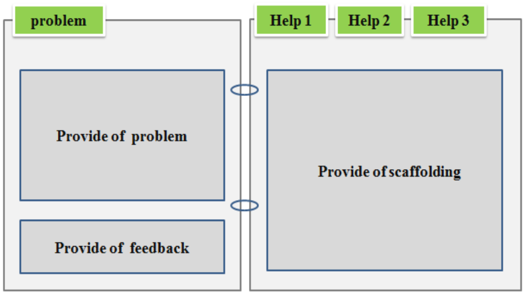
This study surveyed the most difficult subject among 'Basic Subjects of Information technology' for 100 students in the S technology high school students of Suwon city, South Korea. According to the results, they regard 'Principle of the computer' as the most difficult one. Therefore, the learning system of the scaffolding teaching strategy for 'Principle of the computer' has been developed.

2.2 Teaching- learning model and form design based on scaffolding

In this study, learning objectives, learning contents, problem-solving and lesson evaluation are designed based on the scaffolding teaching strategy as shown in the <figure 1>. Form design based on scaffolding provided on problem solving step is shown in the <figure 2>.



<Figure 1> Teaching-Learning Model

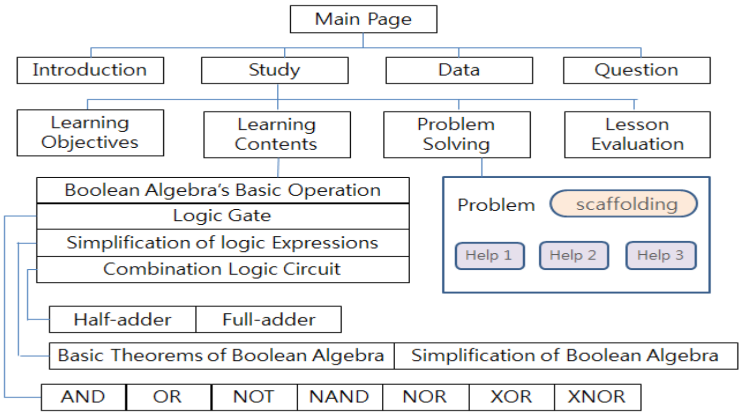


<Figure 2> Page Format

2.3 System overview

2.3.1 Overall system architecture

Main Page is composed of [Introduction], [Study], [Data], [Question] and Quick Menu is composed of [e-mail to master], [Chatting], [Q/A].



<Figure 3> Overall system architecture

2.3.2 Problem -solving page

In problem solving page, questions and help with 3 steps are provided the left frame and right frame respectively. Help providing method is a kind of software based scaffolding types. Help 1 and help 2 provided properly according to each question.

Help 3 is restructured according to this study using chatting source. Also, it is made into screen that can provide scaffolding that can give custom help for individuals during solving problems.

3. Method

3.1 Subject

The subjects of this study were 74 students who were in their first year in S technology high school located in Suwon City, South Korea. Experimental group and the control group were composed of 37 students, respectively.

3.2 Hypothesis

The following hypotheses were formulated:
(1) In 'basic of information technology' subject's principle of computer chapter class, there would be difference in academic achievements and learning satisfaction between classes that adopted scaffolding teaching strategy and classes that didn't.
(2) There will be correlation between academic achievement and learning satisfaction

3.3 Research design and process

The research design of the study are as follows:

Experimental group	O_1 (pretest)	X_1	O_2 (posttest)
Control group	O_3 (pretest)	X_2	O_4 (posttest)

Groups were divided into the experimental group and the control group, both were proved to be the same by pretest. Experimental group was provided with class model that applied scaffolding teaching strategy and control group was provided with class model that did not applied scaffolding teaching strategy. To minimize of the impediment factors that can affect research's validity and reliability, the conditions such as both groups' rate of progress, problem solving, learning contents, teacher's verbal instructions, etc, were same.

4. Research Result and Analysis

4.1 Pretest

Result of pretest of academic achievement (Independent sample t-test):

	Person	Collective averages	Standard deviation	t	p
experimental group	37	56.75	19.44	.058	.954
control group	37	56.48	20.57		

The pretest results of experimental group and control group showed significant level of 0.954 which means there is no significant differences between both groups (p>.05). So, experimental group and control group can be considered as same group.

4.2 Posttest

4.2.1 Academic achievement

Result of posttest of academic achievement (Independent sample t-test):

	Person	Collective averages	Standard deviation	t	p
experimental group	37	71.05	17.65	2.950	.004
control group	37	57.78	20.90		

The posttest results of academic achievement were 71.05 for experimental group and were 57.78 for control group. Experimental group had 13.27 higher academic achievement and significant level was .004 which meant there is statistical significant difference (p<.05).

4.2.2 Learning satisfaction

Result of test of learning satisfaction (Independent sample t-test):

	Person	Collective averages	Standard deviation	t	p

experimental group	37	21.62	4.63	3.498	.001
control group	37	16.70	7.19		

Average of learning satisfaction for the experimental group was 21.62 which was 4.92 higher than that of control group. Significance level was .001 which meant that there is statistically significant difference ($p<.05$).

4.2.3 The correlation between academic achievement and learning satisfaction

Using simple correlation analysis statistical method, the correlation between academic achievement and learning satisfaction is observed. Using the most general Pearson correlation coefficient, we obtained that correlation coefficient between academic achievement and learning satisfaction's is +0.780, which is high. The results explain that high academic achievement means high learning satisfaction.

5. Conclusion

This study tried to apply scaffolding teaching strategy into web based instruction that made individual learning possible and to increase academic achievement and learning satisfaction in technical high school's 'basic of information technology' subject.

For this purpose, simple survey was conducted by S technical high school in Suwon city. Teaching-learning model was development based on the survey results.

As a result, the experimental group taught by the scaffolding teaching strategy showed more positive changes than the control group in academic achievement and learning satisfaction. In addition, a high correlation between academic achievements and learning satisfaction was observed; and therefore, it was concluded that the group with higher academic achievement is more satisfied with what they have learned.

References

[1] Ministry of Education, Science and Technology(2002). Basic of Information Technology. *National institute of National Text Book of KwangWoon University.*

[2] Seoung Mi KIM (2004). The Effects of scaffolding teaching strategy in web based learning environments on learning achievements according to cognitive methods. Master Thesis. Graduate School of Education of Korea national university of Education.

[3] Young Ja SHIN (2005). Design and Implementation of WBI learning program Applying to scaffolding. Master Thesis. Graduate School of Education of A –Ju University.

[4] Tae Wuk LEE, Chang Geun KIM(1998). The Curriculum in Computer Education. *Publisher of Hyungseul.*

[5] You Jin JANG(2004). The Educational significance and directions of scaffolding instruction: in view of the brain science. Master Thesis . Graduate School of Education of Seoul National University Of Education

[6] Biemns, H, & Simons, P. (1992). Learning to use a word processor with concurrent computer-assisted instruction. *Learning and Instruction*, 2, 321-338.

[7] Chee, Y. S. (1995). Cognitive apprenticeship and its allocation to the teaching of Smaltalk in a multimedia interactive learning environment. *Instructional Science*, 23, 133-161.

[8] Guzdial, M. (1995). Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments*, 4, 1-44.

[9] Lieberman, D.A., & Linn, M. C. (1991).Learning to learn revisited: Computers and the development of self-directed learning skills. *Journal of Research on Computing in Education*, 23, 373-395.

[10] Winnips, K.(2001). Scaffolding-by-design as a model for online learner support. Retrieved June 20, 2002 from http://rilw.emp.paed.uni-muenchen.de/2001/papers/winnips.pdf

AR-based Learning Support System for Inorganic Chemistry

Masaru OKAMOTO[a*], Ryoya SUMIDA[a] & Yukihiro MATSUBARA[a]
[a]*Graduate School of Information Sciences, Hiroshima City University, Japan*
*okamoto@hiroshima-cu.ac.jp

Abstract: In this paper, we proposed AR-based learning support system for inorganic chemistry. In order to perform experiments in virtual environment, markers are utilized as control interface and arranged markers are recorded by USB camera. Virtual environment is created from recorded image and various CGs (such as beaker, flame and so on) corresponding to markers. By operating markers, user can perform various experiments (such as flame test). Additionally, so the system can give questions and hints, user acquires knowledge of chemical reaction by solving questions in virtual environment. The verification of the proposed system was conducted by analysis of 12 subjects' learning results.

Keywords: Augmented Reality, Experience-based Learning, Inorganic Chemistry

Introduction

The importance of learning from experiences is strongly described by many teachers and researchers [1]. In order to understand the physical phenomena, our group proposed VR-based learning support system. Virtual environment can show simulation of dynamics experiments for learner. This simulation helps learner to understand the physical phenomena intuitively. In the fields of chemistry learning, experience-based learning support systems were also proposed [2], [3]. Konishi developed of an Intelligent Practice Support System (IPSS) for high school chemistry [2]. IPSS can solve problems and evaluate learner's problem solving process using the result. However, as for field of inorganic chemistry, IPSS is not taken into consideration in study. In the field of inorganic chemistry, it is important for learners to observe an experimental result and process of experiments in Japanese High School education.

In this paper, we proposed AR-based learning support system for inorganic chemistry. Although the experience-based learning support systems using AR technology were developed [4], [5], most of all systems need actual instruments. By this limitation, there is difficulty of preparation and risk. So, in the proposed system, a virtual experiment environment is constructed without the actual instruments. In order to perform experiments in virtual environment, markers are utilized as control interface. Virtual environment is created from recorded real image and various CGs corresponding to markers. By operating markers, user can perform various experiments (such as flame test). Additionally, so the system can give questions and hints, user acquires knowledge of chemical reaction by solving questions in virtual environment. Therefore, there is possibility that learner can find his/her mistake through experiments corresponding to given question. Thus, by using proposed system, leaner can learn about chemical reaction in inorganic chemistry through trial and error.

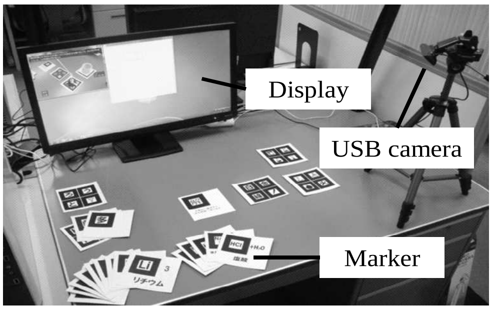


Figure 1: The structure of the proposed system

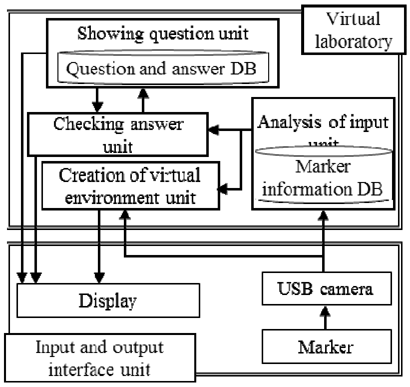


Figure 2: The system structure of the proposed system

Table 1: Examples of Markers corresponding to instruments and item for experiment

	Instruments	Solutes	Water solutions
An Example Image of Markers			

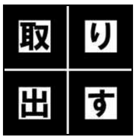
1. AR-based learning support system with questions for inorganic chemistry

1.1 System structure

Figure 1 and 2 show the overview and structure of proposed learning support system respectively. As shown in Figure 1, USB camera records image of user's operation in order to create the virtual environment for experiments based on real image and CGs. Simulation of experiments and creation of virtual environment are carried out by computer, and in display virtual environment (processes and results of experiments) is displayed. In order to construct the virtual environment from real image recorded by USB camera, user's operation must be recognized from real image. Then various markers are utilized for recognition of user's operation. By putting and moving marker in recorded area, user's operation of items, used for performing experiments, are easily recognized. Table 1 shows examples of markers used in proposed system. Used markers are classed into 6 groups. The set of multiple markers is used as marker corresponding to instruments (such as burner and beaker). By putting these marker in recorded area, this system understand that user utilize the corresponding instruments. Additionally corresponding CGs are displayed near the marker in virtual environment. Then, using solutes and water solutions have to be selected for performing experiments which user wants to conduct. In order to indicate the using solutes and water solutions, markers shown in Table 1 are used. In the proposed system, 17 solute markers and 6 water solution markers are prepared.

However, in order to perform experiments in virtual environment, not only the markers as instrument and so on, but also markers corresponding to operation are required. So operation markers shown in Table 2 are used for showing user's intention of operation. Moreover, in the proposed system, practice questions of inorganic chemistry are used in order helps user's learning. For answering a question, user can put marker corresponding answer command in recorded area. By using these markers, user can operate various items in virtual environment learn the chemical reaction about inorganic chemistry.

Table 2: Examples of Markers corresponding to operation by learner

	Operation for water solutions	Adjustment of Solutes' parameters	Checking leaner's result of experiment as answer
An Example Image of Markers			

(a) Creation of copper ion solution (b) Dipping a platinum wire into solution (c) Putting a wire into the flame

Figure 3: Example of experimental process flame color test.

Next subsection shows the chemical experiments which can be carried out in virtual environment of proposed system.

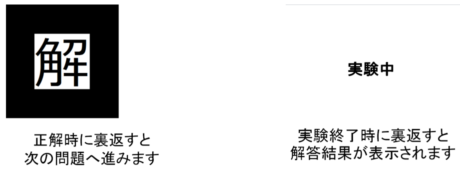
1.2 Chemical experiments in virtual environment

In this system, user can perform the three chemical experiments (1: flame color test, 2: precipitation of ion, 3: positive ion analysis). These experiments can be selected by putting instrument markers (shown in Table 1) which are used in experiments.

For example, a virtual environment of flame reaction experiment is shown in Figure 3. In Figure 3(a), four markers corresponding to instruments and a solute marker are put in recorded area. By putting or moving solute marker as copper ion near the beaker marker, user can create water solution containing copper ion (shown in Figure 3(a)). In Figure 3(b), by moving a platinum wire marker near a beaker marker, a virtual platinum wire is dipped into beaker and a wire have gotten copper ion. In next process, so user should put a wire into the frame in virtual environment for checking change the frame color, platinum wire marker is moved near the burner marker (Figure 3(c)). Then, user can observe the characteristic color of the flame (green) when the wire is in the edge of the flame. As described above, user can conduct the frame reaction experiment by using only the proposed system. Additionally, in virtual experiment, there is no risk by using flame (burner) and no necessity to wash the instruments under the experiment.

By performing these experiments, user can see and know the chemical reactions which are given by textbook and reference book used in Japanese high school. However, it may be difficult for learner to learn chemical knowledge from only these experiments. Additionally, in order to perform experiments effectively, various operations in virtual environment including some operations which cannot be performed in real experiments should be designed.

Next subsection describes about the questions-based learning process and designed operation method for proposed system.



(checking the result of experiment) (during experiment)

(a) An example of presentation of a question in virtual environment

(b) Checking marker

Figure 4: Question and marker for checking answer.

1.3 Question for learning about inorganic chemistry and operations in virtual environment

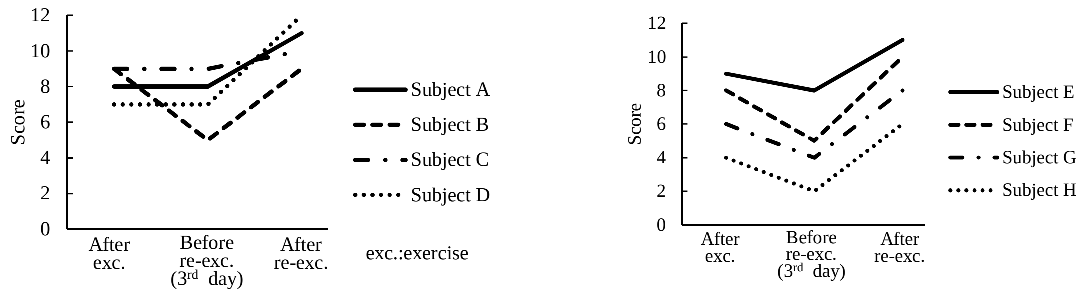
In this subsection, at first, question-based learning approach is described. The question is displayed on the upper part of virtual environment. Figure 4(a) shows an example of presentation of question in virtual environment. In this question, “What is ion which can change the color of flame into blue-green?” is written in Japanese. By presenting a question, user perform experiment in order to find answer about presented question. In proposed system, user’s answer corresponds to results of experiment. User can show his/her situation (experiment is finished or not) by checking marker (Figure 4(b)). This marker is printed at both sides. When this marker is turned, system evaluates the answer (result of experiment). If user makes mistake, hint is displayed on the underside of virtual environment and user perform experiment again based on given hint. After having a correct answer, by turning this marker again, next question is given for user.

Thus mentioned above, by answering to the given question through the experiment in virtual environment, the user can learn about inorganic chemistry. Additionally, markers, corresponding to some functions, help user to perform learning and experiment efficiently.

2. Evaluation experiments

In this section, in order to evaluate the learning effectiveness of proposed system, learning experiments about chemical reaction were conducted. Six subjects (A, B, C, D, E, F, G and H: graduate students and college students) participated in the experiments. Each subject took tests three times. The first test was conducted after the subjects learned about the chemical reaction using the proposed system. The second test was conducted three days after the first test. After the second test, all subjects learn the chemical reaction of inorganic chemistry again using the proposed system. The subjects, that reviewed chemical reaction with the proposed system, were confirmed his/her score by the third test. The number of questions, given by the proposed system, is 12. Similarly, the number of questions written in test is also 12. In the experiments, the subjects were divided into two groups (Group 1 and 2). In order to avoid the influence of given questions, the used questions and test for Group 1 were different from Group 2’s that. The number of questions of each experiment (flame test, precipitation of ion, analysis of positive ion) is 4 respectively.

The results of test are shown in Figure 5. At the first test, most of subjects belonging to Group 1 could get the high score (60% or more). Although the scores of Group 2 are totally smaller than the scores of Group1, it is confirmed that subjects can learn about chemical reaction of inorganic chemistry by using the proposed system. Next, the results of second test show that most of subjects’ score is smaller than the scores of first test. However, the scores of the third test increased from the scores of the first test by relearning using the proposed system. (see results of the third test in Figure 5). From these results, it is confirmed that the proposed system can help learner to learn the chemical reaction of inorganic chemistry. Additionally, it is expected that subjects' learning results improve by learning repeatedly using this system.



(a) Scores of Subject A-D (b) Scores of Subject E-H

Figure 5: The results of scores of all subjects.

3. Conclusions

This paper proposes an AR-based learning support system for inorganic chemistry. The system utilizes markers as control interface in order to perform experiments in virtual environment. Based on positions of arranged markers, virtual environment is constructed from recorded image by USB camera and CGs (such as beaker, flame and so on). By operating markers, user can perform various experiments (such as flame test). Additionally, so the system can give questions and hints, user acquires knowledge of chemical reaction by solving questions in virtual environment.

In future works, we plan to increase the number of subjects and test times for detailed verification. And, we would like to improve the proposed system by adding the effective feedback information for user in learning process. Furthermore, questions and suitable hints given by this system should be investigated.

Acknowledgements

This work was partially supported by Grant-in-Aid for Young Scientists (B) No. 22700819 from MEXT Japan.

References

[1] Dewey, J. (1997). *Experience And Education*. Free Press.

[2] Konishi, T., Okada, Y., Iizuka, D. & Itoh, Y. (2010) Development of an Intelligent Practice Supporting System for High School Chemistry, *Proceedings of the 18th International Conference on Computers in Education* (pp.66-70).

[3] Nanko, R., Okada, Y., Konishi, T. & Itoh, Y. (2008) Constructing Intelligent Virtual Laboratory for High School Chemistry to support learners' consideration, *Proceedings of the 16th International Conference on Computers in Education* (pp.105-109).

[4] Asai, K., Kondo, T., Kobayashi, H. & Sugimoto, Y. Y. (2011). Browsing Lunar Surface Using Tabletop Augmented Reality at Exhibit in Science Museum, The Journal of Information and Systems in Education, 10(1), 11-31.

[5] Sano, A. (2010). A Web Application for Creating Real-Size Augmented Reality Content without 3D Modeling Skills, The Journal of Information and Systems in Education, 10(1), 32-38.

Students' self-control and learning outcome in a university blended course

Yue ZHU[a*], Wing AU[b] & Gregory C R YATES[c]
[a]*School of Education, University of South Australia, Australia*
[b]*School of Education, University of South Australia, Australia*
[c]*School of Education, University of South Australia, Australia*
* linda.zhu@unisa.edu.au

Abstract: Self-control has been regarded as a critical factor which influences people's success in their lives, including their academic achievement. This paper reports a study which aimed to investigate the impact of self-control on a group of 94 university students' learning in a blended course. Data about their learning experiences were collected via pre-test, post-test and weekly reports during their studies. Their capability of self-control was measured by a self-control trait scale. It was found that the self-control could predict the students' learning outcomes. Path analyses revealed that the effect was mediated through course participation.

Keywords: Self-control, course participation, learning outcome, blended learning

Introduction

By combining the strengths from both traditional and online learning, blended learning has become increasingly popular in higher education in Australia. Learning in this environment, the students need to take good control on their study by maintaining the motivation and effort and resisting the factors distracting their study. Self-control can be regarded as one's capability to modify or adapt oneself in order to remain a better and more optimal fit between self and world [1]. Self-control enables people to make plans and carry them out in the face of difficulties and challenges [2]. Therefore, self-control refers to a personality trait, needed to achieve long term goals through being able to control one's impulses and to resist threat to those goals [3].

Self-control is assumed to be associated with the students' learning outcomes. In Tangney's [4] research, students used study strategies more effectively had better learning outcomes. Mischel [5] found the participant in their research with lower level of self-control had troubled psychological portrait in their lives. Moreover, the impact of self-control could influence people's every aspect of their lives, such as educational qualification, health, wealth, and criminal offending [6]. Although the causal relationship between self-control and learning achievement is widely accepted, further research should be done in this area, especially in a blended learning environment. In a web-based learning environment, having good self-control could help students maintain their focus and effort on learning and keep their course engagement in order to achieve their learning goals. More research is needed to study how self-control can predict the students' learning outcome, e.g., mediated through their course participation or engagement. Moreover, it is worthwhile to investigate further, in a blended course, what aspects of the course engagement (e.g., online interaction, the amount of time spent online and offline) will contribute to the students' learning achievement.

Methodology

The present paper reports a research project which examined the learning of 94 university students who took a blended course which combined traditional lectures and workshops with online interactions and online access to course materials. 94 students completed the online pre and post-test questionnaires at the beginning and the end of the semester respectively. During the semester, the students were asked to complete online brief study reports for 6 weeks, which mainly asked about the hours study spent for the course, perceived difficulty level, stress, and workload. 74 out of the 94 students participated in the weekly report survey. This paper presents the effect of self-control on the students' learning outcome.

The students' personality traits of self-control and learning outcome were measured in the pre-test and post-test respectively. The measurement on self-control used in the present study was the Self-control Scale by Tangney et al. [4], which was a questionnaire with 13 items from the long form of 36 items. For each item, students would answer a self-referring statement based on a five-point scale from "Not like me at all" to "Very much like me". The post-test also measured the students' learning outcomes, was indexed by the aggregate percentage score recorded for assignments completed and marked within the course.

Their course participation was reflected by the data from the students' posts in the course online forum each week and weekly reports tapped the students' course experience over 6 weeks. The records form the students' online discussion were coded in the database in terms of number of contributions, latency, and contribution length.

Results

Among the 94 participants, 52% were above the mid-point (39) of the items of self-control, which included students' capability of deferring gratification, resisting temptation, managing efforts, and achieving long-term goals. The students' overall learning outcome was the final marks they received within the course, which was the aggregate of the results from assignments and online forum contribution based on each assessment's weighting. The full score was 100. By the end of semester, 93 students out of 94 had passed the course, among whom 23% achieved Distinction and High Distinction (75-100), 46% received Credit (65-74), and 20% obtained Pass (50-64), and 1 student (1%) failed (below 49).

In the weekly reports, there were different factors such as the number of hours they spent online (online hours) and offline (offline hours) for the course, difficult level, work load, and stress level. The number of the weekly reports the students finished was counted as the report frequency. Finally, the students rated the course experience as positive and negative, and the instances of both positive and negative ratings were counted as the course rating. The available figures were averaged across the 6 weeks for obtain a weekly average figure for each participant. Several factors in the online forum were accounted in the students' course participation, which are elaborated as the followings: *online contribution frequency --- the number of the posts they sent to the online discussion forum; online contribution length --- the sum of the words of the students' posts.*

From the results of correlations, it shows that the students' learning outcome is correlated with their capability of self-control, online contribution frequency, online contribution length, report frequency and difficult level at a significant level (table 7).

Table 7 *Correlations between learning outcome, self-control, and the factors during learning process*

	1	2	3	4	5	6	7	8	9
1. Learning outcome	–	.25*	.51**	.52**	.25*	.26*	.27*	.14	.22*
2. Self-control		–	.23*	.19	.05	.17	.14	-.19	-.01
3. Online contribution frequency			–	.61**	.25*	25*	.34**	.25*	.07
4. Online contribution length				–	.32**	.23*	.35**	.29*	.06
5. Report frequency					–	.16	.90**	.66**	.09
6. Online hours						–	.21	.14	.28*
7. Positive ratings							–	.59**	.14
8. Negative ratings								–	.18
9. Difficult level									–

Note. (a) n = 74, (b) *. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

The concept of the students' course participation was initially constructed on the key variable - online contribution frequency, which was also correlated with online contribution length, report frequency, online hours, positive ratings, and negative ratings. All these variables were added in the construct of SmartPLS model [7, 8] one by one progressively to ensure that the validity of the construct for each stage could be maintained. Based on the result of the construct validity, four variables (online contribution frequency, online contribution length, report frequency, and online hours) contributed to the construct of course participation (AVE = .48, Cronbachs Alpha = .63). Therefore, course participation was indexed by four independent factors: (a) online contribution frequency, (b) online contribution length, (c) report frequency, and (d) online hours. Table 8 illustrated the distribution of the students' course participation. It is found that the students' learning outcome was related with self-control and course participation at a significant level (Table 9).

Table 8 *Frequency of weekly reports*

	Mean	*Median*	*SD*	*Skewness*	*Kurtosis*	*Minimum*	*Maximum*
Course participation	8.82	8.63	2.73	.78	1.08	4	18

Note. (a) n = 74

Table 9 Correlation between self-control, course participation, and learning outcome

	1	2	3
1.Learning outcome	–	.25*	.58**
2.Self-control		–	.23*
3.Course participation			–

Note. (a) n = 74, (b) *. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

Relationships between self-control and learning outcome were investigated using the partial least squares approach (PLS) for path modeling. The initial model tested is shown in Figure 4. This figure depicts the expected relationships between self-control, course participation, and outcome, as based upon initial expectations.

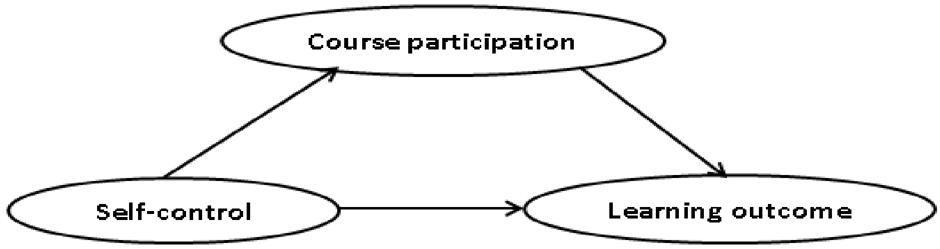


Figure 4 Initial model of the impact of self-control on students' learning outcome

Figure 5 presents the final depiction of the significant pathways of the relationship between self-control, course participation, and outcome. Non-significant pathways were trimmed to produce a parsimonious and descriptive model. 34% of the variance in learning outcome was predicted by the variables (self-control and course participation) (R^2 of .34). In the final path modeling, self-control has a direct influence on course participation. However, the effect is mediated through course participation. The direct and indirect effects of self-control on the students' learning outcome account respectively 46% and 54% of the relationship (Table 10).

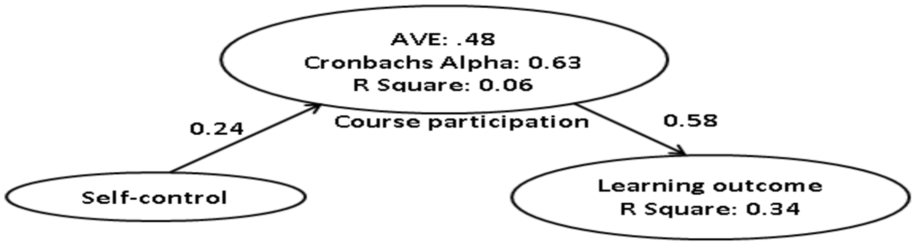


Figure 5 Final model of the relationship between self-control, course participation and outcome

Table 10 Mediation analysis: impact of self-control on learning outcome

Bivariate correlation	.246
Total effect accounted for in the PLS model	.246
Direct effect (self-control to learning outcome)	.114 (46%)
Mediated through course participation	.132 (54%)

Note: The percentage figures refer to percentage of the variance accounted for with regard to direct and indirect effects. The indirect effects account for 54% of the relationship.

Discussion and conclusion

It was noted that the students' scores on a personality scale of overall disposition to exercise self-control, which was measured at the beginning of the study period, could significantly predict their learning outcome by the end of the course. The finding in the present study adds to the construct validity of the scale published by Tangney [4]. It is also important to point out that the prediction of self-control on the students' learning outcome was mediated through their course participation, the hours spent for the course online, active participation in online forum and the project, which could represent the effort and hard work the students put for the course.

As there are different variables tapping the students' learning process and the main factors were correlated with each other (Table 7), construct development is necessary to form a single resolution. After testing the construct validity through PLS model, a single factor resolution has been approved, which also improve the consistency of the students' behavior in the course regarding different aspect of learning process. This implied that the students with active course participation would be motivated to put effort to various learning tasks in terms of the time they spent for the course, participation in online discussion and the research itself. This highlighted the importance of online engagement, in terms of the time spent online for researching, browsing, practicing and the motivation to participation in online forum. Therefore, in a blended learning environment, the students' online engagement can be regarded as a critical factor for their success.

Finally, in addition to the influence of self-control on the students' achievement in traditional classroom [9], the findings of the research emphasized the importance of the positional self-control in an online learning environment, which was also reinforced by Tsai [10]. In Tsai's study, three domains of e-learning strategies are identified in the mode --- perceived-skill, affection and self-regulation domains. It would be worthwhile to have further investigation on the factors of the students' learning process, such as participation in online forum, the motivational strategies and cognitive strategies they applied, the impact of online interaction between the students and teachers on their learning, as the certain parts of the learning process through using online learning components in a blended course remained unknown in the present research.

The limitations of the presentation study concerned the data regarding the weekly report. The brief questionnaire was conducted weekly, but was returned in only about half the cases. The data of students' participation in their learning in terms of online hours were averaged across weeks for the purpose of analysis, which could make it available that a meaningful unit could be derived from each individual participant. However, the extent to which this constituted an accurate index of the amount of time the students spent overall remained unknown.

References

[1] Rothbaum, F., Weisz, J. R., & Snyder, S. S. (1982). Changing the world and changing the self: A two-process model of perceived control. Journal of Personality and Social Psychology, 42, 5-37.

[2] Vohs, K. D., & Baumeister R. F. (2004). Understanding self-regulation: An introduction. In R. F. Baumeister & K. D. Vohs. (Eds.), Handbook of self-regulation research, theory, and application. New York: The Guilford Press.

[3] Baumeister, R. F., & Tierney, J. (2011). Willpower: rediscovering the greatest human strength. New York:Penguin.

[4] Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success, Journal of Personality, 72 (2), 271-322.

[5] Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. Science, 244(4907), 933-938. doi: 10.1126/science.2658056

[6] Moffitt TE, et al. (2011), A gradient of childhood self-control predicts health, wealth, and public safety, Proceedings of the National Academy of Sciences, 108 (7), 2693-2698.

[7] Vinzi, V. E., Chin, W. W., Henseler, J., & Wang, H. (2010). Handbook of partial least squares. Heidelberg: Springer.

[8] Ringle, C. M., Wende, S., & Will, A. (2005). SmartPLS 2.0(M3) Beta. Hamburg: University of Hamburg.

[9] Feldman, S. C., Martinez-Pons, M., & Shaham, D. (1995). The relationship of self-efficacy, self-regulation, and collaborative verbal behaviour with grades: Preliminary findings. Psychological Reports, 77, 971-978.

[10] 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.

Creation of a Topic Maps-Based Wiki with an Article Similarity Measurement

Shu Matsuura[a*], Motomu Naito[b] & Hiromi Toyota[c]
[a]*4-1-1 Nukuikita, Koganei, Tokyo 184-8501, Tokyo Gakugei University, Japan*
[b]*3-3-1 Yshiki-cho, Takahama, Aichi 444-1331, Knowledge Synergy Inc., Japan*
[c]*3350 Honmachida, Machida, Tokyo 194-0032, Honmachida Higashi Elementary School, Japan*
*shumats0@gmail.com

Abstract: We created a pilot case of a Topic Maps-based wiki site to exchange ideas about children's behavior in elementary schools. The topic map of this site consisted of the article and subject topic types. The subject topic type consisted of topics classified into areas such as "behavior," "competence," "field," and "school time." Each article was registered as an instance of the article topic type and associated with relevant subject topics. To measure the similarity between two articles and to relate articles on the basis of the similarity, we used the Tanimoto Similarity. To improve similarity-based retrieval, it was suggested that more specific subject topics characterize the articles.

Keywords: Topic Maps, Wiki, Tanimoto Similarity

Introduction

Since 1998, it has been indicated that elementary schools in Japan often face difficulties in managing classes. [1]. To overcome this situation, both experienced and inexperienced teachers need to share the various problems they face in the classrooms and their solutions. For this purpose, we created a pilot case of the wiki site [2] based on Topic Maps (ISO/IEC JTC1/SC34) [3]. The topic represents the subject. The topics are interconnected through their associations. Information resources are linked to the corresponding topic by the occurrences [4]. Topic Maps fits to the construction of a wiki site [5, 6].

Our wiki site aims to let teachers who write or read articles find information related to it. In general, while an article represents at least one subject, it can be related to various contexts, even if the writer has little intention of doing this. In other words, if two articles share certain common factors, they are similar. Reading articles that have certain degree of similarity may evoke some hints on solving the problem from a different aspect.

This paper introduces a simple method for associating articles. We classify articles as individual topic instances. Then, we ask writers to choose topic names related to the article. Thus, individual articles are characterized by sets of chosen topic names. We evaluate the similarity between any two articles by calculating Tanimoto Similarity.

1. Method

1.1 Educational issues in children's daily life

In this section, we list educational issues faced by elementary school children. We use these particular issues as topic instances to associate with wiki articles.

Competence. Children's education aims to have children attain social skills. OECD has defined the key competency for living and cooperating in the modern society from a broader standpoint [7]. We consider five categories of children's competency—physical strength, intelligence, willingness, practice, and communication. Table 1 shows the indications for assessing the competence of the five categories. These indications have been selected through discussions of elementary school teachers led by one of the authors.

Table 1. Children's competency in school life.

Category	Competency types	Indications
Physical strength	Posture	*Standing and sitting postures*
	Group gymnastic skill	*Ball games, group games*
	Individual gymnastic skill	*Run, jump, apparatus gymnastics*
	Health	*Likes and dislikes in food, illness*
Intelligence	Reading skill	*Reading aloud, comprehension, Kanji*
	Writing skill	*Writing letters, figure, sentences*
	Arithmetic skill	*Arithmetic*
	Logical thinking	*Vocabulary, comprehension, thinking, expressing*
Willingness	Expressiveness	*Smiling, laughing*
	Perseverance	*Tenacity*
	Will for living	*Positive thinking, not depressed*
	Will to act	*Positive attitude*
Practice	Cooperation with friends	*Play with friends*
	Cooperation in work	*Act in cooperation*
	Roles in daily life	*Day duty, activity in charge*
	Rules in daily life	*Following rules*
Communication	Listening to	*Looking at, listening to others*
	Assessment of situation	*Behave according to the situation*
	Expression	*Tell others what the child feels or thinks*
	Sympathy	*Guess what others feel*

Behavior. Children's problems appear in their behaviors. Indications in the behavioral types consist of *school refusal, truancy, antisocial behaviors, rude, nonsense, forgotten, perverseness*. They supplement the assessment of competency. Describing children's behavior is required particularly when the problem is not directly attributable to any specific competency factor.

Space and Time. Many of the children's problematic behaviors occur at a specific time and place. We categorized the typical time periods of children's activities in Japanese schools into the following 8 types, as "going to school," "morning assembly," "classrooms," "intermissions," "school lunch," "cleaning rooms," "social activities at school," and "getting out of school". Under these time period types, 26 indications of children's activities were located in total.

The category "field" consists of 7 location types and 14 indications. The types of "home," "school," and "classroom" specify physical location, while the types "with friends," "with relatives," "local community," and "external community" indicate that the fields are characterized by human relationships. The classified items are not mutually exclusive, but instead supplemental in characterizing the children's environment.

1.2 Constructing the topic map

We constructed an RDBMS topic map and a website on the basis of a topic map development suite Ontopia 5.1.0. [8]. Our topic map consists of "article" and "subject" topic types. The article type contains individual articles written by wiki users.

The article type is associated with the following four types by the broader_narrower association. The "cares_article" type includes instances of articles on teacher's troubles or worries. The "rules_article" type includes useful rules that are applied in the class and work to solve these problems. The "suggestions_article" type includes hints, ideas, and experiences that solve problems faced in class management. Finally, the "teachers_word_article" type includes teachers' words that either were beneficial or, in contrast, not useful in solving the children's problem.

When an article of one of the four article types was uploaded, an article instance of the type is created. Posted article automatically gets a unique base name and a public subject identifier. Finally, the contributor associates his or her article with the subject topics by the "article_related_with_subject" association, to characterize the meaning of the article.

For subject topic type, we subdivided the type into two categories of the "article_subject" and the "article_situation". The article_subject was further divided into four types of "competence," "behavior," "school time," and "field," which were described in Sec. 1.1. The subject topic instances correspond with indications described in the above section.

The "article_situation" is further divided into "teacher's_reflection" and "worry." The "teacher's_reflection" type concerns with the instructor's attitude toward a child or a class, including 5 topic instances of "child_assistance_mind," "class_assistance_mind," "class_management," "educate_child," and "watching_child." The "worry" topic concerns who is particularly worrying, including 3 instances of "children_worry," "parents_worry," and "teacher_worry."

1.3 Similarity calculation

In this system, wiki writers characterize their articles by the set of subjects associated with the article. Although this is an indirect method of characterizing articles, it offers a simple approach to measure the similarity between two articles. We regard the two articles as having similar features if they have common subjects associated with them.

To measure the similarity, we calculate the Tanimoto Similarity between the sets of subjects associated with the articles. Tanimoto Similarity is the rate of intersection of the union of two sets, which assigns "1" for equivalence and "0" for no similarity [9]. If we assign the set of associated subjects A of an article a and the set of subjects B of an article b, the Tanimoto Similarity T_{ab} between articles a and b is written as $T_{ab} = |A \cap B| / |A \cup B|$.

2. Results and Discussion

At present, a total of 102 articles have been written: 67 articles for the cares topic, 14 for the suggestion topic, 14 for the rules topic, and 7 for the teachers_words topic.

Figure 1a shows a semi-log plot of the histogram showing the similarity between any two articles; 64.7% of the articles had fewer than 9 associations. The frequency of the article pairs decayed almost logarithmically as their similarity increased. The articles having 3 or 4 associations were the most frequent, although many associations were mutually exclusive and could not be selected for the same article. 33.5% of the pairs showed a similarity ranging from 0.1 to 0.3. In this study, the range of similarity for showing similar articles is set to more than or equal to 0.3. This value covers only 5.6% of the article pairs.

Here we consider a simplified uniform model, in comparison with our wiki. We consider the articles $A = \{A_1, \dots, A_N\}$, where A_i has m_i ($\leqq M$) associated subjects. Here M is the total number of associated subjects, and the number m_i of associated subjects is chosen uniformly at random from 0 to M. The subjects are chosen randomly from the set of subjects

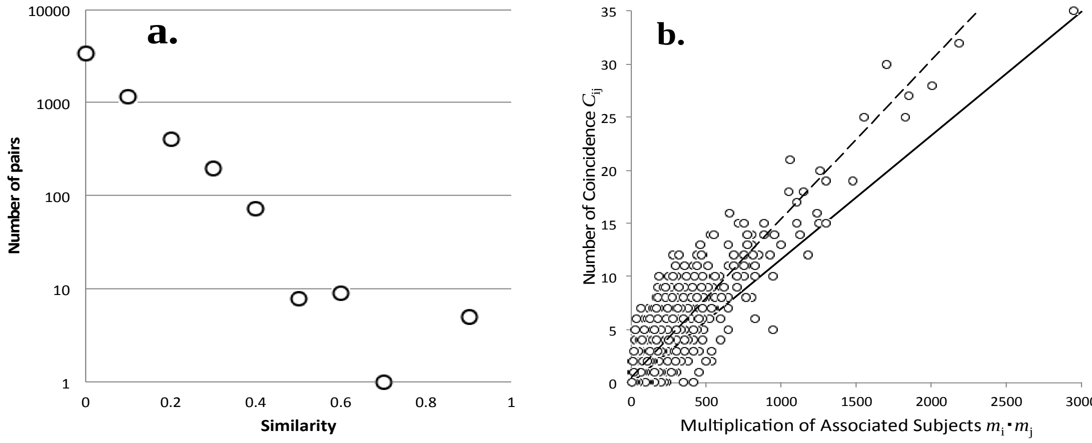


Fig. 1a, b. Left, a, semi-log plot of histogram of the similarity of article pairs. Right, b, plot of C_{ij} verses ($m_i \cdot m_j$) of pairs of actual wiki articles. A bold solid line indicates the slope of M^{-1} = 1/86. The dashed line indicates a fitting line whose slope is approximately 1/67.

S, where $S = \{s_1, \ldots , s_M\}$, for every A_i. Similarity between A_i and A_j is calculated by applying the Tanimoto Similarity on the sets of associated subjects m_i and m_j.

We define the number of equivalent subjects in the sets of associated subjects as "the number of coincidence C_{ij}." The number of coincidence C_{ij} between A_i and A_j is expected proportional to the multiplication $m_i \cdot m_j$, where the coefficient of proportion will be M^{-1}. This relationship is expressed by an equation $C_{ij} = M^{-1}(m_i \cdot m_j)$ (1).

Figure 1b shows a plot of C_{ij} verses $m_i \cdot m_j$ on the actual wiki articles. The solid line shows eq. (1), where M is the real number of subject instances; M = 86. The plotted points are concentrated at lower values of $m_i \cdot m_j$. In addition, the plotted points show higher values of C_{ij} as compared with the line of eq. (1) with M = 86. The dashed line shows the relationship with M = 67. This implies that the range of selection of associated subjects was effectively smaller than the possible number of choices. To increase the number of articles having high similarity, we have to advise wiki authors to associate a single subject with as many viewpoints as feasible.

Finally, we consider the relevance of subject topics in specifying the articles. Table 5 shows six subtypes of subject topics and their actual usage in characterizing the articles. While the Article_situation types have only a few instances, larger numbers of articles are associated per instance, in comparison with the Article_subject types.

Table 2. Number of subject topic instances associated with articles.

Subject type	Subject sub-type	Number of instances selected by contributors	Mean number of associated articles
Article _subject	Behavior	7	6.4
	Competence	35	8.1
	School_time	26	8.5
	Field	8	14.8
Article _situation	Teacher's_reflection	5	19.4
	Worry	3	31.0

In general, an article consists of several specified conditions such as "who is it about," "what," "when," "where," "why," and "how did it happen." These factors are associated with the instances of Article_subject type. Thus, in many cases, the Article_subject instances are regarded as narrower than the article topic instances.

On the other hand, the instances of Article_situation subject type do not characterize what is specifically described in the article. Rather, these instances classify the articles from a broader perspective. Then, the instances of Article_situation subject type are associated with many article topic instances.

To measure the similarity in the articles' specificity, the similarity measurement for the narrower subjects is more effective than that for the broader ones. Improvement in the interface is required to increase assignment of narrower subjects and at the same time decrease the burden of checking many items.

Our approach in this study was to let contributors reconsider what subjects their text could be associated with even though such relationships was not explicitly written in their texts. Another important analysis method is obviously the text mining, which explore new trends from large number of texts [10, 11]. The text mining methods will be effective to extract trends in the wiki articles if the sufficient number and variety of articles are collected.

3. Conclusion

We created a pilot case for constructing a wiki site based on Topic Maps to share the problems, suggestions, and effective rules concerning the life of elementary school children. The articles were stored as article topic instances and were associated with various subject topics including children's behaviors, core competencies, and life skills. On the basis of the retrieval of associated topics, the articles were automatically organized and rendered searchable in the wiki site. In addition, the Tanimoto Similarity calculation evaluated the similarity of subjects associated between two articles. Because the contributors' choice of associated subject influences the effect of this method, the interface needs to be improved for appropriate choice of subject with low burden.

Acknowledgements

This study has been funded by a Grant-in-Aid for Scientific Research (C) 21500842 from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] National Institute for Educational Policy Research. http://www.nier.go.jp/shido/centerhp/unei.pdf (in Japanese).
[2] Wiki site of this work, http://tm.u-gakugei.ac.jp/ca/k/ (in Japanese).
[3] JTC 1/SC 34. http://www.itscj.ipsj.or.jp/sc34/.
[4] Pepper, S. The TAO of Topic Maps. http://www.ontopia.net/topicmaps/materials/tao.html#d0e632.
[5] Cerny, R. (2008). Topincs Wiki – A Topic Maps Powered Wiki. *Lecture Notes in Computer Science*, 4999/2008, 57-65.
[6] Thomas, H. (2007). Design Principle for a Topic Maps Wiki –The Wiki Way of Topic Maps. http://www.informatik.uni-leipzig.de/~tmra/2007/slides/redmann_TMRA2007.pdf.
[7] OECD. The Definition and Selection of Key Competencies, Executive Summary. http://www.oecd.org/dataoecd/47/61/35070367.pdf.
[8] Ontopia code google. http://code.google.com/p/ontopia/.
[9] Rogers, D. J. & Tanimoto, T. T. (1960). A Computer Program for Classifying Plants. *Science* 132, 1115-1118.
[10] Hearst, A. M. (1999). Untangling Text Data Mining, *Proc. of ACL-99*, 3-10.
[11] Feldman, R. & Sanger, J. (2006). *The Text Mining Handbook*, Cambridge University Press.

Practice of Subject Report Revision Process Following Class and Learning Materials Design Method Based on Instructional Design

Shigeru SASAKI[a*] and Hiroyoshi WATANABE [a,b]
[a]*School of Science and Engineering, Teikyo University, Japan*
[b]*Learning Technology Laboratory, Teikyo University, Japan*
*sasaki@ics.teikyo-u.ac.jp

Abstract: Authors have proposed the class and the teaching-materials design method based on the instructional design processes, and developed the support tool for the learning materials development following our method. One of the goals of this research is realizing student assistant participation to the learning materials development. In this research, the revision procedure of the subject report was devised as the part of our learning materials design and development method. Following this method, we have revised the subject report by student assistant participation.

Keywords: Learning materials development, class design, instructional design, e-learning

Introduction

To design learning materials for self-learning courses, instructional design (ID) concepts and systematic models are vital and helpful. However, the development of learning materials that strictly following the ID models can be extremely difficult and time consuming. Therefore we propose a class design and the learning materials development method based on the ID models. In our method, we introduce the "contents outline" that focuses especially on the contents in addition to the "class outline" that covers the activities of the overall class.

In general, learning materials are repeatedly used in universities and the same problems are consistently shown on subject report, then students might simply share their answers and hand in a copy. Therefore, we need to revise the issues of subject reports for each class to resolve this problem. The class design procedures based on the ID concepts include an evaluation and revision process [1, 2]. In this case, learning materials are revised if needed as a result of verification and the revision of the subject report is performed. It is expected that the ID methods for the evaluation and revision process are useful in the revision of subject reports. We have devised the revision procedure of report subjects based on our learning materials development method in which the class outline and the contents outline are used. We have carried out the practice of the subject report revision with student assistant participation following our method.

1. Background of the practice

Authors have been developing learning materials by student assistant participation [3]. At the beginning, we asked the student assistants for creation of learning materials based on the class outline, which focuses on the activity of the entire lesson. However, since the performance of those contents did not reach the level that the teacher had expected, the learning materials were not used in the classes.

To avoid such a situation, we introduced a “contents outline,” which focuses specially on the details of learning materials, in addition to the class design. With this methodology, the learning materials developed by the student assistants were closer to what the teachers had expected and the overall development was clearer.

Since the two design outlines were revised to develop learning materials and carry out the evaluation and revision process, the revision work had becomes complicated. Therefore, we developed a support tool titled “Class/Contents Outline Editor” (COEdit), which linked the related items so that they can be edited together [4].

2. A class design and the learning materials development method

The ID concepts and systematic design models have been the subject of significant focus, especially in regard to the development method of e-learning course design and learning materials. In systematic models, such as ADDIE model, the output from the prior step is used as an input to the following step. In the systematic process of the ADDIE model, there are feedbacks from the "Evaluate" step to the other steps. Dick and Carey’s model is a famous systematic model that contains feedbacks from the formative evaluation step [1].

However, it is difficult to strictly follow all the procedures in the systematic ID models. Morrison, Ross and Kemp proposed the nine element ID model, in which their order and selection are not predetermined [2]. This model appears useful when utilized it in a real situation because the ID models can be arranged according to the actual conditions and environment.

Nakai et al. described the nine step model [5], which was based on Dick and Carey’s model [2]. Their procedures are systematic and suitable for class design at the university level: Steps from 1 to 5 cover the design of the entire course; Steps from 6 to 8 focus on the design of each class; and Step 9 evaluates the learning materials.

For our method, the class outline was created following the procedures of Steps from 6 to 8. Then the proposed contents outline was created based on the class outline, which focuses on the learning activities. The composition of the class outline is shown in Figure. 1.

- Objectives of the class
- Test items of the class
- List of the items that should be explained to achieve the objectives
- A flow of the module
    - Introduction
    - Contents of the learning material
    - Learner’s activities
    - Test items of the class
    - Progressive studies

Figure 1. Composition of the class outline

On the other hand, the contents outline emphasizes on the composition of the learning materials as the primary goal. The flow of creating the contents outline based on the class outline is as follows.

1. Design the composition and the flow of the learning materials themselves for the entire class.

2. Based on the composition and flow of the learning materials, design the composition and the appearance of the page, as well as clarify the details of the items that should be explained.
3. Develop the contents of each page such as explanations and assessments.

It appears that to make the contents outline similar to the concept of rapid prototyping [6] by Jung et al., we must, first, clarify the details of the learning materials.

Numerous ID models include an evaluation process. Since the evaluation and revision process only describes about the problem analysis and measures, it rarely describes the procedure of concrete revision. For this research, the revision of learning materials through student participation has been created, and learning materials for subject reports have been revised along with the procedure.

3. Support Tool for Class and Learning Materials Design

The authors of this study have developed a support tool for the design of class and learning materials (COEdit). One of the main goals of this tool is to edit corresponding items in the two outlines (class and contents) through an evaluation and revision process. If a new line is added, then the information is linked to the line. Since linked information is simultaneously copied, the items that have the same linked information can also be simultaneously edited. This is a convenient function for the revision of the class outline and the contents outline.

4. Procedure of Subject Report Revision

The revision of the subject report was performed through the procedure shown below. Step2 *(2)*, *(3)*, and *(4)* were conducted during the meeting with the student assistants.

(1) Evaluation of subject results (the degree of achievement)
When the class is completed, the results of the subject reports are evaluated.
(2) Check the terms that are required for the subject
Based on the class outline, the aim of the class is confirmed, and the terms that are required to achieve the goal are checked. These are shared between the teachers and the student assistants.
(3) Create the class outline
The idea, based on the information acquired by the procedure *(2)* and the existing subject reports, is shared by the teachers and the students.
(4) Create the contents outline
The contents outline is created based on the class outline in the procedure (3).
(5) Create the learning materials
Based on the class outline and the contents outline, student assistants create the contents of a subject report page.
(6) Evaluate the created learning materials
Teachers evaluate the subject report page created by the students. They also verify whether explanations or expressions differ widely from what they had originally envisioned, and if required, they correct them.

5. Results

A “Programming 4” course was offered for second-year students in the department of human information systems, and the contents included Java application programming. The subject report of “Programming 4” was revised in the procedure shown above. The learning materials of the course were developed by student participation [3] and the class outline and the contents outline were created before content development. Two students, both graduate students, participated in the development process. A practice result is shown along with the procedure described above.

In Step *(1)*, the results of the subject report in 2010 were verified with no significant problems. In regard to Steps *(2)*, *(3)*, and *(4)*, teachers and student assistants held a meeting and discussed the overall goal of the class, the terms, and the questions from the former subject reports. This was achieved by displaying the class outline on a computer using the COEdit tool. To allow the difficulty of the questions to be comparable to the former ones, small revisions were made to the older questions. The ideas of the questions were recorded on the clause of the “subject report” in the class outline.

Next, the information on the contents corresponding to the modified part in the class outline was added to the contents outline. Again, the COEdit tool was used for editing the class outline and the contents outline. The time spent for the meeting of Steps *(2)*, *(3)* and *(4)* was approximately from 20 to 30 minutes per subject for each class. In addition, three patterns of the subject report were created for each class.

In regard to Step *(5)*, the students created the subject reports based on the class outline and the contents outline. The contents of the subject report were created and written in HyperText Markup Language (HTML). Finally, the contents and the source codes were submitted to the teacher.

For Step *(6)*, the teacher evaluated the contents of the subject reports. Only minor details such as the notation of the variables in the Java program or certain expressions were pointed out by the teacher.

After completing the subject reports, the students were interviewed about their creation of the subject report contents. In regard to the ease of creating the contents and the workload, there were opinions such as “there were enough directions to create the contents,” “it is very clear what kind of thing should be made,” and “I thought that the gap between the teacher’s idea and the students’ idea was decreased by the meeting.” In addition, there were the following positive opinions: “the knowledge which was uncertain until now could also be studied making the program of a subject, and an understanding was able to be deepened,” and “although having investigated about Java was serious, making the contents was easy.”

According to these positive reactions, it appears that the students were able to also acquire knowledge about Java programming, and the workload was reasonable for them.

6. Discussion

We have revised the design of the class activities and the learning materials in connection with the subject reports. In addition, the contents of the subject reports were developed through student participation and the learning materials developed in this study were used in actual classes.

In practice, reasonable workloads were applied to both the teachers and student assistants. In this case, the teachers were required to check the design of the class and propose the idea of a subject report during the meeting with the student assistants in Steps *(2)*, *(3)*, and *(4)*. However, it appears that the teachers’ workload was sufficiently eased since they left the entire creation process up to the student assistants. On the other hand, the contents of the subject reports created by the student assistants were actually used in the class. Although the student assistants did not devise unique learning materials, they did complete their work efficiently and acquired some aspects of learning materials development. Therefore, it is suggested that the proposed method was effective and the teachers and student assistants were convinced during this practice.

In addition, the function in that teachers and students could share and edit the class outline and the contents outline was also very useful and such functions will be implemented to the COEdit tool in the future. Finally, since the procedure introduced in this practice was specialized for the subject report revision; this procedure should be extended as a future revision method applicable to general classes that focus on learning materials revision.

7. Conclusion

In this study, the subject report revision procedure of existing class design and learning materials was proposed based on the class design and the learning materials development method following the ID concepts. By working through this procedure through student participation, high-quality learning materials that could be efficiently used in classes were developed. In the subject report revision process, the class and the learning materials design support tool COEdit proved to be helpful and time efficient for the teachers as well as the students. Perhaps in the future, the subject report revision procedure would be extended as a method applicable to general classes that focus on learning materials revision.

Acknowledgements

This research was supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 22500938, 2010.

References

[1] Dick, W., Carey, L., and Carey, J.O. (2001) Systematic Design of Instruction (5th ed.), Addison-Wesley Educational Publishers Inc.

[2] Morrison, G. R., Ross, S. M., and Kemp, J. E. (2004) Designing Effective Instruction (4th ed.), John Wiley & Sons Inc.

[3] Sasaki, S., Arai, M., Furukawa, F., and Watanabe, H. (2008) Web-based Learning Material Development with Less Experienced Staffs' Participation, Proc. of International Conference on Computers in Education 2008 (ICCE 2008), 445–450.

[4] Sasaki, S., and Watanabe, H. (2010) Development of Class and Learning Materials Design Tool based on Instructional Design, Proc. of International Conference on Computers in Education 2010 (ICCE 2010), 265–269.

[5] Nakai, T., Yamasato, T., Nakajima, H., and Okada, T. (2003) e-Learning Handbook, Manahouse, Japan.

[6] Jung, I., Kubota, K., and Suzuki, K. (2008) Instructional Design by the OPTIMAL Model–Effective Method for Blended Type e-Learning, Tokyo Denki University Press.

Learner's Behavior Detection System for Mentoring in e-Learning

Arisa Tanaka[a*], Naoki Matsumoto[b] & Takahiko Mendori[a]
[a]*School of Information, Kochi University of Technology, Japan*
[b]*Graduate School of Engineering, Kochi University of Technology, Japan*
*130350w@ugs.kochi-tech.ac.jp, mendori.takahiko@kochi-tech.ac.jp

Abstract: In this paper, we describe learner's behavior detecting system for mentoring in e-Learning. It is important for learner to keep her/his motivation in e-Learning. Therefore, existing e-Learning system and LMS (Learning Management System) is to realize mentoring for learner based on learning history. However, past mentoring cannot cover in learner's learning. Then, we focus on the learner's behavior in her/his learning. We propose and develop a learner's behavior detection system for mentoring.

Keywords: Mentoring, Learner's behaviors, Learner's Behavior Detecting System

Introduction

In recent years, WBT(Web Based Training) and other e-Learning has been introduced in enterprises and school training. WBT does not need special devices. Thereby, learner can learn on any occasion on e-Learning. However, learning in WBT is difficults for learner. Learner feels difficult to keeping her/his motivations for learning. There is a problem on WBT, that learner drops out of learning on the way. There is support called mentoring to solve this problem [1]. Mentoring is a kind of learning support method. Person who perform mentoring is called mentor. Mentor communicates with e-mail or chat to her/his supports learner's learning.

In face to face lesson, teacher understands immediately learner's intention. Teacher can guess learner's intention from her/his behavior and facial expression. As a result, teacher can flexibly consider efficient learning. For example, if learner yawned, teacher would guess learner has not concentrate. Teacher can induce learner's consideration by her/his questions. In e-Learning lesson, mentor cannot understand correctly learner's intention. As a result, mentor cannot perform mentoring according to learner's situation. The information that can be recorded for mentoring on existing e-Learning is only learning history and send/receive e-mail history.

Prior studies have been proposed the recording system of learner's intention by mouse click [2]. In that study, learner uses the system on learning in e-Learning. If learner cannot understand learning contents, she/he clicks the link. The system records a timing that learner clicked the link. Mentor can understand learner's status of learning from the recorded information by the system. The studies had been shown two problems of the proposed indication method. One problem is real-time properties of the recorded information is low. If learner has doubt to learning contents during learns, she/he wants to indicate it to mentor. Learner is necessary clarify the reason why that she/he has doubt. As a result, learner discontinues clicking. Even if learner clicks the link, she/he cannot indicate timing of doubt.

Another problem is learner cannot concentrate to learning because her/his consideration is induced to clicking. If learner hesitates to clicking, she/he misses learning contents at that time. These problems has not been solved yet.

1. Learner's behavior Detecting System

We propose a method of using learner's behavior to solve these problems. We develop a system that detects and records learner's behavior. The system records image of learner's behavior with Web camera and performs easy image processing. The system detects learner's behavior from results of the image processing. Table 1 summarizes the correspondence of the behavior to record and the specific meaning.

Table 1. The frequency of learner's behavior on each system

body part	behavior	specific meaning of behavior
Facial movement	Head nod	agree, consents, understands, sympathize, satisfy
	Lean of head	doesn't understand, Uneasiness, Hesitation, Worry, Distrust
	Head shake	denial, refusal, mistake
Hand movement	Thumbs up	agree, consents, understands, sympathize, satisfy
	Paper	request the stop
	Hand waving	doesn't understand, Uneasiness, Hesitation, Worry, Distrust

If mentor can understand learner's behavior, we improve learning effect of e-Learning. However, an effective method to understand learner's behavior does not exist. We design and implement of such a system which detects learner's behavior and records detection result. We increase opportunities of performing mentoring. The record of learner's behavior improves the interactivity of the mentor and the learner on e-Learning. As a result, mentor can perform more effective mentoring.

In this research, we use only Web camera. Because teacher does not uses special devices in face to face lesson. The purpose of this study is to understand learner's situation from same information as face to face lesson. Therefore, our propose method don't use any special devices.

2. Prototype System

2.1 Overview of Prototype System

Figure 1 indicates that configuration of our system. The system is composed of the learner's behavior detection system and the enhanced Moodle.

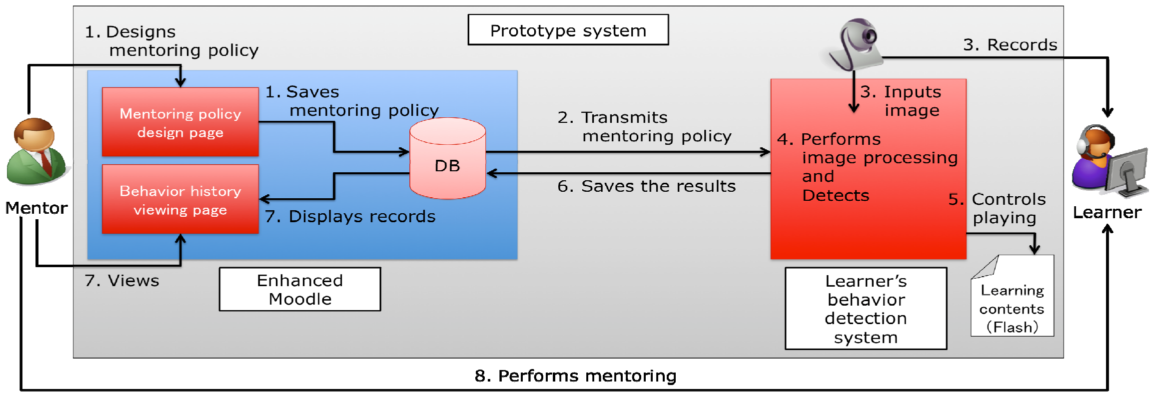


Figure 1. Overview of the prototype system

The learner's behavior detection system is system to detect learner's behavior. The system performs image processing to the image from Web camera. The system detects learner's behavior from the result of image processing using OpenCV [3]. The system saves the result of detected learner's behavior to the database of the enhanced Moodle.

The enhanced Moodle is system that we developed by adding a module to Moodle [4]. We enhanced the mentoring policy design page and the behavior history viewing page to Moodle. The mentoring policy design page is module of Moodle which to design freely mentoring policy. Mentoring policy is timing of mentoring and the method of control playing learning contents. The enhanced Moodle saves mentoring policy to internal database. The behavior history viewing page is module of Moodle which to display learner's behavior history. The enhanced Moodle loads the recorded learner's behavior from internal database. The enhanced Moodle displays learner's behavior history by graph form using pChart [5]. Mentor performs mentoring based on learner's behavior history.

1.1 Overview of the Learner's Behavior Detection System

Figure 2 indicates that learner's active screen with this system. Learner sees two windows with Web browser while her/his learns. Two windows are the learning content window and the mentoring window. The learning content window plays learning contents. The mentoring window gives feedback to learner. We have included a mentoring avatar on the mentoring window. There are three roles of the mentoring avatar.

- tells learner what the system has detected behavior
- shows the message, so as to shift status
- performs mentoring based on mentoring policy

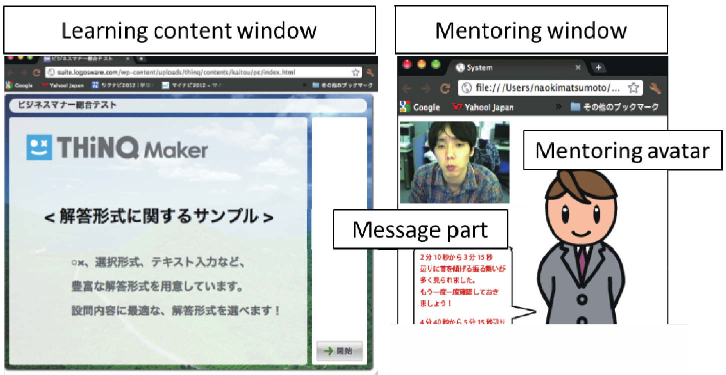


Figure 2. Learner's active screen with this system

If this system records learner's behavior, the motion of mentoring avatar changes simultaneously. Figure 3 indicates the motion of mentoring avatar. A left area of the mentoring window is a message part. In the message part, the mentoring avatar shows mentoring message based on mentoring policy.

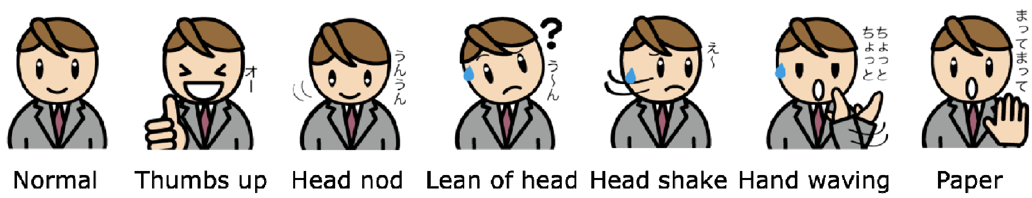


Figure 3. Motion of the mentoring avatar

2. Preliminary Experiment

2.1 Purpose

The purpose of this experiment is to elucidating the predominance of recording method from behavior. We compare the method of prior research with our method. In this experiment, we confirm two items. Another is frequency of learner's behavior. We confirm it from our system and the image from Web camera. Other is learner's load level that learner feels. We analyze it from questionnaire results.

2.2 Overview

We developed three systems for this experiment. System 1 is system that it record clicking text link. If learner's intentions correspond with text link, she/he clicks it. The system records kinds and timing of clicking text link. System 2 is system that it record clicking image button. If learner's intentions correspond with image button, she/he clicks it. The system records kinds and timing of clicking image button. System 1 and system 2 record learner's intentions from mouse click. System 3 is system that it detects learner's behavior. System 3 detects facial and hand movement as learner's behavior. System 3 records a result of detected learner's behavior and a timing of behavior. Each system respectively can record six kinds of learner's intentions. The recording items on each system corresponds to behavior in table 1. Figure 4 indicates that each screen of system 1 and system 2.

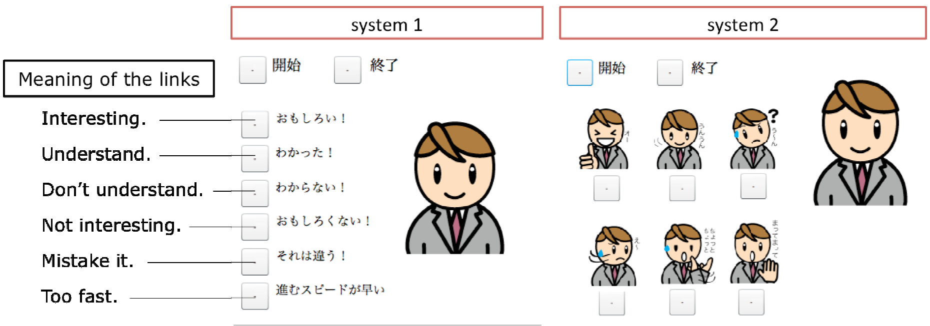


Figure 4. Each screen of system 1 and system 2

We conduct this experiment to two subjects. The subjects uses each above- mentioned three system. We record the subject with video camera during this experiment.
The procedure of this experiment is as follows.

1. Explain the outline of this experiment to subjects
2. Start playing learning contents and recording the subject with video camera
3. Make the subjects fill in questionnaire after learning

2.3 Result

Table 2 summarizes that frequency of learner's behavior on each system. We understood that the learner's behavior detection system be able to record learner's behavior. We confirmed operation of the learner's behavior detection system from this experiment.

All subjects described "System 1 or 2 is better than system 3". However, one of subjects described “I felt satisfaction to indicating intentions by system 3 when the detecting accuracy was good. I felt that system 1 and system 2 better than system 3, because system 1 and system 2 accurate detects intentions”. All subjects described "I wanted to freely behave during experiment". We understood that all subjects felt the load to indicating.

Table 2. The frequency of learner's behavior on each system

System 1							
Subject	Total	Interesting	Not interesting	Understand	Don't understand	Mistake it	Too fast
A	10	2	0	2	4	0	2
B	11	0	2	1	7	1	0
System 2							
Subject	Total	Thumbs up	Head shake	Head nod	Lean of head	Hand waving	paper
A	14	1	1	1	9	0	2
B	13	2	2	2	5	1	2
System 3							
Subject	Total	Thumbs up	Head shake	Head nod	Lean of head	Hand waving	paper
A	14	2	0	2	7	0	3
B	10	0	0	3	7	0	0

3. Discussion

We compared the records of system 3 with the image from video camera. As a result, we confirmed misdetection of the system. For example, subject's behavior is "paper", nevertheless, the system detects "thumbs up". One subject described "I felt shyness when the system misdetects". In other words, misdetection involves to subjects. Therefore, we need to consider and improve the detecting accuracy enhancement.

The subjects tended to restricted herself/himself to behavior on system 3. The subjects described "If I freely behave, system 3 detects it as intention". We need consider the recording environment that does not restrict the learner's behavior.

We did not be able to elucidate the effectiveness of our system from result of this experiment. Because the number of subjects is too few. Therefore, we need to conduct the evaluation experiment to confirm the effectiveness of our system.

4. Conclusion

In this paper, we proposed the learner's behavior detecting system for mentoring in e-Learning. At first, we described the learner's behavior in e-Learning and its meaning. Next, we developed the learner's behavior detection system. Furthermore, we conducted evaluation of the system and its evaluation. Finally, we discussed our system based on the result of the evaluation. Thereby, we clearly showed that the problem of our system and the evaluation. In future work, we are going to design the evaluation experiment moreover prove the effectiveness of our system.

References

[1] Takeshi, MATSUDA., Nobuyuki, HONNA., & Hiroshi, KATO. "Development of e-Mentoring Guideline and its Evaluation." *The Educational Technology Research Journal 29*.3 (2005) : 239-250. (in Japanese).

[2] Seiji, ATARASHI., & Kozo, SUGIYAMA. "Research and Development of Education Support System Activating Interactions Between a Teacher and Students." *IEICE Technical Report ET2001-111* Mar. 2002: 87-94. (in Japanese).

[3] "*Intel Open CV*." Retrieved 1 February 2012 <http:www.intel.com/technology/computing/opnecv/>.

[4] "*Moodle Docs*." Retrieved 1 February 2012 <http://moodle.org/ja/>.

[5] "*pChart a PHP Charting library*." Retrieved 1 February 2012 <http://pchart.sourceforge.net/>.

Development of Moodle Plug-ins that support SCORM 2004

Yosuke MORIMOTO[a*], Kiyoshi NAKABAYASHI[b], Hidenori SUGIYAMA[a] & Junji SHIBASAKI[a]
[a]*Center of ICT and Distance Education, the Open University of Japan, Japan*
[b]*Faculty of Information and Computer Science, Chiba Institute of Technology, Japan*
*morimoto@ouj.ac.jp

Abstract: We have proposed a new architecture for e-learning systems, named ELECOA, and developed Moodle plug-ins that support it. ELECOA is an architecture that fulfills both function extensibility and content interoperability, and the Moodle plug-ins we developed can run content based on ELECOA. We have also implemented the functions of SCORM on ELECOA, so the plug-ins can also run SCORM content.

Keywords: e-learning system, courseware object, SCORM, Moodle

1. ELECOA: An Extensible Architecture for e-Learning Systems

As evidenced by the open educational resources (OER) movement of recent years, there is a lot activity related to the promotion of sharing, distributing, and reusing learning content. Interoperability and reusability of learning content is key in such activities, but it is widely recognized that conventional e-learning systems and content specifications lack sufficient function extensibility. This is often because newly added functions on existing e-learning platforms may conflict with existing learning content, or it can simply be too difficult to add new functions. Even if new functions can be added without any problems, the content will not be interoperable. Sharable Content Object Reference Model (SCORM), which is the de facto standard of e-learning, has the same problem. In response to this problem, we have proposed a new architecture for e-learning systems, named ELECOA (Extensible Learning Environment with Courseware Object Architecture), which fulfills both function extensibility and content interoperability requirements (Nakabayashi & Morimoto, 2012; Nakabayashi, Morimoto, & Hada, 2010). Figure 1 shows the architecture of ELECOA. We propose using a courseware object, which is a program module that implements various educational functionalities, as a layer rather than combining it with the platform as in the conventional architecture. Content are run by the courseware objects which are assigned to them, and introducing new functions or extending existing functions is done by adding new courseware objects. Since this addition does not affect existing courseware objects, function extensibility can be assured. Interoperability can also be assured by distributing courseware objects with the content.

We have established some rules related to content structure, content packaging format, communication method between courseware objects, etc. in order to achieve the ELECOA framework. For example, ELECOA content must be structured hierarchically, which is a common structure for e-learning content. Courseware objects are assigned to respective nodes (root, branch, and leaf) of a content tree. A courseware object assigned to a node is responsible for the behavior of the sub-tree under the assigned node. This makes it possible to implement different pedagogical strategies in different sub-trees and to distribute only specified sub-trees of content. More details are described by Nakabayashi and Morimoto (2012) and Nakabayashi, Morimoto, and Hada (2010).

Although the concept of ELECOA is independent of any programming language, we do need to choose one for use in the actual system. We chose PHP to implement the courseware objects and the platforms for running them. Courseware objects are implemented as PHP classes and are instantiated when the content is launched. In order to investigate the feasibility of ELECOA, we implemented the SCORM 2004 Sequencing and Navigation Specification (Advanced Distributed Learning, 2006) on it. Results showed that it is possible to implement a set of courseware objects that are fully compliant with SCORM 2004 SN.

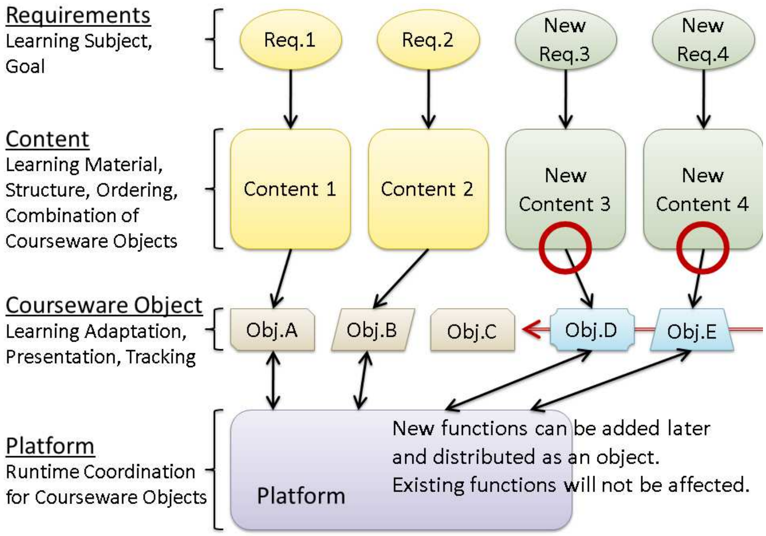


Figure 1. The architecture of ELECOA.

2. Implementation of ELECOA on Moodle

Because courseware objects are independent of platform, they can be used on any e-learning system that supports ELECOA. We have developed Moodle plug-ins that support ELECOA. As mentioned above, we also have developed courseware objects that support SCORM 2004. This means that the plug-ins can support SCORM 2004 by using the courseware objects.

There are three plug-ins: an “activity module”, a “block” (for grade reports), and a “course format”. The latter two are dependent on the activity module. The activity module, named mod_elecoa, works in the same way as mod_scorm (the built-in SCORM module of Moodle) from the point of view of the user. mod_elecoa can launch ELECOA content that are run by the courseware objects assigned to them. Users can thereby run SCORM content (both 1.2 and 2004). When the content is uploaded and added to a course as an instance of mod_elecoa, its manifest file is converted to that of ELECOA. Although mod_scorm itself is able to run SCORM content, mod_elecoa cannot; it is the function of SCORM-compliant courseware objects that are pluggable.

Figure 2 shows a screenshot while running a content on mod_elecoa. The menu area is in the upper right and the content area is in the lower right. If the content is a SCORM content (or more accurately, an ELECOA content that uses SCORM-compliant courseware objects), navigation menus are shown in the menu area and an SCO or an asset is shown in the content area. The structure of the content and the titles and current status of the nodes are shown in the TOC.

An early version of mod_elecoa (Morimoto, Nakabayashi, Sugiyama, & Shibasaki, 2012) had performance problems that were mainly related to the speed of loading Moodle core libraries. In that early version, the Web browser communicated with the Moodle core several times per user action, e.g., clicking the "Continue" button. Each communication took several hundreds of milliseconds, which degraded the overall performance. Although we have improved the performance of mod_elecoa by reducing the number of communications with the Moodle core, further improvement is still required.

The developed plug-ins are compliant with Moodle 2.x and are distributed under the GPL via http://elecoa.ouj.ac.jp/ (Note: currently this Web site is available in Japanese only). "ELECOA Player", which is a standalone e-learning system that supports ELECOA, is also distributed under the Modified BSD License via this Web site.

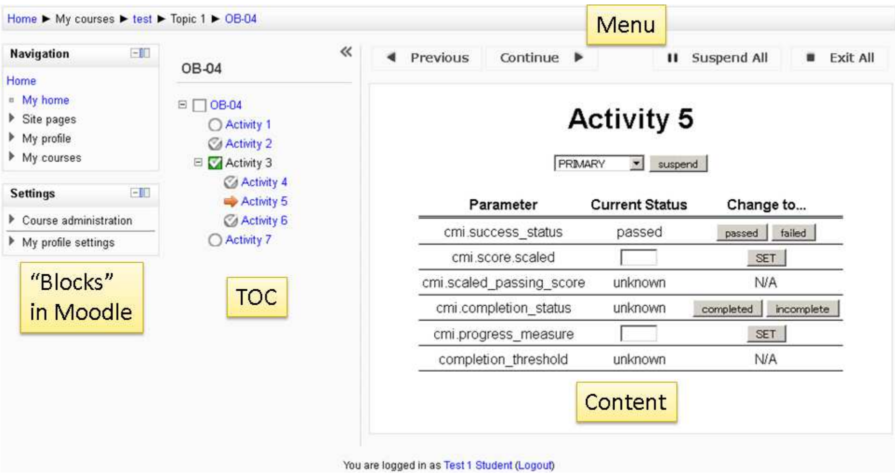


Figure 2. A screenshot of mod_elecoa.

3. Conclusion

In this paper, we described the concept of an extensible architecture for e-learning systems named ELECOA. This architecture features "courseware objects" that are program modules used to implement various educational functionalities. We were able to implement courseware objects which are fully compliant with SCORM 2004 SN. We also developed a set of Moodle plug-ins that are compliant with ELECOA. By using the courseware objects for SCORM, these plug-ins are also compliant with SCORM.
As of this writing, mod_scorm supports very few functions of SCORM 2004. The developed plug-ins should therefore be quite useful for disseminating SCORM and Moodle.

References

[1] Advanced Distributed Learning (2006). SCORM 2004 3rd Edition Specification.
[2] Nakabayashi, K., & Morimoto, Y. (2012). Design and Implementation of Object-oriented Architecture for Extensible Learner-adaptive Self-learning System (in Japanese). *Transactions of Japanese Society for Information and Systems in Education, 29(2)*, 97-109.
[3] Nakabayashi, K., Morimoto, Y., & Hada, Y. (2010). Design and Implementation of an Extensible Learner-Adaptive Environment. *Knowledge Management & E-Learning: An International Journal, 2(3)*, 246-259.
[4] Morimoto, Y., Nakabayashi, K., Sugiyama, H., & Shibasaki, J. (2012). Development of a Utilitarian UI for a Moodle Activity Module That Supports SCORM 2004 (in Japanese). JSiSE Research Report, 27(1), 11-16.

Improve Students' Reading by Taking a Question-Based Learning Process on E-books

Li-Hsiang HSIAO*, Hsin LEE, Shing-Tsaan HUANG, Liang-Yi LI, and Gwo-Dong CHEN
Department of Computer Science and Information Engineering, National Central University, Taiwan
**shair78227@gmail.com*

Abstract: This study developed a question-based learning process embedded on an e-book to help students' textbook reading. We designed three reading phases: preview, reading, and review according to the reading strategy. In each phase, students were assisted by different questions. To evaluate the effect, we used this e-book in an experiment involving twenty-five students studying on a textbook content of computer science. Our research objectives are to test the effects of the system on academic performance and to evaluate whether the system use affects learner's motivation and active reading.

Keywords: questions; reading; e-book; learning process

Introduction

Questions in a textbook are generally provided in each chapter. An author may place preview questions at the beginning of a chapter to tell the reader what they will encounter and review questions at the end to improve their memory of the chapter. However, there are some issues with this book editing style: First, a chapter scope may be too large, and readers cannot remember excessive knowledge; as a result, they have difficulties to answer those questions. Second, different reading procedure may require different types of question. Structured questioning strategies significantly affect text memory and text learning [1]. On the other hand, teachers may have a requirement to ask their own questions to students. Even students may also want to bring up their questions. Thus, how to put the suitable question in each reading procedure is the second issue.

Therefore, this study attempts to address the above issues by adding a question-based auxiliary tool into a textbook. In addition, because of e-book's controllable characteristics and its increasing popularity, we digitally build the auxiliary tool. In summary, this paper brings up a question-based learning process educational e-book system for a solution.

1. Related work

1.1 Questions in Leaning

A teacher ask question in a classroom to develop interest and motivate students, to evaluate students' preparation, to develop critical thinking skills and inquiring attitudes, to review and summarize, and to nurture insights by exposing new relationships [2]. Some researches summarized that high-quality learning includes mastery of content, understanding of concepts, and development of explicit strategies for asking good questions and exploring new ideas [3].Answering questions during a lecture can encourage students to engage some cognitive processes like organizing what they have learned, mentally integrating it with other knowledge, and a need to pay attention in what they are learning [4].

1.2 Reading with Questions

This study aims to use questions to help students reading. Questioning is a cognitive scaffolding assistance to stimulate and guide learning [5]. It is undoubted that question induce meaningful processing facilitate reading [6]. A research found that reading guided by questions is better in immediate recall and delayed recall than careful reading and re-reading without questions [7].Some educators thought that students should be encouraged to develop as independent readers by asking their own questions in order [8]. Self-generation of questions during study can lead to improved performance on a test of comprehension [9].

2. System Design

2.1 Question-based Learning Process in E-book

Our study divided question-based reading process into three preview-read-and-review phases.A "QA-format" e-book was designed for users to be able to add question dynamically with scaffolding for a prescribed reading process. The system has three steps in a reading phase: preview, read, and review.

In the preview phase, students can read questions about the knowledge structure of the section they will meet.

In the reading phase, students can see questions by the side of textbook content in each page. These questions were asked by other readers who has read this page before. The system also provides a function to add comment to the existing questions. The purpose this step is to improve students' motivation and chance of reconsideration.

Finally, the review phase provides questions to evaluate what they have digested. Similar to the preview step, questions were created by teachers. In this step, a student can reflect on how well they have understood the author's ideas.

2.2 System Description

For portability, a "QA-format" e-book is web-based application software. Students can read the e-book on any device that supports network and web browser. The e-book contains two kinds of screen: home page and content page.

Home page (Fig. 2) lists all names of chapters and sections in the book and presents the contents in a hierarchical tree structure. In the content page (Fig. 2), the screen is divided into four parts: tool bar, question area, comment area, and textbook content.

Tool bar is on the top of page, it shows the user identity, number of questions, button of creating a new question, and button to look preview questions again. The remaining areas outside the toolbar from top to bottom are questions, contents, and comments.

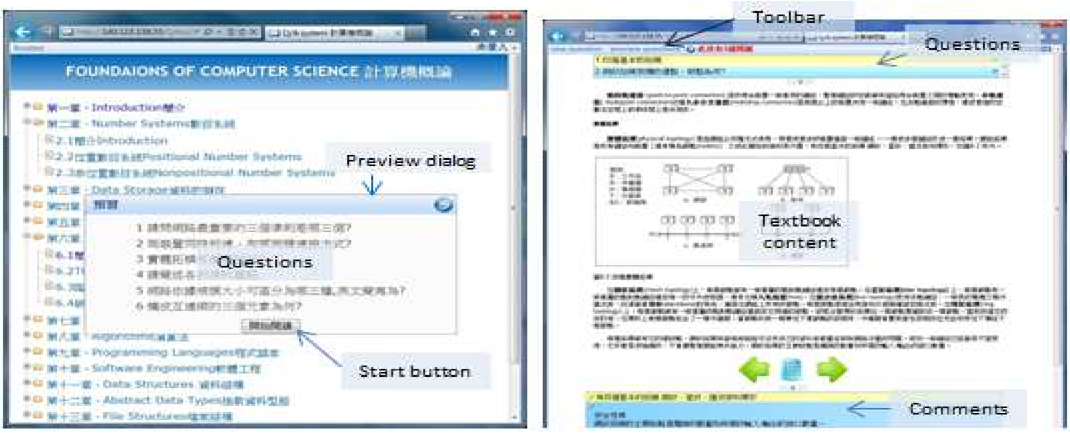


Figure 1. Home Page and Content Page

3. Conclusion

About question-based reading: Through the interviews and observations, we have seen that integration of e-book and questions successfully increased students' motivation of reading, and students gave positive evaluation of this reading approach.

Future "QA-format" e-book design: After this research, we found the e-book system could be improved in the future. First, self-questioning during reading is an important issue. Participated students suggest that adding a more complete design would increase their motivation of asking question. Many studies have mention that asking question by students' own could facilitate more effective instructional procedures [10]. Second, there would be needed a teacher editing interface for general use. Third, some students indicate that they are more familiar with the paper book than e-book.

References

[1] [1] T. M. Gunn (2008). The Effects of Questioning on Text Processing *Reading Psychology*, vol. 29, 405-442.

[2] Kathleen Cotton (2001). Classroom Questioning *School Improvement Research Series*

[3] Wendy Martin, Andy Gersick, Hannah Nudell and Katie McMillan Culp (2002). An Evaluation of Intel® Teach to the Future *Education Development Center/Center for Children and Technology*

[4] Julie Campbell and Richard. Mayer (2009) Questioning as an Instructional Method: Does it Affect Learning from Lectures? *Applied Cognitive Psychology*, vol. *2*, 745–759.

[5] Moll, L. C.. (1990) *Vygotsky and Education: Instructional implications and applications of socio-historical psychology.*

[6] Anderson, R.C., and Biddle, W.B.On (1975). asking people questions about what they are reading. In G.H. Bower (Ed.), *The psychology of learning and motivation, vol. 9* , pp. 9–129.

[7] Eleanor Holmes (1931). "Reading Guided by Questions versus Careful Reading and Re-Reading without Questions," *The School Review, Vol. 39* , 361-371.

[8] Bernstein, Susan Lynne. (1973).The effects of children's question-asking behaviors on problem solution and comprehension of written material Unpublished doctoral dissertation, Columbia University,

[9] Marli E.D.A Andre and Thomas H. Anderson(1978 - 1979). The development and evaluation of a self-questioning study technique,*Reading Research Quarterly*, Vol. 14, No. 4, 605-623.

[10] Dennis R. Knapczyk and Gary Livingston, (1974). "The effects of prompting question-asking upon on-task behavior and reading comprehension", *Journal of applied behavior anaysis*, vol. 7, 115-121.

Adapting Guidance and Externalization Support Features to Program and Algorithm Learning Support Environment

Satoru KOGURE [a*], Makoto OKAMOTO[b], Yasuhiro NOGUCHI[a], Tatsuhiro KONISHI[a], Yukihiro ITOH[c]
[a]*Faculty of Informatics, Shizuoka University, Japan*
[b]*Graduate School of Informatics, Shizuoka University, Japan*
[c]*Shizuoka University, Japan*
*kogure@inf.shizuoka.ac.jp

Abstract: In this paper, we proposed novel learning support environment for understanding programs and algorithm by emphasizing three worlds: programs world, domain world and operations world. To lead learners deeply understand programs and algorithm, we suggest "guidance feature" and "externalization support feature" in learning environment. Hereby, we guess that learners could bridge their gap between programs world and operations world by these features.

Keywords: learning support environment, program and algorithm learning, externalization

Introduction

In order to learn programs and algorithms on elementary programming education, the learners perform three typical learning procedures: understanding algorithm, understanding behavior of sample program, and writing code [1][2][3][4][5][6].

In our study, we focus on "understanding algorithm" and "understanding behavior of sample program". Then, we propose three worlds which learners should consider with learning programs and algorithm. These worlds are "programs world" (including abstract statements in program code), "target model world" (including concrete representation of the target problem) and "operations world" (including concrete operation sequences for the exact input data).

We believe that it is necessary for learners to understand the relationship between three worlds in order to understand a program and algorithm implemented in the program. There is our previous study[6] that emphasizes the relationship between three worlds. In the research, we categorized three programming levels and construct the programming learning support environment which has some suitable functions for each level's learners. Then we evaluated the environment in code reading. Based on the policy of this research, we construct extended programming learning support environment which has two novel features: **"guidance feature"** (to lead learners to perform suitable process at any time) and **"externalization support feature"** (to lead learners to abstract concrete operation sequences to a description of general procedure).

1. Learning Support Policy

It is relatively easy for programmers to understand the relationship between the domain world and operations world because the both of them are composed by concrete data However, it is difficult for learners to understand a difference between programs world and other two worlds. In program world, concrete data is abstracted, and operations with concrete data are structuralized.

We focus learners that <u>are able to reproduce the operation sequences for concrete data (L1)</u>. We design our system to lead learners to <u>be able to generalize operation sequences (L2)</u>, and to <u>be able to understand the relation of programs world and two other world (L3)</u> from (L1) level. We present programming learning support environment for learners that reached L1 level but does not reach L2 or L3 levels. Our proposed support environment has two novel features, "guidance feature" and "externalization support feature". "Guidance feature" is function that leads learners to perform a suitable process on the learning environment according to learner's understanding state. In order to realize "guidance feature", our proposed learning environment support to indicate guidance to learners based on learning scenario. "Externalization support feature" shows learners some candidates of the description externalizing some operation sequences, and leads learners to judge a correctness of the description by choosing correct one from some candidates. By these two features, learners deeply understand the relationship between program codes and operation sequences.

2. System Outline

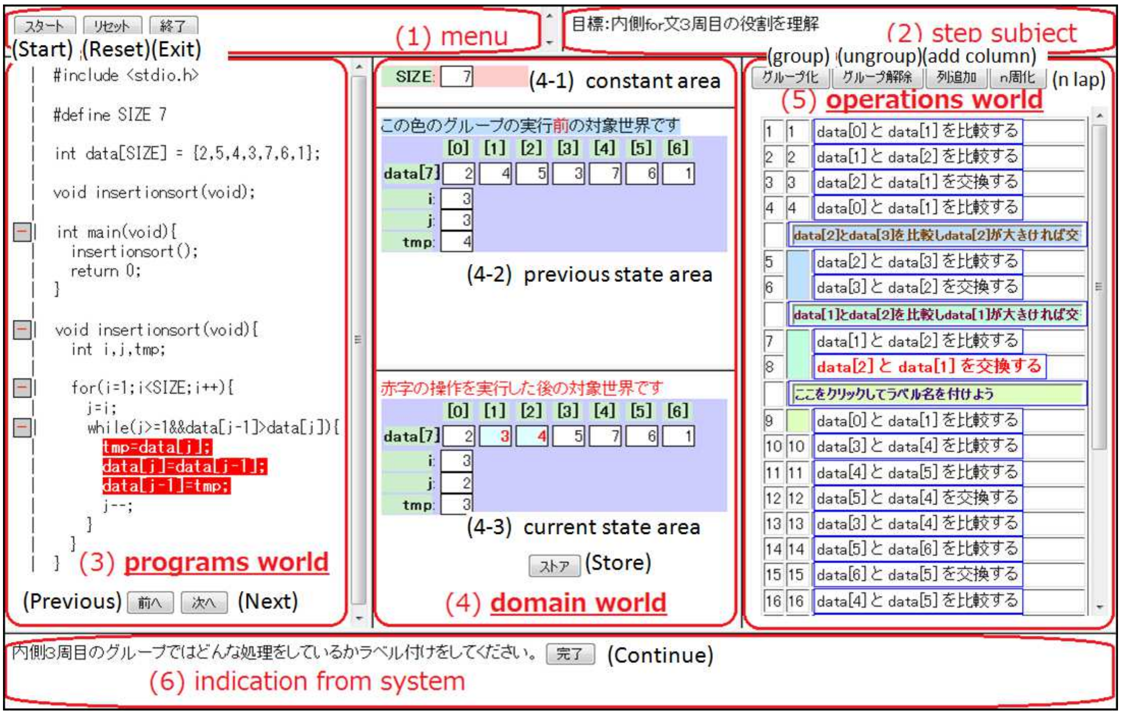


Figure 1. Interface overview.

Figure 1 shows the interface of our proposed learning support environment. Figure 1 (3), (4) and (5) show three worlds that we guess very important concept for programming and algorithm learning. Figure 1 (1) shows a menu of support environment, Figure 1 (2) shows a step subject at a learning process, and Figure 1 (6) shows the indication the system give to learners for suitable learning process according to learner's understanding state.

The system has nine functions for supporting learners. In a previous study[6], we implemented six functions: executing step by step (Figure 1 (3)), packing any block (Figure 1 (3)), storing each state of target world (Figure 1 (4)), making a group of operation sequences (Figure 1 (5)), writing a label made group of operation sequences (Figure 1 (5)), and displaying behavior of grouped operation sequences (Figure 1 (4-2) and (4-3)).

In this study, we proposed new three functions to realize "guidance features" and "externalization support feature". New three functions are **function of indicating guidance (Figure 1 (6)), function of judging correctness of written label, and function of externalizing n laps.**

In pre-experimental results [6], we inscribed that learner wants suitable guidance for learners understanding level in learning programs and algorithm to use learning support environment. The system can lead learners to learn along a prepared scenario by a teacher. The system can give each learner appropriate instructions for each learner's understanding state according to the scenario (**function of indicating guidance**). In order to judges a correctness of learners labeled description, the system can lead learners select the most similar alternative in content to description that learners label in operations world (**function of judging correctness of written label**). In order to bridge a gap between program world (abstract world) and operations world (concrete world), the system lead learners to increase abstract level of operations world using **function of externalizing n laps**. Due to limitations of space, a detailed discussion of these functions is not possible here.

3. Conclusion

We proposed novel learning support environment for understanding program and algorithm. We emphasized the three worlds: programs world, domain world and operations world for understanding program and algorithm. So, we implemented nine functions. Using this support environment, learners will grow understanding and become L3 understanding level. In the near future, we bring implementation of our proposed learning support environment to completion and evaluate it.

Acknowledgements

This study was supported by Japanese Grant-in-Aid for Scientific Research (B) 24300282.

References

[1] Butz, C.J., Hua, S., & Maguire, R.B., (2006). A web-based bayesian intelligent tutoring system for computer programming, *Journal of Web Intelligence and Agent Systems*, 4(1), 77-97.

[2] Fossati, D., Eugenio, B. D., Brown, C., & Ohlsson, S. (2008). Learning linked lists: experiments with the iList system, *Proceedings of the 9th International Conference on Intelligent Tutoring Systems*, 80-89.

[3] Gabor, T., (2009). Algorithm visualization in programming education, . *Journal of Applied Multimedia*. 4(3), 68-80.

[4] Nakahara, T., Konishi, T., Kogure, S., Noguchi, Y., & Itoh, Y. (2009). Learning environment for algorithm and programming where learners operate objects in a domain world using GUI. *Proceedings of the 17th International Conference on Computers in Education*, 59-66.

[5] Noguchi, Y., Nakahara, T., Kogure, S., Konishi, T., & Itoh, Y. (2010). Construction of a learning environment for algorithm and programming where learners operate objects in a domain world', *International Journal of Knowledge and Web Intelligence*, 1(3/4), 273-288.

[6] Satoru, K., Makoto, O., Yasuhiro, N., Tatuhiro, K., & Yukihiro, I. (2011). Programs and algorithm learning environment by visualizing relations among program codes, operations and world model, *Proceedings of 19th International Conference on Computers in Education*, 302-306.

Schema-based Scaffolding for Creating Presentation Documents

Yasuo SHIBATA[a*], Akihiro KASHIHARA[a] & Shinobu HASEGAWA[b]
[a] *Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan*
[b] *Research Center for Distance Learning, Japan Advanced Institute of Science and Technology, Japan*
*shibata.y@uec.ac.jp

Abstract: In order to properly create presentation documents as research activity, it is necessary to get and accumulate experiences in composing the semantic structure that represents what to present and how to sequence the contents presented. However, it is not easy for novice researchers because they have fewer experiences in creating presentation documents. This paper proposes scaffolding for composing the semantic structure of presentation documents with presentation schema that is typical semantic structure embedded in the presentation documents accumulated in a research group. The results of a case study suggest that the schema-based scaffolding contributes to creating presentation documents.

Keywords: Presentation Schema, Scaffolding, Presentation Documents, Research Activity

Introduction

This paper addresses the issue of how to develop skills in creating presentation documents (P-documents for short) consisting of slides in a research group. Creating P-documents involves composing semantic structure that represents what to present and how to sequence the contents presented [1,2]. However, it is hard particularly for the novice researchers who have fewer experiences in composing the semantic structure from their research contents.
In this paper, we introduce a scaffolding method for composing the semantic structures of P-documents with presentation schema (P-schema for short) that is a typical semantic structure of P-documents accumulated in the research group [1]. This paper also reports a case study, which has been conducted for ascertaining whether P-schema contributes to composing better semantic structure and better P-documents.

1. Skill Development for Creating Presentation Documents

In creating a P-document as research activity, it is necessary to segment the research contents into several parts to compose the corresponding slides and to sequence the slides so that the presentation can be readily understood by the audience [1,3]. Such segmentation and sequencing specifies the semantic structure the P-document embeds, which represents what to present and how to sequence the contents presented [1]. The semantic structure tends to be common among the P-documents created in a research group. The common structure could be viewed as heuristics for creating P-documents, which is unique to the group [3].

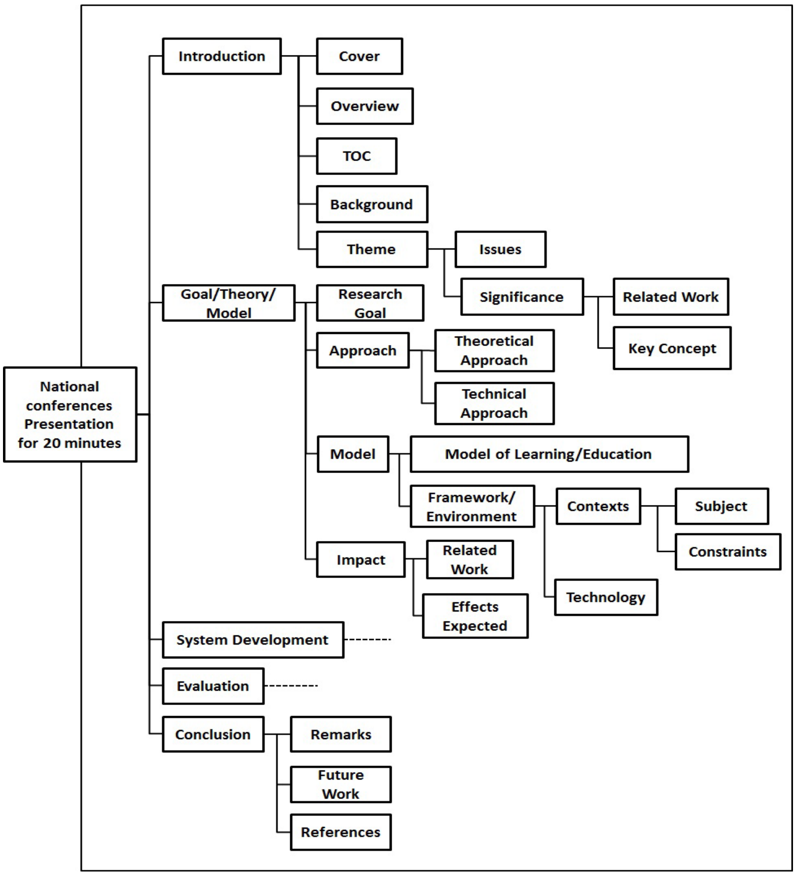


Figure 1: P-schema representing P-documents for national conferences.

On the other hand, the novice researchers have difficulties in segmenting and sequencing the research contents since they do not have a full understanding of the heuristics. In order to develop skill in creating P-documents, it is important to accumulate experiences in composing semantic structures to be embedded in the documents. However, the novice researchers have fewer chances of creating P-documents.

We have accordingly proposed scaffolding for composing the semantic structures of P-documents with P-schema that is typical semantic structure of P-documents accumulated in the research group [1]. P-schema could be defined by the expert researchers [2]. In this schema-based scaffolding, the novice researchers are allowed to compose the presentation structures in authentic research contexts of creating and learning P-documents.

Figure 1 shows P-schema that is defined from P-documents for undergraduate thesis in our research group. In the context of creating a P-document, a novice researcher could segment and sequence the research contents according to the P-schema. In the context of learning a P-document an intermediate/expert researcher created, he/she is also allowed to relate the slides and the P-schema to compose the semantic structure the document embeds. In this way, the novice researchers could accumulate experiences in composing the sematic structures to develop the skill.

In order to ascertain whether P-schema could function well as a scaffold for composing semantic structure and P-document, we have conducted a case study with 4 novice researchers. Each subject was required to create P-documents for his/her undergraduate thesis with and without P-schema. After that, two expert researchers in our research group evaluated the P-documents. As the results, the experts evaluated that P-document created with P-schema had a better semantic structure than P-document created without P-schema, had an understandable sequence of the slides, and made clear what the contents of the slides represent.

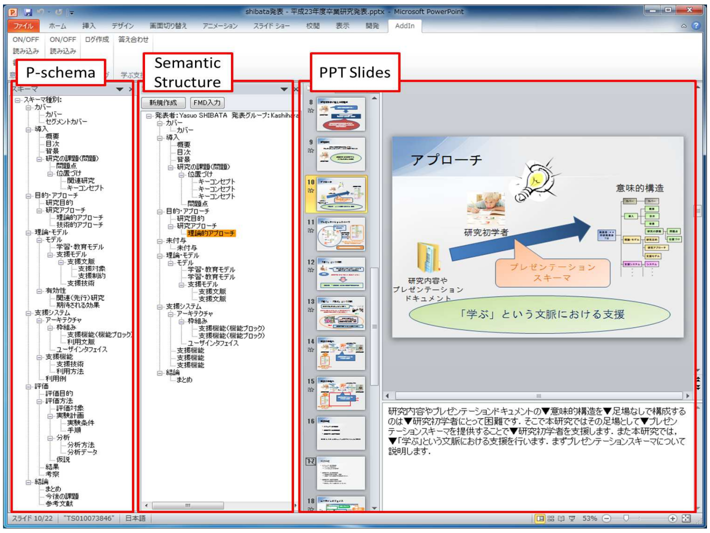


Figure 2: User interface for composing semantic structure of P-document.

2. Scaffolding System

Following the results of the case study, we have implemented a scaffolding system as an add-in of PowerPoint 2010 whose main aim is to help novice researchers compose semantic structures and P-documents. As shown in Figure 2, this system provides the novice with PPT functions and P-schema. The novice researcher can use P-schema in the left pane to compose a semantic structure of the P-document in the middle pane.

3. Conclusion

In this paper, we have introduced schema-based scaffolding where novice researchers could compose semantic structures of P-documents with P-schema that are accumulated in a research group. This paper has also reported a case study, and has demonstrated a scaffolding system with PPT functions. In future, we will evaluate the system in detail to refine the functionality including adaptive and intelligent components.

Acknowledgements

The work is supported in part by Grant-in-Aid for Scientific Research (B) (No. 23300297) from the Ministry of Education, Science, and Culture of Japan.

References

[1] A. Kashihara, and S. Hasegawa: A Scaffolding Framework for Creating Presentation Documents in Research Group. Proc. of SITE 2012, pp. 25-28 (2012).
[2] S. Hasegawa, A. Tanida, and A. Kashihara: Recommendation and Diagnosis Services for Presentation Semantics, Proc. of ICCE2010, pp.285-289 (2010).
[3] K. Saito, A. Tanida, A. Kashihara, and S. Hasegawa: An Interactive Learning Environment for Developing Presentation Skill with Presentation Schema, Proc. of E-Learn2010, Orlando, Florida, pp. 2696-2703 (2010).

Has the time come to discuss technical affordances of standards?

Tore HOEL
Oslo and Akershus University College of Applied Sciences, Norway
tore.hoel@hioa.no

Abstract: Formal standardization communities are searching for ways to improve their approach. Process improvement is often the primary focus for this discussion. This paper explores the possibilities to put more emphasis on the technical affordances of standards. A case study is presented of a proposal to reform European standardization practice. Technical quality issues are now put forward as important metrics. However, the technical quality criteria are still vague and more research is needed to come up with the dimensions for a quality discourse on technical aspects of standardisation.

Keywords: standardization, process improvement, technical standards quality

1. Introduction

Doubt is raised about the positive contribution of standardization to the development of the emergent field of learning technologies [1]. In times of crisis and self-scrutiny, one gets a chance to ask the more fundamental questions of how standardization is handling issues of quality related to processes and outputs. The purpose of this paper is to create awareness of this discussion, based on a small case study of a proposal from the European Commission (EC). EC has launched an "European Union standardization proposal" to be implemented from 2013. The EC wants to speed up the time it takes to make standards; to expand the remit of standards to cover services, management systems, environmental and social issues; and to make sure that appropriate standards developed outside the Europe are being implemented and used [2]. Similar initiatives are seen also in other parts of the world, e.g., in Australia [3]. These initiatives seem to take the technical quality of the standard as given; it is the market relevance and uptake that are identified as challenges to be addressed. However, a good purpose and justification for a standard is no guarantee for the standard to be well scoped and designed [4]. Therefore, the question raised in this paper is whether we see any opening for discussing technical affordances of standards in the current initiatives. And if so, how should this discussion be structured?

2. Related work

Hoel and Mason [5] have argued that qualities of standards should be addressed both in relation to process and product. With 'process' is meant the activities setting up the work towards a standard, i.e., choosing the right standard setting body, organizing a transparent process, ensuring stakeholder engagement, etc. With 'product' is meant the outcome of the standardization process, the standard itself. An improved process is not possible without a better understanding of the relationships between the three parts that make up standardization: process, product, and domain. The last part influences and is influenced by both process and product, as e.g. the domain supports certain processes and is best served by certain standards.

The quality of the standardization product, especially within the ICT domain, has often been discussed from a top-down perspective focussing on principles like correctness, clarity, relevance, comparability, economic efficiency, and systematic design [6]. However, another bottom-up perspective is also possible, discussing if the standard is well-formed, understandable, of the 'right size', etc. [7]. A third approach would be to discuss quality in relation to adoption, market uptake and software quality [8].

3. Case study: Accepting the work of others – adapting European standardization

A two page annex of an EC proposal [2] draws up the "requirements for the recognition of the technical specifications in the field of ICT". Three classes of requirements are identified, which this paper terms Relevance, Process and Technical Quality.

1. **Relevance** requires that the specification is accepted in the market and does not hamper interoperability.
2. **Process** requirements deal with openness, consensus, and transparency, in addition to the mandate and aim of the organization that has developed the specification – all known operational directives of the international standards bodies.
3. **Technical Quality.** The six requirements in the EC proposal related to the Product, i.e. the technical specification in question. *Maintenance*, *availability*, and *intellectual property* rights relate back to the process and the operational qualities of the organization that publishes the specification. These first three requirements pertain to the standard as a document. The last three requirements to externally developed standards relate directly to the technical characteristics of the specification: 1) *relevance*; 2) *neutrality and stability*; and, 3) *quality*.

The *relevance* criterion has two parts: (i) the specifications should be effective and relevant; (ii) specifications need to respond to market needs and regulatory requirements [2]. The first part is partly redundant (defining relevance by being relevant). However, it is noteable that the two parts are not merged, leaving a space related to effectiveness that is separate from the market and regulatory relevance space. This could be interpreted as an invitation to identify and discuss characteristics of the specification in question that are related to how the designed artifacts, e.g., information model, vocabularies, etc., works in a technical implementation.

The *neutrality and stability* criterion has also a mix of market and technical concerns, addressing (i) market distortion, competition and innovation; (ii) preferring performance orientation rather than development based on "design or descriptive characteristics"; and (iii) standards "based on advanced scientific and technological developments". The part about performance orientation, seems to invite to a discussion about principles of Information Systems Design, preferring specifications that 'work' to specifications that are 'developed the right way'.

The last criterion on *quality* is split in two parts. It may be easy to judge whether "standardized interfaces are not hidden or controlled by anyone other than the organizations that adopted the technical specifications". The part on "quality and level of detail" leaves on the other hand more room for discussion. The quality and the level of detail should be "sufficient to permit the development of a variety of competing implementation of interoperable products and services" [2]. Here the level of detail points directly to the design characteristic of the specification. What is the right level of detail? Is it "just enough" or is it necessary to strive towards a level of "correctness", giving an extensive and fully covered representation of the domain in question?

4. Discussion

The backdrop for this study is participant observation over a decade from both European and international ITLET standard groups, and a study of Directives setting up procedures for formal standardization [4]. When setting up ITLET projects in CEN and ISO there is nothing in the procedures that encourages discussions on approach, methodologies and general technical aspects of standards. This may be explained by the Directives [11], which stress methods neutrality. However, it could be argued that more emphasis on questions like rationale and scope, technical approach, base standards, technological context, etc. could strengthen the technical quality of the output of standardization. Therefore, it is interesting that EC requirements raise discussion about the technical affordances of standards as part of an effort to define what a good standard is. Even if the main focus is on acceptance in the market place, e.g., through a due process, technical design qualities are not seen as out of scope or in breach with methods neutrality.

5. Conclusions

This study has described standardization as an interrelationship between the standardization process, the outcome of standardization, and the domain served by standardization. Standard governance has mostly been concerned with process aspects of standardization. Also this case study shows that the issues related to acceptance in the market space seem better defined than technical issues. This points to the need for more theoretical work in this area. However, when a dominant stakeholder as the European Commission opens up for questions related to technical affordances of standards it gives an impetus to explore new avenues for a standardization discourse. What lies in an "effective" standard and the optimal "level of detail" [2] are questions that should be put forward for further research.

6. References

[1] Hoel, T., Hollins, P. A., & Pawlowski, J. M. (2010). On the Status of Learning Technology Specifications and Standards, *International Journal of IT Standards and Standardization Research*, vol. 8, no. 2.

[2] European Commission (2011). Proposal for a Regulation of the European Parliament and of the Council on European Standardisation, Com-2011-315

[3] Standards Australia (2010). Developing Standards – Net Benefit. http://www.standards.org.au/DevelopingStandards/NetBenefit.aspx

[4] Hoel, T., and Mason, J. (2012). Deficiencies of scope statements in ITLET standardization, in G. Biswas et al. (Eds.) (2012). Proceedings of the 20th International Conference on Computers in Education. Singapore: Asia-Pacific Society for Computers in Education

[5] Hoel, T., and Mason, J. (2011). Expanding the Scope of Metadata and the Issue of Quality, in *T Hirashima et al (Eds) (2011) Proceedings of the 19th International Conference on Computers in Education Chiang Mai, Thailand: Asia-Pacific Society for Computers in Education*.

[6] Schuette, R. and Rotthowe, T. (1998). *Lecture Notes in Computer Science*, vol. 1507, no. 20. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 240–254.

[7] Mendling, J., Reijers, H. A., and Cardoso, J. (2007). *What makes process models understandable?* Springer-Verlag, 2007, pp. 48–63.

[8] Int. Standards Org. (ISO): Information technology - software product evaluation - quality characteristics and guide lines for their use. ISO/IEC IS 9126 (1991)[2

English Language Learning in a One-to-One Computing Environment – Impacts and Considerations

Lee Yong TAY[a*], Jin Sheng NG[b], Cher Ping LIM[c]
Shanthi Suraj NAIR[a] & Siew Khiaw LIM[a]
[a]*Beacon Primary School, Singapore*
[b]*Nanyang Polytechnic, Singapore*
[c]*Hong Kong Institute of Education, Hong Kong/China*
*tay_lee_yong@moe.edu.sg

Abstract: This case study research describes and evaluates the impacts of the learning of English language in an ICT-enriched learning environment in a primary school under the FutureSchools@Singapore programme. Students from the school are provided with two-to-one computing devices in Primary 1 and subsequently one-to-one computing learning environment from Primary 2 (i.e., Grade 2) onwards. Students procure their own computing devices from Primary 4. This case study describes how ICT has been used for the teaching and learning of English. The frequency of ICT use in the English classrooms is also presented. More importantly, the impacts in terms of the students' ICT skills, English test score and feedback from students involved are also presented and discussed. The students performed well in the ICT skill test. The students also performed relatively well in the English test as compared to the control school. In general, the students commented that they were satisfied with the use of their notebook computers. They found it interesting, useful and easy to use. However, students complained about the computer viruses, small computer screen, battery lifespan, weight of their notebook computers, slow internet connect speed and internet connect errors. In the self-reported higher-order thinking questionnaire, no significant difference was found for students' problem-solving skills. However, students from the control school reported a statistically higher score as compared to the experimental school for reflective thinking.

Keywords: One-to-one computing, impacts, evaluation

1. Introduction

1.1 Objectives

The main intent of this research study is to obtain a more in-depth understanding of the impact of the use of Information Communication Technology (ICT) in the teaching of English in a primary school under the FutureSchools@Singapore program. This paper also describes how ICT has been used for the teaching and learning of English.

1.2 Background

The school in this case study research is the first primary future school that was set up under the FutureSchools@Singapore program in 2008. The school's core mission is to seek innovative teaching approaches that leverage on technologies to better engage the new generation of young learners. The school has implemented a successful one-to-one program (i.e., one-to-one computer to student ratio) for all its students.

2. Literature review

The literature review focuses on both the positive and negative impacts of the use of ICT in educational settings.

2.1 Positive Impacts of ICT

Research studies indicate positive impacts of ICT use in educational settings. Positive level of learning engagement and motivation [1][10][11][21][22] (Binnur, 2009; Kim & Jung, 2010; Krish, 2008; Softa, 2011; Suhr, Hernandez, Grimes, Warschauer 2010), higher test scores[3][22] (Brescia, Kissinger & Lee, 2009, Suhr, Hernandez, Grimes, Warschauer 2010) and the learning of higher-order related type of thinking skills[13][24] (e.g., problem-solving and critical thinking skills) (Lim and Tay, 2003; Takahira, Ando & Sakamoto, 2002) have been reported.

2.1.1 Higher English Test Scores

Brescia, Kissinger and Lee[3] (2009) find that students who used a computer for one hour per day showed more positive school behaviours and higher reading test scores than those who did not. The research by Suhr, Hernandez, Grimes, Warschauer[22] (2010) shows that 4th grade students in one-to-one computer classrooms outperformed students in the traditional classrooms in English Language Arts tests. Dumanig, David and Jubilado[6] (2011) also found that computer-mediated reading, when used adequately, creates a significant impact on students' reading comprehension skills.

2.1.2 Facilitation of Higher-Order Type of Learning

According to Mouri, Sakamoto, Hatano, and Sakamoto[15] (2002)'s research, schools which actively work on Internet Education produce students who can read critically. Lim and Tay[13] (2003) propose that ICT could be used to facilitate the learning of higher-order thinking via a constructivist's paradigm. Kim and Jung[10] (2010) report that digital textbooks have a positive effect on students' metacognition, self-regulated learning, self-efficacy, problem-solving, motivation, and self-reflection.

2.2 Issues and Considerations with the use of ICT in Educational Settings

Some of the authors who have reported the positive impacts have also raised concerns[11][12][15][22] (Krish, 2008; Lai & Pratt, 2004; Mouri, Sakamoto, Hatano & Sakamoto, 2002; Suhr, Hernandez, Grimes & Warschauer, 2010) such as insignificant impacts on student's English test scores[27][26] (Tilfarlioğlu, 2011; Tse, Yuen, Loh, Lam, Ng, 2010) and ICT as a distractor besides adding to the complexity to students' learning[2][11][22] (Blumenfeld, Kempler & Krajcik, 2006; Krish, 2008; Lai and Pratt, 2004; Suhr, Hernandez, Grimes and Warschauer, 2010).

2.2.1 Insignificant Impacts on Students' Performance

Tilfarlioğlu[27] (2011) concludes that while Web2.0, a language teaching tool, provide an opportunity for English Language learning, few students see its educational purposes. Tse, Yuen, Loh, Lam, Ng[26] (2010) also find that blogs do not have much influence on students' academic performances in English reading-comprehension in Hong Kong. There was also little evidence to prove that regularly consulting the Internet was associated with high grades on English reading tests. Blumenfeld, Kempler and Krajcik[2] (2006) caution that the use of ICT, even based on learning theories, may not bring out the intended effects as cognitive engagement requires more than the use of ICT alone.

2.2.2 ICT Complicates the Teaching and Learning Process

2.2.2.1 Additional Efforts and Support Required

Krish[11] (2008) discusses that the complexity of the classrooms and teachers' roles have also been increased. Instructors are required to be well-versed not only in the language, but also in technological tools. More effort and time is needed to prepare for lessons. Zare-ee and Shekarey [29](2010) collate a list of problems such as lack of infrastructures[14][17][18] (Mehlinger & Powers, 2002; Pelgrum, 2001; Rossberg & Bitter, 1988), lack of adequate training for teachers and staff[8][20][23][28] (Jacobson & Weller, 1988; Schrum, 1999; Strudler & Wetzel, 1999; Willis, Thompson & Sadera, 1999), weak technical support[19] (Schrum, 1995), lack of time, software, knowledge of available information technology resources, and unavailability of computer resources. Design elements used in ICT can also be a distraction from actual learning targets.

2.2.2.2 Computer and the Internet as a Distractor

Suhr, Hernandez, Grimes and Warschauer[22], (2010) find that students use their laptops at home in a variety of homework- related and entertainment activities. Writing papers was the most common, followed by browsing the Internet, managing photos, and playing games. There were only two homework-related items with entertainment activities dominating student's time. Another issue is that in the first year of implementation, much time was spent teaching students basic computer skills and less on content learning. Lai and Pratt[12] (2004) also present comments from the teachers that there is misplaced focus on the quantity of students' research instead of the quality. With the wealth of information available on the web, many students also have poor researching skills and critical reading ability. Plagiarism is also a serious problem. The positive effects of ICT were mostly superficial and changed neither the philosophy nor the teaching pedagogy of teachers to a large extent. Takahira, Ando and Sakamoto[24] (2007) have shown that increased usage of the Internet has resulted in a decline in their capability to produce information.

2.3 The Positive Impacts and Concerns Raised – ICT in Education

The above literature review seems to suggest that ICT in educational settings could bring about higher level engagement and motivation among learners, improve English test scores and also facilitate the learning of higher-order type of thinking skills or be a device that distracts the students from their academic or learning tasks.

3. Research Design and Methods

A case study approach is used in this research study to look into the impacts and considerations of the ubiquitous use of ICT in the teaching and learning of English in a primary school context, focusing on the pioneer batch of Primary 4 students who has been with the school since its inception from January 2008.

The following data were being collected for analysis and triangulation – (1) interview with the five English teachers; (2) the students' frequency of in class ICT use survey; (3) feedback from the 167 Primary 4 students; (4) document review of scheme of work, (5) the students' English test scores; and (6) the students' self-reported questionnaire survey.

4. Key Findings

4.1 Interview with teachers

A total of 5 teachers taught the seven Primary 4 English classes. All the teachers were interviewed individually so that the responses from the teachers could be independent. These sessions took the form of informal interviews; based on how they had integrated ICT into their lessons. All the interview sessions were tape-recorded and transcribed for analysis.

4.1.2 ICT and the teaching of English

The English quiz module in the learning management system is the most frequently used tool. The self-marking task and item analysis available in the quiz module was time-saving and simple. Students were given comments immediately and were more driven to reflect and get the right answer. Teachers also set up blog sites for the sharing of online teaching materials and for students to upload their compositions so their classmates and teachers could exchange pointers to further improve the flow and idea of their writings.

On top of that, the English department also subscribed to an online reading comprehension portal where students could learn by accessing the online digital stories available at their own speed. The English teachers piqued their students' interest through the writing of digital stories, a main instructional methodology for the picking up of languages and media literacy techniques. The students first employed a suitable software programme to formulate a digital story. Teachers provided feedback and the finished stories were circulated on the school network and internet through blog sites, so everyone could look at it and offer their opinions or encouragement. The students were engrossed and enjoyed the process of turning their ideas into words that were accompanied by vibrant pictures and music. ICT expedited the presentation of their stories.

4.1.2 Teachers' Reflections and Thoughts about Students' Learning

Teachers who were interviewed unanimously reflected that the use of ICT in their English classes were engaging and motivating in learning and academic tasks. They also felt that it could facilitate the learning of the higher-order type of thinking skills. However, they also raised concerns, just like the literature reviewed, that ICT seemed to engage and motivate their students but the teachers were not certain that this would lead to better performance in students' English test scores. Teachers also shared that the use of ICT required more effort to prepare lessons, technical support and resources (i.e., technological infrastructures).

4.3 Feedback from students

In general, the students found the notebook computers interesting, useful and easy to use. However, the students also complained about the computer viruses, battery lifespan, and computer weight, slow internet connect speed and internet connect errors.

The Primary 4 students were asked to respond on how they felt about their use of their personal notebook computers. 167 out of the 225 students responded in a positive and encouraging manner. The positive comments were 3.8 times more than the negative, with mainly technical related complains.

4.5 Students' English Test Scores

All the Primary 4 students (aged 10) took a 38-item (48 marks) English test with a comparative control school. The test scores were converted to 100%. The paper was in the form of multiple-choice questions, fill-in-the-blanks and open-ended responses. The students were tested on their vocabulary, grammar, cloze, editing and comprehension. A total of 440 Primary 4 students took the English test with 216 from the school and 224 from the control. The mean scores were 73.87 and 67.38 between the school and comparative control school, respectively. The mean difference between the schools was significant at t (434) = 4.414, $p < 0.001$. The effect size between the two means according to Cohen's d was 0.4057 (small-medium effect size). A non-linear regression analysis on the frequency of ICT usage during English lessons and the students' English results during their first semester assessment shows that the frequency of ICT usage during English lessons had an impact on the students' English test scores.

4.6 Students' self-reported questionnaire survey on critical problem-solving and reflective thinking (higher-order type of thinking skills)

A total of 890 students (Primary 4 students) participated in a self-reported questionnaire survey (i.e. 434 from the school and 456 students from the control school) on their critical problem solving and reflective thinking. The self-reported questionnaire survey items were presented using a 7-point Likert response scale with 1 = strongly disagree and 7 = strongly agree. On average, the questionnaire did not take more than 5 – 10 minutes to complete. The questions were derived through literature search, discussion among the authors and validation with experienced teachers.

4.6.1 Principal components analysis

An exploratory factor analysis was conducted using the principal components analysis. In particular, the varimax rotation was employed onto the 15 items. In accordance to recommendations from Kaiser[9] (1960), components with eigenvalues greater than 1.0 were adopted. Two components with eigenvalues more than 1.0 were yielded and it composed of 50.65% of the total variable explained. Most of the factor loadings were above 0.5 which bears practical significance and was therefore retained for further analysis[7] (Hair, black, Babin and Anderson, 2010). Furthermore, similar test with oblique rotation (Promax: Kappa = 4) showed consistent results of two constructs. To validate the items within each of the two constructs, a reliability analysis, in the form of Cronbach's[5] alpha was used. In accordance to recommendations by Cronbach[5], scores of the two constructs were 0.771 and 0.898 and these were considered acceptable and good. With the above analysis, we therefore derived two constructs, namely, problem-solving and reflective thinking. Based on initial literature review, these are in turn related to higher-order thinking skills.

4.6.2 Confirmatory factor analysis

Confirmatory analysis was used to examine the factorial structure of the 15 item scale and AMOS 17.0 was used for this analysis. A proposed model was built and model fit was assessed by a number of indices. The Mardia's coefficient (1970), which is a standard measure of multivariate normality, was first consulted. The coefficient, 109.087, was lower than the recommended threshold of 225 (15X17) and based on the recommended value of p*(p+2), the requirement is satisfied hence the data for this study was deemed to be adequate for confirmatory factor analysis.

Next, we turn our attention to the model fit indices. As chi-square has been found to be too sensitive to sample size, the ratio of chi-square to its degree of freedom was used (CMIN/DF) and scores of 3.0 or below (CMIN/DF = 1.925) are deemed to be an acceptable fit between the hypothesized model and sample data [4](Carmines and Mclver, 1981).

4.6.3 Model Comparison

In accordance to suggestions by Noar[16] (2003) and Timothy et al. [25] (2010), fit indices of alternative models allow for comparison of proposed instruments. Based on the R^2 values of the items, the item with the lowest R^2 (item 1, $R^2 = 0.210$) was removed from the proposed model and a 14-variable model was re-run. The Tuker-Lewis Index and the Comparative Fit Index both fell below 0.95 (TLI = 0.916 < 0.95, CFI = 0.928 < 0.95), indicating a less than acceptable fit. On this basis, we concluded that model 1 with 15-items and two constructs was retained as the model of better fit

4.6.4 Means Testing (T-Test) on each of the two constructs

We further want to investigate if there were any difference between the 2 schools in terms of the creative problem solving and critical reflective thinking constructs. We derived mean scores for each of the constructs and compared the scores between the schools with results as follows, using the *t*-test procedure

4.6.5 Problem-solving and reflective thinking

Based on the Levene's test of equality of variances, we proceed to look at the results to the variance assumed *t*-test, with results $t = 1.089$, $p = 0.277 > 0.05$. We therefore conclude that there we no statistical difference between the 2 schools in terms of the mean problem-solving scores. For reflective thinking, based on the Levene's test of equality of variances, we proceed to look at the results to the variance not assumed *t*-test, with results $t = -3.923$, $p = 0.000 < 0.05$. We therefore conclude that there we is a statistical difference between the 2 schools in terms of the mean reflective thinking scores with mean scores of the control school higher at 6.41 against the school in this study at 5.9474.

5. Discussion and conclusion

The above findings seem to suggest that students' level of engagement and motivation was high when ICT was used in their English lessons. In addition, the students' English test scores were, on average, significantly higher than the control school. The students also performed well in the ICT skills tests as compared to the national cohort. On an average, students' frequency of ICT usage during English lessons was recorded as 5.16. Computer use was frequent during English lesson and in general, students enjoy using ICT for learning.

The interview with the teachers and review of documents reflect sentiments similar to the literature reviewed, that more effort and extensive technological infrastructures were required. At times, ICT was a device of distraction. No difference was observed on students' creativity in problem-solving between the experimental and control schools. However, students' from the control school reported a significantly higher mean score for critical reflective thinking as compared to the experimental school. Although the students' English test score were higher than those of the control, the regression analysis indicated that the r^2, which is the variance explained, only accounted for 3.1 percent. In other words, the use of ICT only accounted for 3.1 percent of students' English performance.

All in all, it seems that ICT does have an impact on the learning of English in the primary or elementary level, especially in areas of engagement, motivation and test scores. Students generally favour the use of ICT for their learning of English. However, it is important to consider the extra effort, manpower and technological infrastructure overheads when embarking on such an endeavour. It is also worth the effort to further explore how to reduce students' off task activities with their computers. The relationship between ICT use and the facilitation of the acquisition of higher-order type of thinking skills also deserve more in-depth study and exploration; in terms of how could this aspect be more accurately measured.

References

[1] Binnur, G.I. (2009). Effect of technology on motivation in EFL classrooms. *Turkish Online Journal of Distance Education, 10*(4), 136-158.

[2] Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2006). Motivation and cognitive engagement in learning environments. In R. K. Sawyer, (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 475-488). New York, NY: Cambridge University Press.

[3] Brescia, W., Kissinger, D. & Lee, S.K. (2009). Computer use and academic development in secondary schools. *Computers in the Schools, 26*(3), 224-235. DOI: 10.1080/07380560903095204

[4] Carmines, E.G., &Mclver, J.P. (1981). Analyzing models with unobserved variables: analysis of covariance structures. In G. W. Bohrnstedt, & E. F. Borgatta (Eds.), *Social measurement: Current issues* (pp. 65-115). Beverly Hills, CA: Sage.

[5] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

[6] Dumanig, F. P., David, M. K., Jubilado, R. (2011, November). *Computer-mediated reading and its impact on learners' reading comprehension skills*. Paper presented at the Brighton Business School, University of Brighton, UK.

[7] Hair, J.E., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Sandle River, New Jersey: Prentice-Hall.

[8] Jacobson, M. & Weller, M. (1988). A profile of computer use among the University of Illinois humanities faculty. *Journal of Educational Technology Systems, 16*(2), 83-97.

[9] Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141 – 151.

[10] Kim, J. H. Y. & Jung, H. Y. (2010). South Korean digital textbook project. *Computers in the Schools, 27*(3/4), 247-265. DOI: 10.1080/07380569.2010.523887

[11] Krish, P. (2008). Language learning in the virtual world: Instructors' voices. *International Journal of Pedagogies and Learning, 4*(4), 113-129.

[12] Lai, K. W., Pratt, K. (2007). Positive to a degree the effects of ICT use in New Zealand secondary schools. *Computers in the Schools, 24*(3/4), 95-109. DOI: 10.1300/J025v24n03_07

[13] Lim, C.P. & Tay, L.Y. (2003). Information and communication technologies (ICT) in an elementary school: Engagement in higher order thinking. *Journal of Educational Multimedia and Hypermedia, 12*(4), 425-451.

[14] Mehlinger, H. D., & Powers, S. M. (2002). *Technology & teacher education: A guide for educators and policy makers*. Boston: Houghton Mifflin Company.

[15] Mouri, M., Sakamoto, A., Hatano, K., & Sakamoto, T. (2002). The effects of computer use in high school on students skills of practical use of information. *Japan Journal of Educational Technology, 26*(Suppl.), 85-90.

[16] Noar, S.M. (2003). The role of structural equation modeling in scale development. *Structural Equation Modeling, 10*(4), 622-647.
[17] Pelgrum, W. J. (2001). Obstacles to the integration of ICT in education: results from a worldwide educational assessment. *Computers & Education, 37*(2), 163–178. DOI: 10.1016/S0360-1315(01)00045-8
[18] Rossberg, S. A., Bitter, G. G. (1988). Microcomputer infusion project: A model. *Techtrends, 33*(5) 24-28. DOI: 10.1007/BF02786308
[19] Schrum, L. (1995). Educators and the internet: a case study of professional development. *Computers and Education. 23*(3), 221-228.DOI: 10.1016/0360-1315(95)00012-B
[20] Schrum, L. (1999). Technology professional development for teachers. *Educational Technology Research and Development, 47*(4), 83-90. DOI: 10.1007/BF02299599
[21] Softa, B. (2011). Learning environment effect and use of technology in the study of English language. *Problems of Education in the 21st Century, 35,* 127-137
[22] Suhr, K.A., Hernandez, D.A., Grimes, D., Warschauer, M. (2010). Laptops and fourth-grade literacy: Assisting the jump over the fourth-grade slump. *The Journal of Technology, Learning and Assessment, 9*(5), 6-45.
[23] Strudler, N., & Wetzel, K. (1999). Lessons from exemplary colleges of education: factors affecting technology integration in preservice programs. *Educational Technology Research and Development, 47*(4), 63 - 81. DOI: 10.1007/BF02299598
[24] Takahira, M., Ando, R., Sakamoto, A. (2007). Effect of Internet use on development of information literacy: A panel study with Japanese elementary school children. *Computers in the Schools*, *24*(3/4), 61-82. DOI: 10.1300/J025v24n03_05
[25] 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 & Education, 55, 1764-1771.
[26] Tse, S. K., Yuen, A. H. K., Loh, E. K. L., 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.
[27] Tilfarlioğlu, F.Y. (2011). An international dimension of the student's attitudes towards the use of English in Web 2.0 technology. *The Turkish Online Journal of Educational Technology*, *10*(3), 63-68.
[28] Willis, J., Thompson, A., & Sadera, W. (1999). Research on technology and teacher education: current status and future direction. *Educational Technology Research and Development, 47*(4), 29-45. DOI: 10.1007/BF02299596
[29] Zare-ee, A. & Shekarey, A. (2010). Comparative study of the use of ICT in English teaching-learning processes. *Turkish Online Journal of Distance Education, 11*(2), 13-22.

An Integrated GPS-supported Outdoor Exploratory Educational System—EagleEye

Morris S. Y. Jong[a*], Eric T. H. Luk[b], Jimmy H. M. Lee[c]
[a]*Department of Curriculum and Instruction*
[abc]*Centre for the Advancement of Information Technology in Education*
The Chinese University of Hong Kong, Hong Kong
[*]mjong@cuhk.edu.hk

Abstract: EagleEye is an integrated GPS-supported educational system for supporting students and teachers respectively in pursuing and facilitating exploratory learning in outdoor fieldtrip activities. This system has four components, including the (1) Location-based Exploratory Resource Authoring Tool, (2) GPS-supported Exploratory Platform, (3) Repository Server, and (4) Teacher Console. A preliminary study, which involved 40 participants (38 students and 2 teachers from a school) adopting EagleEye in an outdoor fieldtrip activity, was carried out to investigate their perceptions of this system. It was found that EagleEye brought desirable fieldtrip experience to the students. The teachers also perceived positively the educational potential of EagleEye for outdoor fieldtrip activities from both technical and pedagogical perspectives.

Keywords: EagleEye, exploratory learning, GPS-supported outdoor learning, mobile learning

1. Introduction

With the advancement of mobile technology in the recent decade, there have been educational researchers endeavouring to study the educational potential of mobile learning and how mobile devices can be integrated into the process of education. For example, Sharples et al. [1] developed a framework theorizing about mobile learning. This framework informs researchers of (1) the design of mobile technology for educational purposes, and (2) the analysis of learning taking place in mobile contexts. Shih et al. [2] developed a mobile learning platform for scaffolding students in the course of inquiry-based learning in the context of social science education. Looi et al. [3] transformed an existing science curriculum into a mobilized inquiry-based science curriculum for delivery via 1:1 mobile devices. Boticki et al. [4] investigated the integration of mobile learning and CSCL (computer-supported collaborative learning), namely, mCSCL, for promoting students' collaborative experience and sharpening their skills of communication, negotiation and decision-making. In this paper, we discuss our new mobile learning initiative on harnessing the Global Positioning System (GPS) technology to support outdoor fieldtrips.

The GPS is a space-based satellite navigation system, providing reliable geo-location information anywhere on the earth [5]. Based on the GPS technology, we have developed an integrated mobile exploratory educational system, namely, *EagleEye*, for empowering students and teachers respectively in pursuing and facilitating exploratory learning in outdoor fieldtrip activities. Early this year, we conducted a preliminary study to probe into students' and teachers' perceptions of using EagleEye according to their experience in an outdoor fieldtrip with the adoption of this system. This study involved 38 students and 2 teachers from a school. We used a mixed research approach with a combination of student questionnaire-based survey, teacher interview, and researcher observation.

2. Rationale behind Our Initiative

Knowledge cannot be separated from how and why it develops [6]. Nevertheless, the knowledge taught at school are usually fragmented into small and unconnected pieces. The original intention is for making learning easier, but this often ends up neglecting the rationale behind the knowledge itself, creating unrealistic learning content and context, and rendering the whole learning process boring [7]. In the view of *constructivist education*, learning should be an active process in which students construct knowledge by interacting with rich and authentic learning environments ([8], [9], [10]). *Experiential learning* [11] is one of constructivist learning paradigms being advocated in today's education, while one of the pedagogical approaches to implement this learning paradigm is *fieldtrips* [12].

Fieldtrips place learning in contexts that can never be replicated in classrooms [13]. Evidence has shown that, in comparison with classroom activities, students in fieldtrip activities are more motivated [14], more willing to transfer, apply, and anchor knowledge [15], as well as more likely to generate greater subject matter interest [16].

Nadelson and Richard Jordan [17] categorize fieldtrips into two common genres. The first genre is indoor-based, for example, a class visiting to a museum, science centre, institution, etc. The second genre is outdoor-based, for example, a class visiting to a park, forest, wetland, villages, or other sites with specific natural or heritage settings. The latter is usually adopted in geography, ecology, or cultural education for providing students with opportunities to pursue exploratory learning in real-life, real-world environments. In this paper, our focus is on outdoor fieldtrips. For writing convenience, the term "fieldtrip(s)" refers to "outdoor fieldtrip(s)" hereafter.

Conventionally, students' fieldtrips are led by teachers [12]. Before a fieldtrip activity, a teacher will design a set of paper-based worksheets for scaffolding his/her students in groups to observe, experience, and reflect, in accordance with some specific learning objectives of the fieldtrip. Usually, these worksheets contain open-ended questions for facilitating the students during the fieldtrip to think about and response to. At the end of the fieldtrip, the students have to submit the teacher their "answers" to these questions in written format. These submissions will sometimes be used for assessing the students' performance in the fieldtrip.

Students and teachers, however, do encounter various difficulties in pursuing and facilitating conventional fieldtrips activities. Before our development of EagleEye, we conducted in-depth interviews with 5 senior secondary students and 5 secondary teachers respectively from three different schools. The students and teachers respectively had rich fieldtrip-participation experience and rich fieldtrip-facilitation experience. According to their experiences and observations in the past, they elaborated on a number of problems arising frequently in conventional fieldtrip activities, as summarized below.

Problem 1: Learning taking place in a teacher-centred manner. The student-to-teacher ratio in a fieldtrip is usually large. Typically, one to two teachers facilitate a class of 40 students (or even more). Lacking enough "manpower," the teachers will bring the whole class to the designated exploratory spots on the fieldtrip site one by one in a designated order, without allowing the student groups to plan and frame their own exploratory route. This violates largely the original advocacy of the fieldtrip approach to facilitate learners to learn in a constructivist, student-centred fashion.

Problem 2: Students' learning motivation cannot be sustained. At the beginning of a fieldtrip, students' learning motivation is usually high. They are happy and excited (see

Figure 1a), because outdoor activities are rare relatively in comparison with everyday classroom activities. However, their motivation will often decrease gradually in the course of the fieldtrip. A reason is that the students have no right to control the time being spent on a particular exploratory spot in correspondence with their own learning pace and interest. The time for staying at each spot is determined by their teachers, not themselves. Further, every time when a class of 40 (or even more students) arrives at a single exploratory spot (usually a small area), it is too crowned for every student to observe and experience the context around and the happenings (e.g., social and cultural events) therein (see Figure 1b).

Problem 3: More effective scaffolds are needed. In a fieldtrip, most of the learning scaffolds (guiding/open-ended questions) are given to students through worksheets. These worksheets are mainly text-based (or sometimes with images, such as, maps, pictures, etc.) presented in a static manner. During the fieldtrip, the students can only use texts to respond to the questions by writing on the spaces aside or below the questions, or on the opposite slides of the worksheets. This kind of paper-and-pencil scaffolds, nonetheless, does not appeal to today's students who are, in Prensky's [18] terms, "digital natives." They are eager to have technological and multimedia elements for supporting their learning process.

Problem 4: Collaboration among students is weak. In a fieldtrip, although students are usually divided into groups and asked to pursue the exploration collaboratively, most of them just care about whether they can complete the worksheets in hand before the end of the fieldtrip. Instead of having much discussion with their groupmates, the students spend a lot of time on copying the information from the fieldtrip site (see Figure 1b). They hope that the information can help them to answer the questions on the worksheets.

The aim of our development of EagleEye is to mitigate the above problems. Besides the employment of mobile technology, from the pedagogical perspective, we adopt Jonassen et al.'s [19] framework, "meaningful learning with technology," as the foundation for shaping the design and implementation of EagleEye.

(a) Beginning of a fieldtrip: Happy and exciting faces

(b) During a fieldtrip

Figure 1. Students in a Conventional Fieldtrip Activity

3. EagleEye

EagleEye consists of four core components. The following will elaborate on the specific function of each component.

3.1. Location-based Exploratory Resource Authoring Tool (LERAT)

The LERAT is a PC-based software tool for teachers to create location-based exploratory resources for running on GPS-enabled tablets. Each of these resources is a combination of a *map* and a number of location-based *exploratory scaffolds* for facilitating students to pursue exploratory learning during a fieldtrip. The interface of the LERAT is shown in Figure 2. When a teacher creates a location-based exploratory resource, first of all, he/she needs to import a map (in image format)[1] into the LERAT for specifying the geo-area corresponding to his/her planned fieldtrip. This imported map becomes the base (the background) of the resource. Further, he/she needs to conduct simple calibration by inputting the corresponding real-world latitude and longitude values[2] into the corners of the map. After that, the teacher can set up, at each designated exploratory spot on the map, a location-based exploratory scaffold. Usually, each scaffold contains a guiding question for hinting the students to explore that spot. These scaffolds will then be presented as "hotspots" (see the circles in Figure 2). Various templates (multiple-choice, true/false, fill-in-the-blank, open-ended question types, etc.) are available in the LERAT for assisting the teacher in designing the scaffolds. After developing the resource, he/she will upload it the RS (see Sub-section 3.3).

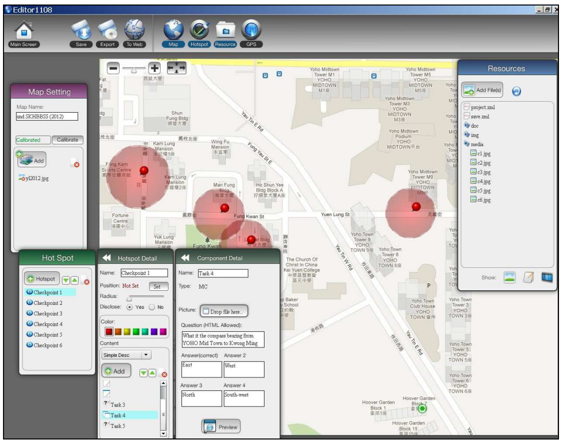


Figure 2. Location-based Exploratory Resource Authoring Tool (LERAT)

Figure 3. GPS-supported Exploratory Platform (GEP)

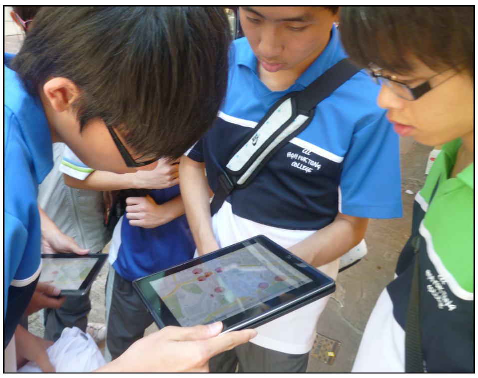


Figure 4. "Hotspots" Popping-up on Students' GEP

3.2. GPS-supported Exploratory Platform (GEP)

The GEP is a tablet-based software application (i.e., an App). The current version of the GEP is designed for running on Apple™ iPads, and is available for free download from the Apple™ App Store. The interface of the GEP is shown in Figure 3. Before a fieldtrip,

[1] This image source can be obtained simply by screen-capturing it from free online maps, such as, Google™ Maps (http://maps.google.com) or Bing™ Maps (http://maps.bing.com).
[2] These values can also be obtained easily from free online maps, such as Google Maps™ or Bing Maps™.

students will first connect to the RS (see Sub-section 3.3) to download the corresponding location-based exploratory resource (designed by their teacher) to their GPS-enabled tablets. During the fieldtrip, the students will open the resource with the GEP. An "avatar" (see Figure 3) will appear on the map to indicate their current geo-location in the real world. Based on the ongoing GPS signals received by the GEP, the "hotspots" (exploratory scaffolds) embedded in the resource will pop up automatically when the students step in physically the corresponding geo-locations on the fieldtrip site (see Figure 4). These hotspots will guide them to observe, experience, and reflect during the course of the fieldtrip. They can also submit their responses to the questions (presented at the hotspots) through the GEP to the RS (see Sub-section 3.3). Hence, their teacher can retrieve those responses from the RS through the TC (see Sub-section 3.3). Unlike conventional fieldtrip activities where a whole class crowds into each exploratory spot simultaneously, with the GEP students can work in small groups to plan and frame their own exploratory route and how much time to be spent at each spot, according to their own learning pace and interest.

3.3. Repository Server (RS) and Teacher Console (TC)

The RS has three main functions. Firstly, it is for teachers to upload their created location-based exploratory resources. Secondly, it is for students to download the resources to their GPS-enabled tablets before the fieldtrips. Thirdly, it is for storing the students' on-going responses to the questions (exploratory scaffolds) during the fieldtrips. The TC is a web-based (browser-based) platform connected to the RS. It aims at enabling the teachers to retrieve their students' responses to the questions after the fieldtrips. Those data will provide the teachers with useful information for assessing and debriefing their students.

4. Research Design

The aim of the present study was to probe preliminarily into students' and teachers' perceptions of EagleEye after experiencing its use during a fieldtrip activity. A secondary school in Hong Kong was invited to participate in the study. The participants involved a class of 38 Grade-9 students (aged 16.3 in average) and 2 geography teachers. The students had some experience in participating conventional fieldtrips when they were in lower grades. The teachers had rich experience in organizing conventional fieldtrips.

The fieldtrip in this study was conducted as an extra-curriculum event of geography education. The fieldtrip site was a "semi-rural" area located at the New Territories in Hong Kong— a village called Lung Yeuk Tau. From the educational perspective, this fieldtrip aimed at letting the students explore the impacts of urbanization on the conservation of the traditional Chinese culture in Hong Kong. In fact, the teachers had organized similar fieldtrip activities (on the same site, with the same educational aim) but in a conventional fashion (with paper-based worksheets) for their former students in previous years. The following sub-sections delineate further the design of this study.

4.1. Teacher Enablement Training

Four weeks before the fieldtrip, we conducted a two-hour training session for the teachers. The training provided them with hand-on practice on the use of the LERAT, GEP, and TC of EagleEye. After that, we requested the teachers to convert their paper-based worksheets used previously in the Lung Yeuk Tau fieldtrip into a location-based exploratory resource with the LERAT for the present study. In addition, we asked them to practice the use of the GEP and TC by themselves. Within a week, the teachers created the resource.

4.2. Setting of the Fieldtrip Activity

The 38 students were divided randomly into 8 groups (4 to 5 students per group). Each group was given a tablet (Apple™ iPad). The fieldtrip activity was held in a morning, taking around 3.5 hours to finish (excluded the round-trip travelling time between the school and the fieldtrip site). The first part (20 minutes) was a briefing session (taking place in the school) in which the teachers briefed the students on the aim of the activity, the operation of the GEP, as well as some safety issues. The second part (2.5 hours) was the fieldtrip in which the students worked in groups with the GEP. The third part (40 minutes) was a debriefing session (taking place after coming back to the school) in which the teachers retrieved (with the TC) and discussed the students' responses to the questions (exploratory scaffolds) embedded in the location-based exploratory resource.

4.3. Data Collection

We observed the entire study (including the teacher enablement training, briefing session, fieldtrip, and debriefing session). A post-activity student survey was conducted right after the completion of the debriefing session. The questionnaire of the survey consisted of 10 items, in 5-point Likert scale. Table 1 describes the design of the questionnaire, while Table 2 (in Sub-section 5.1) displays the 10 items. The first eight items requested the students to rate their present EagleEye-supported fieldtrip experience in comparison with their past experiences in conventional fieldtrips. These items were designed with respect to the four problems in conventional fieldtrips that we have discussed earlier in Section 2. We wanted to study the possibility of adopting EagleEye to mitigate these problems. The last two items requested the students to rate their overall perceptions of the use of EagleEye. After finishing the survey, we interviewed the two teachers together for gathering their perceptions of the adoption of EagleEye in fieldtrip activities.

Table 1. The Design of the Questionnaire Items

Aim	Items	Corresponding Problem (see Section 2)
To study the possibility of adopting EagleEye to mitigate the problems in conventional fieldtrips	Items 1, 2	*i) Learning taking place in a teacher-centred manner.*
	Items 3, 4	*ii) Students' learning motivation cannot be sustained.*
	Items 5, 6	*iii) More effective scaffolds are needed.*
	Items 7, 8	*iv) Collaboration among students is weak.*
To study the students' overall perceptions of the use of EagleEye	Items 9, 10	/

5. Findings

We discuss the findings from two perspectives. First is the students' perceptions based on the post-activity student survey results and our observation. The latter is the teachers' perceptions in accordance with the teacher interview and also our observation.

5.1. Students' Perceptions of the EagleEye-supported Fieldtrip

The questionnaire return rate was 100%. Table 2 shows the descriptive statistics *(M: Means, SD: Standard Derivation)* of the students' responses to each items.

The results align with what we observed in the study. From the students' point of view, EagleEye did provide them with a favourable fieldtrip experience *(see Items 9 &10)*. In addition, they rated their experience in the present fieldtrip more desirable than their past experiences in conventional fieldtrips. This indicates EagleEye did mitigate, to a certain extent, the problems discussed in Section 2. Firstly, EagleEye made the fieldtrip more active and student-centred (see Items 1 & 2, mitigating Problem 1). Secondly, EagleEye motivated and engaged the students in a greater extent (see Items 3 & 4, mitigating Problem 2). The students were provided with better support through EagleEye (see Items 5 & 6, mitigating Problem 3). Lastly, EagleEye promoted the collaboration among the students (see Items 7 & 8, mitigating Problem 4).

Table 2. The Post-activity Student Survey Results

Item		M	SD
Comparing to my past fieldtrip experiences, ***in the fieldtrip today …*** **5: Strongly Agree 4: Agree 3: Neutral 2: Disagree 1: Strongly Disagree**			
1	I participated more actively because of EagleEye.	4.26	0.49
2	I had better control on the process of exploration according to my learning pace and interest because of EagleEye.	4.43	0.50
3	I am more motivated in the process of exploration because of EagleEye.	4.31	0.49
4	I was more engaged in the process of exploration because of EagleEye.	4.23	0.36
5	The location-based features of EagleEye *(the indication of my geo-locations, and the just-in-time pop-up of hotspots)* provided me with better support in the process of exploration.	4.21	0.50
6	The multimedia and interactive features of EagleEye *(the hotspots, and the questions presented therein)* provided me with better support in the process of exploration.	4.25	0.48
7	I had more discussion with my groupmates because of EagleEye.	4.37	0.67
8	I have better collaboration with my groupmates because of EagleEye	4.26	0.63
In an overall manner, ….			
9	The use of EagleEye offers me good fieldtrip experience.	4.48	0.45
10	I am eager to have more chances to have EagleEye-supported fieldtrips in the future.	4.47	0.46

5.2. Teacher' Perceptions of the EagleEye-supported Fieldtrip

Owing to the limited spaces of this paper, we discuss only the teachers' perceptions of EagleEye from the technical and pedagogical perspectives. From the technical point of view, the teachers found EagleEye was easy to use, and appreciated the user-friendliness of the LERAT, GEP, and TC. They elaborated—

> *We are not smart in IT and computing in fact. However, we didn't have any difficulties in converting the paper-based worksheets into the location-based exploratory resource with the LERAT for the present fieldtrip. We also didn't have any difficulties in getting familiarized ourselves with the use of the GEP and TC. The overall operation of EagleEye is quite user-friendly.*

The teachers' elaboration aligns with what we observed in the study. During the teacher enablement training, the teacher did well in the hand-on practice on EagleEye. During the briefing session, they were able to explain the operation of the GEP to their students fluently. During the debriefing session, they could retrieve their students' responses to the questions (exploratory scaffolds) through the TC without any problems. From the pedagogical point of view, the teachers perceived the educational potential of EagleEye positively. Specifically, they appreciated its ability to engage and motivate the students during the fieldtrip and make the whole fieldtrip more student-centred. They elaborated—

> *We have organized fieldtrips over many years, but we have never seen students so motivated and engaged like the ones in the fieldtrip today. We are also happy to see today the students could plan and frame their own exploratory route, and have more control on their own exploration ……The performance of the students today is impressive.*

6. Conclusion and Further Study

In this paper, we have delineated our mobile learning initiative, EagleEye, an integrated GPS-supporting students and teachers respectively in pursuing and facilitating exploratory learning in outdoor fieldtrip activities. We have also discussed our perceptional study for investigating students' and teachers' experience in using EagleEye, in which we got some initial success and encouragement. However, as highlighted in the paper, the study was a preliminary one. More research work has to be done on further studying the educational effectiveness of EagleEye, in terms of, for example, advancing students' knowledge and exploratory skills, promoting students' subject matter interest and affection for the natural environments, etc. In addition, in order to enhance the existing use of EagleEye, we would like to know if we need a new pedagogical approach which is different from the typical 3-phase approach (briefing, fieldtrip, and debriefing) adopted for long in conventional outdoor fieldtrip activities. All of the above have been added to our research agenda.

References

[1]. Sharples, M., Taylor, J., & Vavoula, G. (2005). *Towards a theory of mobile learning*. Retrieved May 20, 2012, from http://www.mlearn.org.za/CD/papers/Sharples-%20Theory%20of%20Mobile.pdf

[2]. 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.

[3]. Looi, C. K., Zhang, B., Chen, W., Seow, P, Chia, G., Norris, C., & Soloway, E. (2011). 1:1 Mobile inquiry learning experience for primary science students: A study of learning effectiveness. *Journal of Computer Assisted Learning, 27*, 269–287.

[4]. Boticki, I., Looi, C. K., & Wong, L. H. (2011). Supporting mobile collaborative activities through scaffolded flexible grouping. *Educational Technology & Society, 14*(3), 190–202.

[5]. Ashbrook, D., & Starner, T. (2002). Learning significant locations and predicting user movement with GPS. *Proceedings of the 6th International Symposium on Wearable Computers* (pp. 101-108). Retrieved May 20, 2012, from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=01167224

[6]. Dewey, J. (1938). *Experience and education*. New York: Macmullan.

[7]. Papert, S. (1993). *The children's machine: Rethinking school in the age of the compu*ters. New York: Basis Books.

[8]. Hein, G. (1992). Constructivist learning theory. *Developing museum exhibitions for lifelong learning* (pp. 30-34). London: The Stationery Office.

[9]. Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

[10]. Otting, H., & Zwaal, W. (2007). The identification of constructivist pedagogy in different learning environments. In M. K. McCuddy, H. van-den-Bosch, W. B., Martz, A. V. Alexei, & K. O. Morseb (Eds.), *The challenges of educating people to lead in a challenging world* (pp. 171-196). Netherlands: Springer.

[11]. Kolb. D. (1984). *Experiential learning*. New Jersey: Prentice Hall.

[12]. Douglass, D. (2008). *The out-of-classroom experience*. Retrieved May 20, 2012, from http://www.4faculty.org/includes/digdeeper/Outside/outside.htm

[13]. Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal and informal education. *Journal of Science Education and Technology, 16*, 171–190.

[14]. Bransford, J. D., Brown, A. L., & Cocking, R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

[15]. Hofstein, A., & Rosenfeld, S. (1996). Bridging the gap between formal and informal science learning. *Studies in Science Education, 28*, 87–112.

[16]. Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (Eds.). (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: National Research Council.

[17]. Nadelson, L. S., & Richard Jordan, J. (2012). Student attitudes toward and recall of outside day: An environmental science field trip. *The Journal of Educational Research, 105*, 220-231.

[18]. Prensky, M. (2006). *Don't bother me mom – I'm learning.* St. Paul, MN: Paragon House.

[19]. Jonassen, D. H., & Howland, J. (2003). *Learning to solve problems with technology: A constructivist perspective*. Upper Saddle River, N.J.: Merrill Prentice Hall.

Development and Evaluation of Interactive English Conversation Learning System with a Mobile Device Using Topics Based on Learners' Life

Kae NAKAYA[a], Masao MUROTA[a]
[a]*Dept. of Human System Science, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Japan*
{knakaya, murota}@mr.hum.titech.ac.jp

Abstract: Motivating Japanese EFL (English as a Foreign Language) learners is essential if they are to learn English conversation effectively. We developed an English conversation mobile learning system based on a heuristic model of variables influencing the concept of "Willing to Communicate" by MacIntyre et al, with "interactive English conversation", "topics based on interests and life of learners", and "applications which learners can use anywhere and anytime". The results of experiments showed that the "interactive English conversation" style worked well in making learners feel relaxed and making them speak fluently.

Keywords: English Conversation Learning, Lifelog, Second Language Acquisition, Mobile Learning

Introduction

Japanese EFL (English as a foreign language) learners often have low motivation towards English learning. A Benesse Corporation survey held among junior high school students showed that 60 percent think of themselves as not good at English [1]. Similarly, many surveys of Japanese college students report that more than 65 percent of them dislike or hold a low motivation towards learning English [2][3].

In particular, Japanese EFL learners feel strong anxiety towards speaking English in the classroom, which causes hesitation in speaking [4]. As a result, they have low motivation to express what they think. The National Institute for Educational Policy Research shows that more than 60 percent of junior high school students don't like to speak or write what they think [5]. Therefore, they are not active in English conversation learning.

To motivate learners to learn English conversation, it is important to increase their "willingness to communicate" (WTC). According to the "Heuristic Model of Variables Influencing WTC" by MacIntyre et al, WTC directly affects the frequency of L2 (Second Language) use [6]. Thus, more active communication in the L2, and by extension, improvement in speaking ability, is achieved by increasing the learner's WTC. WTC is influenced by two immediate situational factors, which are the "desire to communicate with a specific person" and the "state of communicating self-confidence".

The "desire to communicate with a specific person" is used to mean that people want to communicate with a person who they know well, who holds similar opinions to them, is somebody who they often meet, or are attracted to. When people enjoy talking, the conversation becomes interactive. People tend to talk more if the interlocutor agrees with what they say and an immediate reply is given. If the conversation is strongly-interactive and comfortable, people will desire to talk more.

The “state of communicating self-confidence” is affected by whether one has experienced the situation or developed knowledge about the topic at the moment of conversation. People can communicate with self-confidence on topics they know well. However, when the situation is unfamiliar to them, their confidence is reduced. In the case of Japanese EFL learners, the effect of L2 communication confidence on the WTC is particularly strong [7].

English conversation classes in Japan do not, however, do enough to promote learners’ WTC. Learners only practice and imitate conversation examples from textbooks during class. The materials of the classes encourage passive learning, so learners tend to only reproduce the contents and not express themselves. The materials are not as interactive and attractive as learning with real interlocutors.

In addition, the topics used in class do not necessarily motivate all learners because English conversation teachers can only offer a few topics in one class. Pino suggests that in order to encourage learners to speak English more, teachers should choose topics that are related to learners’ own experiences and avoid ritual domain [8].

In this paper, we developed a system based on the model by MacIntyre et al, which aims to increase motivation of Japanese EFL learners. We propose pseudo-interactive and agreeable English conversation to motivate the “desire to communicate with a specific person”. We also propose topic-based English conversation about learners’ daily lives to make learners acquire the “state of communicating self-confidence”. We propose 2 kinds of contents: topics based on learners’ lifelogs, and topics based on common situations learners face.

To encourage learners to talk in English more frequently, we developed an Android application that learners can use to talk whenever and wherever they want. Japanese EFL learners rarely talk in English in daily life because Japan is a linguistically homogeneous nation. Therefore, with respect to this point, (we believe) the use of a portable device is more suitable than a stationary PC.

The targets of this system are bachelor and masters course students who have achieved a TOEIC level C score [9]. The students are supposed to have already learned all basic grammar, but tend to be reluctant to speak in English.

1. Related work

1.1 Pseudo-interactive English conversation materials

ELIZA[10] is a computer program which offers pseudo-conversation. It replies to users using pattern matching and, in some cases, makes users feel as if they are interacting with another person.

SpeakGlobal is a website that offers learners interactive English conversation learning with robots using artificial intelligence [11]. To communicate, the robots use speech recognition and speech synthesis technology. This way, robots can talk with their own voice, and learners can reply with theirs, making the conversation feel intimate because of its interactivity. In other words, learners are motivated to talk because they can talk directly with the robot and not have to use the keyboard.

However, both of them cannot offer interaction based on the learner’s life. The pattern matching of ELIZA is primitive. Therefore the program can only offer general responses and the interaction often becomes unnatural after some time. As for SpeakGlobal, these robots can only talk about general topics and cannot adapt to learner diversity. In contrast, in order to offer interactive English conversation learning that aims to motivate learners to speak more English, our system features topics based on learners' lifelogs.

1.2 Language learning environments that have a learner-centered design

Much research has tried to tailor materials to a diverse set of learners in order to motivate them to learn English. These kinds of materials are based on a "learner-centered design" that personalizes each material's content. This research has collected learners' data, such as their profiles, location, time, and learning log as parameters, for the personalization of the learning materials. They proposed learning systems based on leaners' data and offer learning contents that are more related to learners [12][13][14]. Personalization is an important factor in the design of language learning materials.

Recently, mobile learning systems have become a hot topic in this field, with many research papers on mobile learning [12][13][14][15]. Mobile devices enable systems to collect learners' data effectively, while also enabling learners to learn anywhere and anytime, because of the compact nature of these devices.

We have adapted personalization and learner-centered design to English conversation learning because present research in mobile language learning has predominantly focused on improving vocabulary skill.

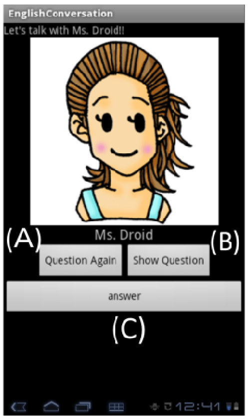


Figure 1 : Basic mode screen

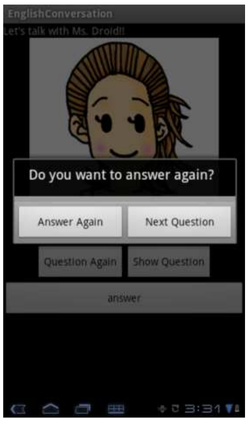


Figure 2 : Confirmation mode screen

2. Design of English conversation learning

2.1 The features of the system

Our system offers a type of English conversation that is closer to in-person conversation. Acting as an interface to the system, the Android application asks a question that the learner answers immediately. Moreover, in the case of "yes" or "no" questions, it includes a branch containing the next question or comment. Learners can therefore feel as if they are talking with the application interactively.

In addition, because conversation topics are related to learners' daily life, the learning is personalized to every learner. If the topics are general, learners are sometimes unfamiliar with them, and, as a result, are reluctant to speak. However, learners can talk more comfortably and with greater self-confidence in our system because of the personalized nature of interaction.

2.2 The procedure of English conversation learning with the system

Conversation learning occurs with the interaction between learner and system. First, after the learner launches the application, the application asks a question based on the learner's life which the learner listens to. The learner can listen again by pressing the "Question Again" button (Figure 1(A)). If the learner wants to read a transcript of what the application says, pushing the button shown in Figure 1(B) will display the transcript onscreen. Pushing the button again will hide the transcript.

After pushing the "answer" button (Figure 1(C)), the learner replies using English speech recognition. The learner has a chance to reply again during the confirmation mode (Figure 2). Third, the application speaks the next question. One instance of conversation learning consists of about 20 interactions. We call this one conversation session.

3. Design of the system

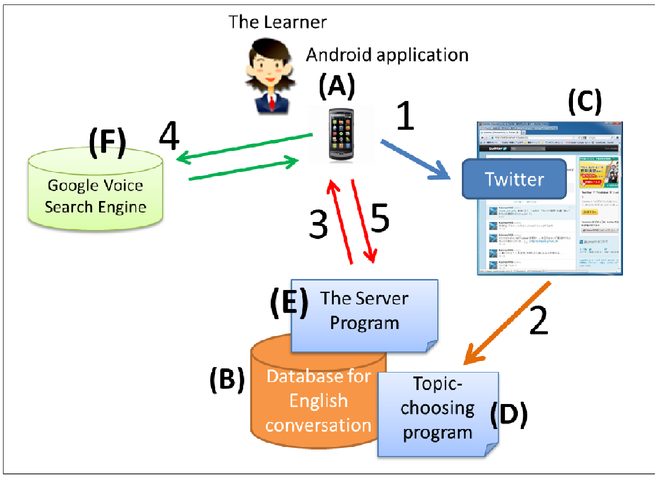


Figure 3 : English conversation learning process

3.1 The outline of the system

Our system offers pseudo-interactive English conversation, and consists of the elements shown in Figure 3. When the learner launches the application, the server sends a question based on the learners' life from the database to the application (Figure 3(3)). The application vocalizes the question using speech synthesis. If the learner replies in English, the application accepts the speech input using Google's Voice Search application (Figure 3(4)) and sends the input to the server (Figure 3(5)). The server then sends the application a new question or comment based on the learners' reply (Figure 3(3)). By repeating this process, the learner can practice English conversation.

1	What classes did you have today? Please tell me one of them.
Example learner responses	I had a database class today.
2	What did you learn from your class today?
Response	I learned SQL commands.
3	Was it interesting for you?
Response	Yes, the class was very interesting.
4	That's good! So you enjoyed the class, right?
Response	Yes, it was a fun.
5	How many students take the same class?

Figure 4 Conversation example in the case of "yes"

1	What classes did you have today? Please tell me one of them.
Example learner responses	I had a database class today.
2	What did you learn from your class today?
Response	I learned SQL commands.
3	Was it interesting for you?
Response	No, it's boring.
4	Oh, really? So please continue studying, and you will feel the class is getting interesting!
Response	All right. I'll do my best.
5	How many students take the same class?

Figure 5 Conversation example in the case of "no"

3.2 *Design of sentences for the conversation*

In this system, we offer "pseudo-interactive English conversation" in three ways. First, we arrange the order of questions or comments in a session to make learners feel as if they are "talking naturally" with the application whatever they speak (an example can be seen in Figure 4 and Figure 5). Second, as the accuracy of voice recognition in Android is not acceptable for longer sentences, the system has a branch for the next question or comment in the case of the learner's reply including "yes" or "no", which are correctly recognized. If the learner replies "yes" to the question (Figure 4 (3)), another question is asked (Figure 4 (4)). Third, the comments include phrases that aim to cheer learners up, and make them feel more relaxed (Figure 4 (4), Figure 5 (4)).

Table 1: List of conversation categories

classes, homework, sports, job hunting, research, conference, laboratory, part-time job, daily life, TV game, TV program, travel, cooking, pets, presentation, house-moving, seminar, shopping, music, mobile phone, movie, diet, comic books, programming, favorite books

3.3 *Topics based on the learner's lifelog*

We prepared 420 sentences categorized into 25 categories of topics relating to the daily life or interests of the learners (Table 1). Each category comprises of several conversation sessions. The system chooses the appropriate category based on the learner's lifelog, which is collected from the learner's twitter account (Figure 3(2)). The system calculates the degree of similarity between the categories and the obtained lifelog, and then chooses the category with the highest score. In case all scores are zero, the system chooses the "Daily Life" category.

Table 2 : List of conversation topics based on commonly occurring situations

order in a restaurant, reserve a room of a hotel, tell what you feel to a doctor, talk in an office, talk in a sightseeing tour, self-introduction, talk with staff in a shopping-mall, talk in the airport, talk in a bank or post office, talk about party invitation

3.4 *Topics based on situations that often happen to learners*

We prepared 10 kinds of topics based on situations that often happen to learners, such as conversations in a restaurant or an airport (Table 2). The system offers 2 topics a day at random. First the system explains to the learner the situation behind the conversation. After that, the system offers a situative question and the learner replies.

4. Evaluation

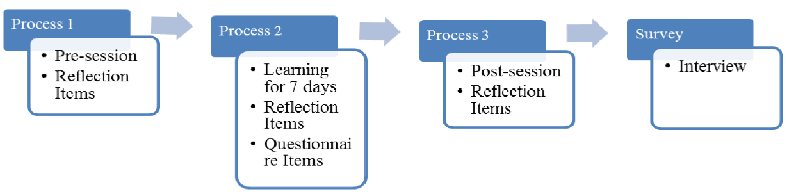


Figure 6 : Experimental procedure

4.1 The aim of the experiments

We conducted an experiment for 7 days, aiming to observe changes in attitude during English conversation. The subjects were split into two groups, one using topics based on learners' lifelogs, and one using situated general topics. After they had used the application to practice English conversation, we measured the differences between the two groups.

4.2 Reflection Items and Questionnaire Items

To observe the 2 kinds of effectiveness, we collected reflections and a questionnaire after every conversation session about willingness of English conversation and about their impression towards the sentences, using a 5-point Likert scale. We defined reflection items concerning willingness as "Reflection Items", which include fun, anxiety, expressing oneself, motivation, relaxation, and fluency. We defined question items about the subjects' impression of English sentences as "Questionnaire Items", which include interests, daily life, friendliness, difficulty with understanding, and difficulty in replying.

4.3 The outline of the experiments

We conducted the experiment from January 17th to 23th, 2012. The subjects were 8 bachelor and masters course students (7 male and 1 female). The mean age was 22.5 (*SD*=1.1). Four subjects used the application with lifelog based topics, and the others used one with general situative topics. We analyzed all data except for 1 subject whose TOEIC score didn't match the target of this research.

Table 3 : Mixed two-way factorial ANOVA table for the Reflection Part

Reflection Items	Source of variation	Sum of squares	DOF	Mean square	F-value	p-value
Fun	Between topics	0.0134	1	0.0134	0.0131	0.9132
	Between terms	0.1801	1	0.1801	0.5368	0.4966
	Interaction	0.4301	1	0.4301	1.2822	0.3089
Anxiety	Between topics	0.8932	1	0.8932	0.3890	0.5602
	Between terms	0.8932	1	0.8932	2.7796	0.1563
	Interaction	0.3575	1	0.3575	1.1125	0.3398
Expressing oneself	Between topics	0.9301	1	0.9301	0.7030	0.4400
	Between terms	1.2515	1	1.2515	5.6142	0.0640 +
	Interaction	0.1801	1	0.1801	0.8077	0.4100
Motivation	Between topics	0.3810	1	0.3810	0.7748	0.4190
	Between terms	0.1205	1	0.1205	0.1948	0.6774
	Interaction	0.3348	1	0.3348	0.5411	0.4950
Relaxation	Between topics	0.8218	1	0.8218	0.5833	0.4795
	Between terms	2.2057	1	2.2057	5.9397	0.0589 +
	Interaction	0.4557	1	0.4557	1.2272	0.3184
Fluency	Between topics	0.9301	1	0.9301	0.4933	0.5138
	Between terms	5.5372	1	5.5372	29.8636	0.0028 **
	Interaction	0.5372	1	0.5372	2.8973	0.1495

Table 4 : A part of mixed two-way factorial ANOVA table for the Questionnaire Part

Questionnaire Items	Source of variation	Sum of squares	DOF	Mean square	F-value	p-value
Daily Life	Between topics	0.8932	1	0.8932	1.0400	0.3546
	Between terms	3.0807	1	3.0807	7.3115	0.0426 *
	Interaction	0.4557	1	0.4557	1.0816	0.3460
Difficulty with understanding	Between topics	1.3393	1	1.3393	0.8855	0.3899
	Between terms	2.7515	1	2.7515	11.0984	0.0207 *
	Interaction	0.2515	1	0.2515	1.0144	0.3601
Difficulty in replying	Between topics	2.1488	1	2.1488	0.7243	0.4336
	Between terms	2.1488	1	2.1488	15.1681	0.0115 *
	Interaction	0.2917	1	0.2917	2.0588	0.2108

4.4 The results of the analysis

4.4.1 Reflection Items

We defined the mean value of the 1st and 2nd day as “the value in the beginning” and the mean value of the 6th and 7th days as “the value at the end”. We statistically compared the value at the beginning with the one at the end to observe any improvement in willingness. The result was analyzed by ANOVA with two independent valuables: topics (lifelog-based and situative general) and terms (beginning and end).

Statistically or marginally significant improvements were observed for three Reflection items on the term factor (Table 3). “Fluency” showed improvement with significance ($F(1,5)=29.8636$, $p<.01$). “Express yourself” and “Relax” exhibited a marginally significant improvement ($F(1,5)=5.6142$, $F(1,5)=5.9397$, $p<.10$ for both). All the values of the three items at the end are higher than the ones in the beginning. In addition, we analyzed the value of Reflection Items from the pre- and post-session of general topics. The Wilcoxon signed-rank test for both applications did not reveal any prominent changes.

4.4.2 Questionnaire Items

The analysis showed statistically significant changes at three items on the term factor. “Daily Life”, “Difficulty with understanding”, and “Difficulty in replying” significantly improved ($F(1,5)=7.3115$, $F(1,5)=11.0984$, $F(1,5)=15.1681$, $p<.05$). The values of both topics at the end are lower than the ones in the beginning.

4.5 The results of the survey

We took a survey of the subjects’ impression and suggestions for improvement of the application after the experiment. The survey showed that the subjects became more relaxed during conversation or motivated to speak English after 7 days of learning with the application because subjects had to think and reply immediately by “speaking” to the system’s spoken questions. Moreover, some subjects said that if they were to use the application over a longer term, it might decrease their anxiety.

On the other hand, sometimes the subjects felt that the interaction was unnatural in an obvious way. Moreover, the subjects that used lifelog-based topics often found topics to be the same or too similar. Most subjects wanted functionality to check and improve their English. They wanted the application to check the grammar of what they spoke, record and playback the conversation, or practice vocabulary they did not yet know.

5. Conclusion and future works

5.1 The effectiveness of interactive conversation by speaking

According to the analysis of Reflection Items and the survey, we conclude that the pseudo-interactive English conversation which was offered in both topics worked effectively in making learners feel that they were speaking English fluently, in a relaxed manner, all using the their own English knowledge. The reason is that every time the system sends a question or comment, and the learner replies immediately, the system replies instantaneously. The subjects repeated this processes many times and became used to expressing themselves in English. The result that showed statistically significant decreases in “difficulty with understanding” and “difficulty in replying” supports this conclusion.

5.2 Necessity to improve the use of lifelog for English conversation learning

We have to improve the way we offer topics and English sentences based on learners' life and improve lifelog collection. This is because the Questionnaire Items of "Daily Life" showed a statistically significant lower change especially around topics based on the lifelog obtained from subjects' tweets. The usefulness of the Twitter lifelog strongly depends on how learners express their life or how often the learners post tweets. The application that offered the lifelog-based topics sometimes chose the same topic. Therefore some subjects became disinterested. We have to not only propose another way of collecting lifelogs but also other ways of offering topics that interest learners or learners want to talk about.

5.3 Support learners in learning English

As many subjects said, the system needs features for learning English, such as grammar checking, offering practice vocabulary for new words, and reflecting on the conversation. These features will motivate learners to not only learn English conversation but also acquire new English skills during conversation learning.

References

[1] Benesse Corporation. (2009). The first survey on junior high school English (in Japanese). http://benesse.jp/berd/center/open/report/chu_eigo/seito_soku/pdf/data_00.pdf(Accessed May 21, 2012)

[2] Maeda, K., Tsumura, S., Koiso, K., Kagada, T. (2010). Survey on English learning motivation –using text mining- (in Japanese). A collection of theses of Osaka University of Commerce, 6(2), 77-90.

[3] Tsumura, S. (2011). Relationship between self-esteem and demotivating factors in English learning. In A. Stewart(Ed.), JALT2010 Conference Proceedings. Tokyo: JALT. 806-815.

[4] Hojo, R. (1996). A Study of Students' Anxiety over Classroom English (6). Bull. Joetsu Univ. Educ., 15(2), 495-506.

[5] National Institute for Educational Policy Research. (2012). The survey about writing English (in Japanese). http://www.nier.go.jp/kaihatsu/tokutei_eigo_2/tyousakekka.pdf (Accessed May 21, 2012)

[6] MacIntyre P. D., Clèment R., Dörnyei Z., Noels K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation, In Modern Language Journal, 82, 545-562.

[7] Yashima, T. (2002). Willingness to Communicate in a Second Language: The Japanese EFL Context. The Modern Language Journal, 86, I, 54-66.

[8] Pino, C. (2009). Overcoming Japanese EFL Learners' Fear of Speaking. University of Reading Language Studies Working Papers, 1, 55-63.

[9] Educational Testing Service. "PROFICIENCY SCALE" , http://www.toeic.or.jp/toeic/pdf/data/proficiency.pdf (Accessed May 21, 2012).

[10] Weizenbaum, Joseph. (1966). ELIZA – a computer program for the study of natural language communication between man and machine. Communications of the ACM, 9, 1, 36-45.

[11] SpeakGlobal. http://speakglobal.co.jp/ (Accessed May 21, 2012)

[12] Ogata, H., Li, M., Hou, B., Uosaki, N. & Yano, Y. (2011). Effectiveness of Ubiquitous Learning Log System. Proc. of the 19th International Conference on Computers in Education (ICCE 2011), Chiang Mai, Thailand. http://www.nectec.or.th/icce2011/program/proceedings/pdf/C4_S7_303S.pdf (Accessed May 21, 2012)

[13] Peterson, S. A., Markiewicz, J. & BjØrnebekk, S. S. (2009). Personalized and Contextualized Language Learning: Choose When, Where and What. Research and Practice in Technology Enhanced Learning, 1-42.

[14] Edge, D., Searle, E., Chiu, K., Zhao, J. & Landay, J. A. (2011). MicroMandarin: Mobile Language Learning in Context. Proc. Of the CHI 2011 Conference on Human Factors in Information Systems, 3169-3178.

[15] Stockwell, G. (2010). Using Mobile Phones for Vocabulary Activities: Examining the Effect of the Platform. Language Learning & Technology, 12(2), 95-110.

The Effect of Learning Spatial Geometry By Mobile Devices

Lin-Jung WU [a*], Sin-Chuan LAI[a] , Kuo-En CHANG[a] , Yao-Ting SUNG[b] & Hsien-Sheng HSIAO[c]

[a] *Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan*

[b] *Department of Educational Psychology and Counseling, National Taiwan Normal University, Taiwan*

[c] *Development of Technology Application and Human Resource, National Taiwan Normal University, Taiwan*

*ljungwu@gmail.com

Abstract: The aim of this research is to develop a hands-on spatial geometry learning system to facilitate students' geometry learning. The system is developed with Duval's four critical elements of geometry apprehension including perceptual apprehension, sequential apprehension, operational apprehension and discursive apprehension. The system supports senior high school students in the process of spatial geometry problem-solving, allowing them to hands-on manipulate spatial geometry graphics and develop their visualization and mental imagery. In total, 58 participants from different classes were recruited. The experimental group used the hands-on learning system, whereas the control group followed the traditional paper-based approach. The study investigates the effects of the hands-on geometry learning system on students' perceptual apprehension, sequential apprehension, operational apprehension, overall spatial geometry scores and learning attitude. The results revealed more learning attitude, and higher perceptual apprehension, sequential apprehension, operational apprehension in the experimental group.

Keywords: spatial geometry, mathematics, mobile devices, mental imagery, imagery and visualization.

Introduction

The visual dimensions of mathematical learning and the value of visual-spatial thinking increasingly have been acknowledged as essential to mathematics education (Clements & Battista, 1992; Zimmermann & Cunningham, 1991). Conceptualization of geometrics is the basis to develop the mental ability to think in abstract terms by using real-world objects as examples to learn to think in terms of abstract objects and finally postulate and prove formally defined objects (Zhou, 1999; Do & Lee, 2009). Spatial ability is a critical skill in geometric learning. Several studies investigate how to improve spatial abilities (Hannafin, 2004; Hannafin, Truxaw, Vermillion, & Liu, 2008; Kaufmann & Schmalstieg, 2003, Yang & Chen, 2010).

Actions of dissection, rearrangement and recomposing promote the imagination and logical thinking through observation and analysis (Clements & Battista, 1992). However, in real class interactions, teachers often are not able to holistically present visualized views to students because of the lack of props as useful supplements, creating misinformation for students. They are often restricted by the formal definitions as well as the logical order of operations putting them in a position that focuses less on the meaning of geometry that begins with observation, operation, categorization, and organization (Battista, 1994). A mathematical concept or problem presented only in textual data are often overtly abstract and cannot be easily understood (Guttiérez, 1996). When teachers focus strictly on mathematical logical signs and describing them by their definitions, concepts, or properties, it might not be effective to help students build a thinking system to construct a spatial image and to manipulate it when trying to solve a problem in 3-D geometry (Garrity, 1998; Gurny, 2003). Skemp (1987) pointed out that, by directly introducing definitions and formula to teach mathematical concepts might be the most concise way for the teachers; however, it is often at the detriment of their students. When acquiring a brand new concept, most students start their first step from the actual scenario of situations (Anna Sfard, 1991).

Krutetskii (1976) had pointed out that reasoning based on visualization was not unique. Different students can visualize the same problem in different logical reasoning ways. Focusing on just one single point of view will make a stalemate (geometrical rigidity) (Hoz, 1981), a state where students dwell around one single point of the visualized geometric image, unable to reason from different viewpoints.

Presmeg (1986) had suggested the following points for the visualization:

i. Single sample of the illustration was often misleading.
ii. Standardized images would easily limit students' understanding of the non-standardized geometric shapes
iii. Fixed images that cannot be manipulated could limit students' ability to reason

Without focused and detail-oriented reasoning analysis, it was often ineffective for students to understand the reasoning process. Even through, it made misunderstanding for students.

Previous studies on spatial geometry emphasized geometric learning activities should focus on the actual observations and operations (Bishop, 1989; Grand, 1990). Duval (1995) stressed the complex or abstract mathematical concepts specific to visually present (for example: space coordinate system), or the entity manipulation aids, inspiring students to develop their spatial visualization. Bishop (1980) believed there are two particular abilities suitable for dealing with geometric objects. The first ability is to interpret the image information. It is involved in visualizing process, one can manipulate and converse the visual representation with the corresponding mental image. The second one is the ability to manage the visualization process. This is an ability to understand how to use shapes, diagrams, and descriptive narratives to present geometry concepts. Clements (1979) emphasized that learners with good visualization and mental imagery would be able to observe different perspectives of objects, judge and manipulate the mental images of those objects.

Duval (1995) discovered that figure through the process of manipulation, description, and reasoning can deepen our understanding of the four critical elements of geometry apprehension (perceptual apprehension, sequential apprehension, operational apprehension and discursive apprehension). Perceptual apprehension is that which we recognize and distinguish properties of shapes. Sequential apprehension is through construct figure from properties. In this case the figural units depend not on perception but on technical constraints (e.g. ruler and compass, primitives in computer software) and mathematical constraints. Discursive apprehension is that we compose family of figures and analyze families of figures. In any geometrical figure the perceptual recognition of geometric properties depends on discursive statements, via descriptive words to explain the properties of geometric system and to reason based on these descriptive constructs. Operative apprehension depends on the way in which a given figure is modeled, for example by dividing it into parts, or by transforming it optically or changing its orientation in the plane. These modifications can be performed mentally or physically. Duval (1995) suggested that operating on a figure in certain ways is critical in the heuristic process and that pupils need to be taught how to do this. However, operative apprehension is not independent of the other apprehensions, discursive and perceptual apprehensions very often obscure operative apprehension.

It is suggested that work with computers may support the development of sequential apprehension and it also might encourage the development of operative apprehension if the software has been designed with this in view (Duval, 1995). Therefore, students will have more sufficient experiences in manipulating geometric shapes and visualizing the figures in order to develop the ability to reason from geometric figures. Computers can be the supplemental tool in assisting students to experiment, trial out, and comprehend the properties of geometry (Lin, C. P., Shao, Y. J., Wong, L. H., Li, Y. J., Lin, C.P. et al., 2011). Moreover, when conducting geometry instructions, teachers should emphasize on how to develop students' ability in visualizing geometric spatial relationships as well as presenting the concepts both holistically and independently (Guttiérez, 1996). Using computers to provide multiple angles of the diagrams, in additional to being able to freely manipulate, and observe geometric relationships can facilitate visual geometry learning and teaching.

Following the advances of technologies, multimedia aided learning system has increasingly been the focus in the pedagogical area. Dixon (1997) discovered that GSP (The Geometer's Sketchpad) had helped grade eight students in understanding conceptual constructs of rotation and mirror images of geometric diagrams. Berta Tünde (2002) believes that computers can present different visual forms of geometric structures and characteristics which benefited students in solving geometric problems. Osta (1998) discovered that computers can provide a rich interaction between learners and the diagrams. Chang, et. al. (2007) proposed that multimedia materials have an evidential effect on students' ability to visualize, analyze, describe, reason, and organize geometric information. The National Council of Teachers of Mathematics even suggested to utilize interactive computer software applications for students to learn geometry (NCTM, 2000).

Since information technology allows people to communicate concepts of multiple views via 3-D that are difficult to convey with traditional 2-D illustrations, it can help students develop the ability to think geometrically and sharpen their skills on geometric studies. Moreover, many researchers believe that when students transit to an area of higher mathematics, the ability to visualize spatial relationships becomes even more important since the complexities of visual recognition would have increased dramatically (Smith, 1964; Fennema, 1977). It highlights the importance of spatial geometry teaching and learning.

Compared to personal computers, mobile devices (PDA, Tablets, etc.) are portable and capable of performing calculation tasks which make them ideal as supplementary tools for classroom learning (Trimmel & Bachmann, 2004; Hennessy, 2000; Sung, Y. T., Chang, K. E., & Wu, L. J., 2007). They enable one to interact directly with what is displayed, rather than indirectly with a cursor controlled by a mouse. Secondly, it lets one do so without requiring any intermediate device that would need to be held in the hand.

According to these reasons, this research has utilized mobile devices to develop interactive geometric learning practice system, allowing students to actively perform operations on the live shapes as a method of practicing, guiding students' cooperation and discussion to elevate the students' spatial and geometrical learning results.

1. Hands-on Spatial geometry learning activity and GeoPlay system

GeoPlay (Spatial geometry hands-on learning system) and its spatial geometry learning curriculums are designed based on the four elements of logical reasoning Duval (1995) had proposed. In order to meet the learning needs, the experimental tools developed in this research have been used in a classroom setting helping students to immediately apply and build their logical reasoning roadmaps after the instructor has provided formal instructions. The research focus on students' actual practice to facilitate them to develop the other three apprehensions. The learning activities are as follows:

- Perceptual apprehension vs. questions with text-format: When learners click the sample exercises, the first problem will be presented as text format (Figure 1). As in the general text questions, learners can think and generate answers accordingly from the description of questions. Their perceptual apprehension will form through learners' understanding of the text description. They can solve problems by their mental imagery, however, their mental imagery is not necessarily the same as the given graph, which may have the wrong perception of understanding. If learners choose to answer the question in this stage, the system will immediately give feedback after learners sent the answer. Learners may share the feedbacks with peers or teachers.
- Sequential apprehension vs. questions with figures-format: Younger students might particularly face greater challenges in understanding abstract geometry concepts or problems simply by giving them text descriptions. The process of individuals constructs a figure by using visualization tools (such as paper and pen, computers, or generating image in mind) to aid in analyzing, interpreting, and learning concepts is a kind of sequential apprehension. Constructing figures can be used as a tool to understand abstract geometric concepts (Yakimanskaya, 1991) or to reduce the individual burden of the working memory area. Therefore, learners can choose to click the "Figure" button to transfer text-format questions into figure-format through GeoPlay system (Figure 2). And they can use handwriting function to mark some important key information on the figures. If learners cannot convert the text description to the figures; instead of, they can create their own visual image assisted by given figures.
- Operational apprehension vs. questions with dynamic manipulation: Duval (1995) observed that the figure can be manipulated, and through changing the figures in different ways, learners will get operational understanding to help them solve difficult geometry problems. Operative apprehension is a kind of cognitive process that individuals transfer figures to mental imagery. These modifications can be performed mentally or physically. Thus, if learners still cannot think and reason through the figure-format questions, they can click the "Manipulate" button. System will show the solid geometry components constructed form Cabri 3D which can provide learners manipulate by touch pen (Figure 3). Through manipulating the geometry components (changing the size or its orientation in the plane), they will easily start to think and reason possible answers of the questions. Learners also can improve their skills of mental imagery and visualization through their manipulating to the solid geometry components.
- Discursive apprehension vs. solving questions with dynamic reasoning: Mathematical properties represented in a drawing cannot be determined through perceptual apprehension. A graphic without denomination or hypothesis is an ambiguous representation, so that not everyone will see the same things or the same properties. A graphical presentation of identification will affect the individual perception. That is, although individuals can simply perceived by the figure nature of some of the geometric, but may be wrong. Therefore, some must first be given through speech (denomination and hypothesis) and others can be derived from the given properties. In other words, the discursive apprehension of individual of the figures is not based only on perceptions of figure symbols have yet supplemented by narrative text. Discursive apprehension is the cognitive processes that individual describes a figure through language or narrative text or use the text representation to reason. Therefore, students can click the "Solution" button and show the Flash animation walkthrough reasoning problem-solving process to help learners build their own reasoning abilities.

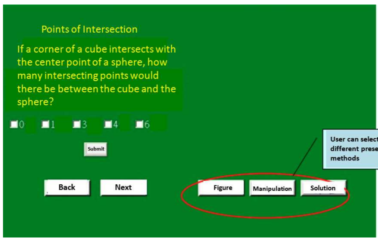


Figure 1. Text presentation of Question

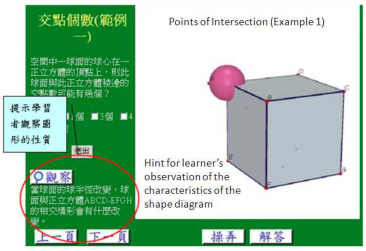


Figure 2. Pictorial Description of the geometric figures

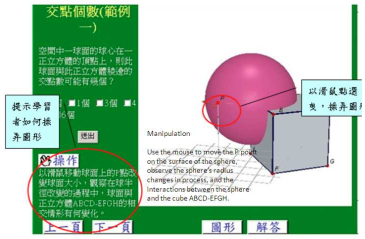


Figure 3. Manipulation of the geometric figures

2. Methodology

2.1 Participants

In order to determine how effectively the Geoplay improve the ability to think geometrically among senior high school students, the study uses the pre- and post-test quasi-experimental design. The study selected two senior high school classes in Taipei City which divided into an experimental group and a control group. There were 31 students in the experimental group (19 boys and 12 girls) and 27 in the control group (13 boys and 14 girls).

2.2 Experiment

A quasi-experimental design was used in which the independent variable was the group (control group or experimental group). The dependent variable was the post-test score for each learning part and the pretest score was the covariance. The experimental group used Geoplay to practice the activities of geometric learning while lecture-based instruction was used for the control group. Moreover, experimental and control group had the same teacher and learning material. After the experimental treatment, the ANCOVA analysis will use to evaluate the significant effects of different learning outcomes.

2.3 Material

2.3.1 Handheld mobile learning software for spatial geometry

The study selected three units from senior high school spatial geometry curriculum, space coordinate system, the spatial relationship between line and plane, spatial intersection between shapes. The students in the experimental group practiced with Geoplay system. Each student was required to complete every activity.

2.3.2 Pre-test and post-test

This experiment involved conducting tests to find out the changes in students' geometry ability after practicing with every activity. The questions were selected from the three units and organized into pretest and post-test. The questions in the pretest and post-test were based on the same concepts and belong to the same question types. There were 7 questions about perceptual apprehension, 11 questions about the sequential apprehension, and 7 questions about operational apprehension, total 25 questions. The questions about perceptual apprehension aimed to assess students' basic space mapping capabilities. This part of the composition is unique and different students will not have much different ideas in the graphics. The questions about sequential apprehension were to assess students' visual ability (visualization) of geometry figure. This part of the composition is not unique, different students may have different ways of composition. The questions about operational apprehension tested students' ability to manipulate graphics in geometry problem solving procedure. Students are required to reason according to the description of questions, based on composition, reasoning and then manipulation of visual graphics, the last inference the answer.

2.4 Procedure

Prior to the experiment, both groups were conducted a 40-min pretest, after which the formal experiment was performed. The students in experimental group practiced geometry learning activities by using GeoPlay, and control group of students involved in a lecture-based instruction. After completing the experiment, each participant was given a post-test.

3. Results

To understand the effect of "Spatial geometry hands-on learning system" (GeoPlay) on students learning of spatial geometry, the single-factor analysis of covariance (ANCOVA) was used to analyze the change of students' pretest and posttest scores. The pretest scores as a covariance to analyze the difference between posttest within the two groups. The results were divided into four part (perceptual apprehension, sequential apprehension, operational apprehension, and overall spatial geometry scores), and were analyzed if posttests in experimental group and control group had significant difference in each part.

The test scored one point for each question, perceptual apprehension part contained 7 points, sequential apprehension part contained 11 points, and operational apprehension part included 7 points, total 25 points. Table 3 lists the pre- and post-test scores of experimental group and control group.

Table 3 Average and standard deviation of test scores for the experimental group and control group in the pretests and post-tests

Groups	N	Perceptual	Sequential	Operational	Overall
Experimental	31	4.61(1.75) 5.61(1.38)	8.45(1.63) 9.45(1.43)	2.79(1.38) 3.29(1.42)	15.86(3.27) 18.36(3.47)
Control	27	4.26(2.57) 4.82(1.98)	7.96(1.81) 8.48(1.76)	2.09(1.46) 2.17(1.49)	14.32(4.91) 15.46(4.38)

The mean and SD values of the posttest in both experimental group and control group had improved after experiment treatment (see Table 3). Thus further analysis of pre- and posttest in both groups to investigate whether there is a significant difference between the two groups.

One-way ANCOVA was used to identify significant differences between experimental and control group post-test scores for perceptual apprehension, sequential apprehension, operational apprehension, and overall spatial geometry scores after eliminating the effects of the pretest scores. First, the homogeneity of the regression coefficients of the total post-test score and post-test score for each part were tested. The F values for the homogeneity of the regression coefficients on the post-test scores for the four parts match the basic hypothesis of the homogeneity of the regression coefficient ($F(1,54) = 1.77, 0.07, 0.15, 0.01$; $p > .05$). ANCOVA could be used to examine these data. The result of ANCOVA showed the scores in perceptual apprehension were not reach statistical significance ($F(1,55) = 3.71$, $p=.06>0.05$) between experimental group and the control group, but the result in sequential apprehension, operational apprehension and overall spatial geometry scores were significantly higher ($F(1,55) = 4.23$, $p=.05$; $F(1,55) = 4.701$, p=.034<.05; F (1,55) = 8.111, p=.01<.05) in the experimental group than in the control group.

4. Discussion and Conclusions

First, the result showed that the neither experimental group nor the control group showed a significance in their perceptual apprehension. The cause of the inconsistency may be that students may solve problems with visual approach, or even they solve problems directly using memorized formulas. Therefore, there is no significant difference on perceptual apprehension between the experimental group by using system and the control group by using paper-based instruction. Second, there are significant differences in sequential apprehension between two groups. This finding is consistent with the researches of Battista (2002) and Clements (1997). GeoPlay guides students in developing visual thinking by means of presenting illustrations of a variety of objects. This method is equally beneficial for developing a student's visual thinking through the use of physical teaching tools to guide the student's geometric thinking. Third, part of visual manipulation, most are more complicated geometric problems, requiring much higher level of multi-cognitive analysis. Students must reason and solve problems by using mental and visual method (visual imagery). Students must use Duval's (1995) operational understanding of the visual image manipulation to reasoning problem-solving. The results showed that performance in operational apprehension part of the experimental group and control group, significant differences can be inferred. By using the "Spatial geometry hands-on learning system" (GeoPlay), there are positive effects on spatial geometry teaching. It's also helpful for the operation of the visual imagery of students' problem-solving. It is in line with Osta (1998) presented that interactive operating environment is conducive to the development of the visual capabilities, such as mental imagery and visualization. Fourth, the Overall spatial geometry scores showed that the experimental group and control group were significantly different. It can be inferred using the "Spatial geometry hands-on learning system" (GeoPlay) can facilitate spatial geometry teaching. Also, through the dynamic manipulation environment provided by the system, the mechanism helps to improve students' spatial geometry learning. It is in line with Battista (2002) theory of geometry learning, and also consistent with previous research investigated that computer interactive learning environment can assist the development on spatial skills (Berta Tünde , 2002; Dixon, 1997). Sixth, the results of the learning attitude questionnaire showed that students hold a positive attitude about the use of "hand-held mobile learning tool in the spatial geometry learning system" (GeoPlay). Most students showed the positive agreement for the use of computers to more clearly show three-dimensional geometry materials, and the manipulation of graphics also help build spatial concepts. In particular, the ratio of up to 93.6% considered that computer images can be shown more specific in three-dimensional geometry, 80.6% of the proportion thought that "hands-on manipulation" will be the best way to help understand the three-dimensional geometric problems, which show that the system can indeed provide effective support in the three-dimensional geometry learning.

References

[1] Battista, M. T. (1994). On Greeno's environmental/model view of conceptual domains: A spatial/geometric perspective. *Journal for Research in Mathematics Education*, *25*(1), 86–94.

[2] Battista, M. T. (2002). Learning geometry in a dynamic computer environment. *Teaching Children Mathematics*, *8*(6), 333-339.

[3] Bishop, A. J. (1980). Spatial abilities and mathematics education---A review. *Educational Studies in Mathematics, 11*, 257-269.
[4] Bishop, A. J. (1989). Review of research on visualization in mathematics education. *Focus on Learning Problems in Mathematics*. Winter Edition 1989, 11, 7-16.
[5] Clements, M. A. (1979). Sex differences in mathematical performance: An historical perspective. *Educational Studies in Mathematics, 10*, 305-322.
[6] Clements, D. H., & Battista, M. T. (1992). *Geometry and spatial reasoning*. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning, 420-464. NY: Macmillan.
[7] Clements, D. H., Battista, M. T., Sarama, J., & Swaminathon, S. (1997). Development of students' spatial thinking in a unit on geometric motions and area. *The Elementary School Journal, 98*, 171-186.
[8] Dixon, J. K. (1997). Computer use and visualization in students' construction of reflection and rotation concepts. *School Science and Mathematics, 97*(7), 352-358.
[9] Duval, R. (1995). Geometrical picture: kinds of representation and specific processings, *in Exploiting Mental Imagery with Computers in Mathematic Education* (Sutherland & Mason eds.), Springer, 142-157.
[10] Fennema, E. & Sheman, J. (1977). Sex-related difference in mathematics achievement, spatial visualization and affective factors. *American Educational Research journal, 14*. 51-71.
[11] Garrity, C. (1998). Does the use of hand-on learning, manipulatives, improve the test scores of secondary education geometry students? (ERIC Document Reproduction Service No. ED 422179).
[12] Gurny, H. G. (2003). High school students' performance on Vandenberg's Mental Rotation Test: Art ability, gender, activities, strategies and ease of taking the test (ERIC Document Reproduction Service No. ED 479372).
[13] Guttiérez, A. (1996). Visualization in 3-dimensional geometry: In search of a framework. In L. Puig & A. Guttiérez (Eds.), *Proceedings of the 20th conference of the international group for the psychology of mathematics education* (pp. 3–19). Valencia: Universidad de Valencia.
[14] Hennessy, S. (2000). Graphing investigations using portable (palmtop) technology. *Journal of Computer Assisted Learning, 16*, 243-258.
[15] Hoz, R. (1981). The effects of rigidity on school geometry learning. *Educational Studies in Mathematics, 12*, 171-190.
[16] Chang, K. E., Sung Y. T., and Lin, S. Y. (2007). Developing Geometry Thinking through Multimedia Learning Activities. *Computers in Human Behavior, 23*(5), 2212-2229.
[17] Kaufmann, H. & Schmalstieg, D.(2003). Mathematics and geometry education with collaborative augmented reality. *Computers & Graphics 27*, 339–345.
[18] Krutetskii, V.A. (1976). *The psychology of mathematical abilities in school-children*. Chicago: University of Chicago Press.
[19] National Council of Teachers of mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.
[20] Lin, C. P., Shao, Y. J., Wong, L. H. Li, Y. J., Niramitranon, J. (2011). The Impact of Using Synchronous Collaborative Virtual Tangram in Children's Geometric.
[21] Osta, I. (1998). CAD tools and the teaching of geometry. In Carmelo Mammana & Vinicio Villani (eds.), *Perspectives on the teaching of geometry for the 21th century*. 128-144.
[22] Presmeg, N.C. (1986). Visualization in high-school mathematics. *For the Learning of Mathematics, 6*, 42-46
[23] Sfard, A.(1991). On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin. Educational *Studies in Mathematics, 22*, 1-36.
[24] Skemp, R. (1987). *The psychology of learning mathematics*. New Jersey: Lawrence Erlbaum Associates.
[25] Sung, Y. T., Chang, K. E., & Wu, L. J. (2007). A new perspective of instructional design for classroom e-Learning: Case studies of mobile learning. Paper presented at the 2007 Mobile and Ubiquitous Learning Workshop, August 15-17. Hong Kong, University of Hong Kong.
[26] Trimmel, M. & Bachmann, J. (2004). Cognitive, social, motivational and health aspects of students in laptop classrooms. *Journal of Computer Assisted Learning, 20*, 151-158.
[27] Tünde, B (2002). Combination of traditional and computer based tools in mathematics education. *Proceeding of 2002 International Symposium Anniversary of Pollack Mihály Engineering Faculty*, 1-9.
[28] Yakimanskaya, I. S. (1991). The development of spatial thinking in schoolchildren (R. H. Silverman, Trans.). (Vol. 3). Chicago: *National Council of Teachers of Mathematics.*

An Implementation of Smartphone-enabled Seamless Learning: a Snapshot Perspective

Elizabeth KOH* & Chee Kit LOOI
National Institute of Education, Singapore
*elizabeth.koh@nie.edu.sg

Abstract: In the "WE Learn Project" at Nan Chiau Primary School, a smartphone-enabled seamless learning curriculum was employed to support the Ministry of Education's Masterplan 3 goals in primary schools. Using a snapshot perspective, this paper presents analyses and findings from the first trial curriculum unit in Primary 3 English. Briefly, the functional snapshot shows that the smartphone-enabled curriculum enhanced the academic achievement of students compared to the traditional worksheet based curriculum.

Keywords: mobile, seamless learning, primary school, English language, smartphones

1. Introduction

Smartphones are being adopted en-masse throughout the world (IDC, 2012). In a similar parallel, more and more smartphones are being harnessed for education (Norris, Hossain, & Soloway, 2011). Past lessons on ICT adoptions in education have taught us that learning and teaching with the tool is more important than the tool itself. Seamless learning has been conceived as an influential pedagogical framework in smartphone-enabled learning (Chan et al., 2006; Looi et al., 2010; Wong & Looi, 2011). In Singapore, the Ministry of Education (MOE) has developed strategic masterplans by which ICT can be harnessed to enrich the learning landscape (MOE, 2008). One of the key goals in the latest masterplan, the Third Masterplan for ICT in Education (mp3), is to encourage the development of self-directed learners. Self-directed learning is an important skill for the students of today in order for them to gain ownership of learning, monitor their learning and manage new situations they encounter (Tan, Divaharan, Tan, & Cheah, 2011). Some research has reported that self-directed learning also leads to higher academic achievement (Zimmerman, 1989).

Aligned with the goal of implementing seamless learning and cultivating self-directed learners, smartphones were adopted in a trial design and implementation of the Primary 3 English language curriculum at Nan Chiau Primary School. This "WE Learn Project" is a scaling up of the seamless learning initiative in Primary 3 Science (Looi et al., 2011; Zhang et al., 2010). By transforming the classroom from the traditional teacher-centered model to a learner-centered one, the project hopes to enhance the learning outcomes of students. The project is still in-progress and this paper reports on the pioneer smartphone-enabled unit in the English curriculum.

Being a preliminary analysis of the ongoing project, this paper examines student academic achievement during the trial implementation. The research question is "how has the smartphone-enabled implementation affected students' academic achievement?" The snapshot perspective is the lens utilized to analyze the implementation. The snapshot perspective is based on snapshot theory (Murray, 2006) which denotes that there are particular junctures in a project implementation whereby the entities are conceptually whole, providing for discourse and discussion.

The paper starts with a brief literature review of smartphone characteristics, seamless learning, and the theoretical lens, the snapshot perspective. Next, the research design is described. This is followed by the analysis of the smartphone-enabled implementation in the school. The paper will end with further discussion, limitations of the study and the conclusion.

2. Literature Review

2.1 Smartphone characteristics

Smartphones have several characteristics. Cochrane and Bateman (2010) identify 16 types of affordances of the smartphone which they capture in a rubric. They include: image capture, video capture, mobile web experience, text entry, GPS, touch screen, application availability, 3G, WiFi, cost, availability in the country, screen size, and portability. These affordances differ among different smartphone devices. The rubric serves to identify which device is more beneficial for adoption in the classroom. Similarly, Looi et al. (2010) identify 12 affordances for smartphone use in school. They are: platform, form factor, mobility, connectivity, applications, voice, battery, durability, cost, support, features (such as camera, pen-based input, ease of use on the interface, voice or audio), and memory storage. These affordances affect the type of learning activities that can be carried out.

Rather than looking at specific features of the phone, other researchers have examined broader characteristics of the smartphone. Patten et al. (2006) develop a functionality framework based on pedagogy and educational applications for handheld devices. They identify 7 categories which are progressive in nature; that is, each latter category consists of earlier functionality. Moreover, the first few categories contain features that are available in laptops or desktop computers, while the latter ones are unique to handheld devices. The categories are: administration (e.g. calendars, grading), referential (e.g. dictionary, e-books), interactive (e.g. drill and test, graphing), microworld (e.g. models of real world domains), data collection (e.g., note taking, sensor readings), location aware (e.g. museum guide, augmented environments), and collaborative (e.g. co-present games, collaborative environments). In summary, there are many characteristics of the smartphone. These can all be harnessed for various pedagogical designs.

2.2 Seamless Learning

Seamless learning has been conceptualized in the mobile learning context (Chan, et al., 2006; Looi, et al., 2011) and is a broad pedagogy that interacts with technology, teaching and learning (Wong & Looi, 2011). Its chief tenant is that learning is a continuous process across formal and informal learning environments (Looi, et al., 2010). Wong and Looi (2011, p. 2367) further unpack 10 dimensions of mobile seamless learning as follows:

1) Encompassing formal and informal learning
2) Encompassing personalized and social learning
3) Across time
4) Across locations
5) Ubiquitous knowledge access
6) Encompassing physical and digital worlds
7) Combined use of multiple device types
8) Seamless switching between multiple learning tasks
9) Knowledge synthesis
10) Encompassing multiple pedagogical or learning activity models.

A key backdrop of seamless learning is the changing of existing instructional designs for improved, innovative learning. To enhance pedagogical models and teaching practices, seamless learning invites participants to remove constraining seams (e.g. conceptual, cultural, and physical) for continuous and sustained learning.

The tenets of seamless learning have been examined in several studies. Sandberg et al. (2011) examined English learning as a second language for 5th grade Dutch students across three conditions: lessons in class, lessons in the zoo with a mobile device, and lessons in the zoo with students allowed to take the device home for a fortnight. The group which took the mobile device home had the highest results. When time was controlled for, there were no differences among the groups. It seems that mobile devices help to motivate students to use their out of class time to learn. Besides motivating students, artifacts created by students in seamless learning environments are also important (Wong, Chen, & Jan, 2011). Lin and Hsiao (2011) examine how visual aids (still images and dynamic animation) on the mobile device affect the learning of English vocabulary. They found that animations helped high school students learn English verbs better. It suggests that seams from older pedagogical models can be removed to enhance learning.

2.3 Snapshot Approach

The theoretical lens adopted in this paper is based on snapshot theory (Murray, 2006). The snapshot theory is derived from the Computer Science discipline whereby software programmers use snapshots at critical junctions to generate discourse in a software development. Entities in a snapshot must be conceptually whole in order to provide a frame for discussion. This research adapts key concepts in snapshot theory and utilizes it as a framework for analysis. Snapshot theory proposes that explaining a phenomenon requires a series of snapshots that have particular characteristics and relationships. These snapshots typically start with an *infrastructure* snapshot (main features of the tool), followed by an *advanced infrastructure* snapshot (further knowledge of the infrastructure composition, enriching the older snapshot), a *functional* snapshot (how the tool functions with the features of the tool), and an *example* snapshot (how the tool works). Snapshots can be weak or complete. A weak snapshot denotes incomplete insight while a complete snapshot encompasses all details to explain the phenomenon.

3. Activity Implementation and Research Design

The project implementation team involves a project manager, two English teachers, the English subject head, a curriculum designer and two allied educators. The team decided to start with the mobilization of a unit in the Primary 3 English curriculum, which had the theme of “mystery”. The smartphone-enabled curriculum was rolled out to three classes in late March 2012 by two teachers. The entire smartphone-enabled curriculum was taught in 12 periods (6 hours) over 2 weeks to the classes. These three classes will be hereby known as red, blue and green. One teacher taught two of the classes (class Red and Blue) while the other taught the class Green. As this was the pioneer activity, an allied educator and the curriculum designer were present during the lessons to support the teachers.

A mixed methods study was designed involving qualitative and quantitative data. Academic achievement is measured in terms of students’ grades. For quantitative data, the implementation team designed an assessment to measure students’ English academic learning. The test had 5 sections covering the following: Vocabulary, Identification of noun and verb, Tenses, Metaphors, and Antonyms. The total score was 30.

A pretest-posttest design was conducted for the smartphone-enabled classes. Two other non-smartphone-enabled Primary 3 classes were chosen as control. These classes used the traditional curriculum which was worksheet-based. For the control classes, only a post-test was administered. The pre-test was administered before the smartphones were used for any learning activity. The post-test was administered after the unit was taught, two weeks later. The questions in the pre-test and post-test were identical. Students took about 30 minutes to complete the test.

For the qualitative data, during this first stage of the project, the researcher took the stance of an observer as participant (Gold, 1985; Kawulich, 2005). Basically this variant of the participant observer methodology allows the researcher to be made known to the members of the implementation team. The researcher's main role is to collect data and the researcher is not considered as a member of the team. The team controls the level and amount of information given to the researcher. It is membership at the periphery. Data was collected from official team meetings, teacher reflections and several lesson observations that the researcher was allowed to observe. The researcher also had informal conversations with members of the team. Field notes were written and data was triangulated.

4. Analysis of Smartphone-enabled English Seamless Learning

4.1 Infrastructure snapshot

This pertains to the tool, the smartphone employed in the project. Here is a list of the main affordances of the smartphone:

- Platform: Windows Phone 7.5
- Form factor: lightweight
- Image capture: back-facing camera
- Internet connectivity: via 3G and WiFi
- Touch screen
- Voice: recorder
- Video: recording and streaming
- Applications: Among other applications available in the Windows marketplace, the project had a specially designed suite of software termed myDesk with three main applications
    - Map-It: mind-mapping application
    - Sketchbook: a drawing tool
    - Blurb: structured note-writer
    - These suite of software was supported by a myDesk learning management system for teachers to view, manage and grade students work

4.2 Advanced Infrastructure snapshot

This pertains to how the features of the smartphone can be harnessed for English language learning. In this study, each student had a smartphone, and the device was with them 24 x 7. This pedagogical choice, the dimension of across time in seamless learning, encouraged students to use the smartphones for learning in class and out of class.

Vocabulary: The unit started with the teacher reading a story which had a mystery theme. Students mapped the story using the smartphone application "Map-It". This encouraged them to remember the new vocabulary. In addition, students used the dictionary application to search for the meaning of new words. Students seamlessly switched between the learning tasks, from the storytelling activity, to dictionary search due to the availability of the smartphone and its applications. Teachers also encouraged students to audio record themselves reading a passage using the audio recorder and to search for the word meanings at home, encouraging self-directed learning and informal learning.

Identification of noun and verb: Using the application Blurb, students were given words such as "spy" and tasked to write sentences using the word as a noun and as a verb. Students were tasked to write a few sentences in class and to write a few more sentences after the class, at their own time. Students were also challenged to come out with these kinds of words, and write a sentence. This learning activity emphasizes seamless learning encompassing formal and informal environments.

Tenses: Students helped each other to take a photo of each other using the camera function and then used Sketchbook to create a disguise. Students learnt about tenses as they annotated what they did in Sketchbook. For instance, after drawing curly blue hair on herself, the student wrote, "I drew a wig on my hair". While this activity was dominantly about personalized learning, to a certain extent it had an element of social learning as students had to cooperate in taking the picture.

Metaphors: For the activity on similes, students either took a picture or drew an image of the simile and annotated it e.g. as busy as a bee. Once again, teachers encouraged students to do this activity out of class. Several students took pictures of ants and flowers from the school garden or at home. Others took pictures of animals like pigs or bees from objects at home. This activity encouraged seamless learning in formal and informal contexts.

Antonyms: Using Map-It, students identified and classified positive and negative antonyms. Students started the activity in class and were encouraged to continue it after school hours, allowing formal and informal learning.

4.3 Functional snapshot

How did the smartphone function to help in student's academic achievement? The paper provides a functional snapshot through examining the quantitative results.

4.3.1 Pretest and Posttest results of smartphone-enabled classes

There were a total of 114 students in the 3 smartphone-enabled classes. Students on average scored 22.69 for the pre-test and 25.88 for the post-test, a difference of 3.18 between the two tests. A paired samples t-test showed a significant difference of $p<.001$ between the two tests. This indicates that the smartphone-enabled curriculum improves the English content knowledge of students.

A sectional analysis was performed and there were significantly higher scores for section 1, 3, and 5(vocabulary, tenses, and antonyms), with most improvements in section 3 and 5. However, there was no significant improvement in sections 2 and 4 (identification of noun and verb, and metaphors). A possible implication derived from the results is that these aspects are difficult for Primary 3 students to grasp. The intervention seems to help students gain more vocabulary but they may not know how to use these words correctly.

The data was compared across the three classes and slight differences were found. For class Red, there was a significant improvement for vocabulary and tenses but little improvement for identifying noun and verbs, metaphors and antonyms. Class Blue had the most improvement, demonstrating significant increases for sections 1, 3, and 5. Class Green showed significant improvement for the understanding of tenses and antonyms but no significant increase for vocabulary. There was also a decline in scores for sections 2 and 4.

As the school practices ability-grouping, class Red was regarded as high ability, class Blue as mixed ability, and class Green as lower ability. These differences in results across classes could be due to the student ability levels and their prior knowledge. For instance, high ability students already have a good grasp of English content and so did not learn much more for antonyms during the lessons. Mixed ability students who may not have much prior knowledge were able to gain the most from the smartphone-enabled curriculum as seen by the higher number of sections that improvement was shown. For class Green, the lack of prior knowledge could have affected their results.

Another possible reason could be how the teachers taught the unit. Teachers gave different amount of challenges and tasks to students. For instance, Class Red was given 5 words to write sentences in nouns and verbs. For Class Blue and Green, students chose one word only. Given a similar amount of time, teachers did not enact the smartphone-enabled curriculum in the same pace, catering to the learning abilities of the students in the class. It could be that given more time and tasks, class Green could have had similar results as the other classes. Nevertheless, the mixed results for class Green suggests that revised strategies in helping lower ability students are needed, especially for students to grasp difficult concepts such as metaphors.

4.3.2 Posttest results of Smartphone-enabled and Non-smartphone-enabled classes

Two other classes acted as control classes (68 students). In the control classes, students were taught the same content using the traditional teaching method which is dominated by worksheets. The mean score for the post-test for the smartphone-enabled classes was higher at 25.88 compared to the mean score of the control classes, 21.25. Welch's t-test was performed between the experimental classes and the control classes. The test showed that the mean scores were significantly different at $p<.001$. This suggests that the smartphone-enabled curriculum intervention helps students gain higher test scores.

Similarly, a sectional analysis was conducted. Students in the experimental classes had higher means for all sections compared to the control. All sections were significantly different except for section 4. This suggests once again that metaphors are a challenge for Primary 3 pupils to understand. Indeed, upon hearing this, the teachers agreed at once how it was conceptually difficult and explained that this is the first time students were taught this. Nevertheless, the smartphone-enabled curriculum compared to the non-smartphone-enabled curriculum enabled the students to improve on the other aspects.

While these results are positive, they could have been skewed due to one of the control classes being of a much lower ability. A further statistical test was conducted to compare between class Blue and a control class which was regarded as being slightly higher in ability. The total score for class Blue was 25.84 while the control class scored 25.86. There were no significant differences between the total scores of the two classes. However, sectional comparison showed significant differences for vocabulary and antonym learning. The smartphone-enabled class had significantly higher scores in these two areas. On the other hand, the control class had significantly higher scores than the smartphone-enabled class in sections 2 and 3. There was no significant difference for scores in section 4.

Based on these results, the smartphone-enabled curriculum seems to help mixed ability students attain scores on par to their higher ability peers. However, in terms of emphasis, the smartphone-enabled curriculum helps to build content knowledge more than the application of that knowledge. As this is a trial phase, revisions to the smartphone-enabled curriculum are needed to enhance students' application of the content.

5. Further Discussion, Limitations and Conclusion

5.1 Smartphone characteristics

One of the instrumental affordances of the smartphone as described earlier is connectivity. In this implementation, the smartphone was connected to the Internet which allowed students to utilize many online applications. Moreover, the specially designed suite of learning software was web-based and relied on Internet connectivity. This design allowed the teachers to access student submissions easily through a web portal. However, during the trial, the Internet server became unreliable and students could not use the web applications. Over the weeks, these technical issues were gradually resolved via great effort by various members and extended members of the team. Teachers also adapted to the various situations and used the smartphone where possible. This highlights the importance of inter-related technological systems in the infrastructure snapshot.

5.2 Seamless Learning

In this implementation, seamless learning has been enacted in various ways. The smartphone-enabled curriculum appropriated several dimensions of seamless learning as espoused by Wong and Looi (2011). In particular, the dimension of formal and informal learning bears further discussion. As noted in past literature, there are many definitions of these terms. Some definitions have sharply delimited the boundary as the physical school environment (Spikol & Milrad, 2008), while others focused on learner autonomy (Looi, et al., 2010) or other variations (Chen, Millard, & Wills, 2008). In this study, informal learning was still more teacher-led. While such informal learning is noteworthy, the study could go further to develop and encourage more student-led incidental learning, to bring about deeper seamless learning.

Nevertheless, seamless learning using the smartphones has generated an incidental motivation effect. Teachers noticed an enthusiasm in the students to re-do their assignments. After going through the errors and misconceptions of student's submitted assignments, many students came forward to submit their assignments again. This was relatively easy for students' to do as they could just access the assignment online on their smartphones and click a button to submit. The increased motivation for the learning tasks could be one of the processes that the smartphone-enabled curriculum could have brought about and which led to the positive results for student's English learning.

5.3 Future Directions, Limitations and Conclusion

As the first trial implementation, the results provide support that the smartphone-enabled curriculum can help students to improve their academic English. Specifically, the implementation seemed to be able to help students in building vocabulary and understanding the different types of tenses. However, the results must be interpreted with some caution. There were several limitations in the rigor of the test. Firstly, there was a short duration of 2 weeks between tests and what is reflected in the test may not be internalized by the students. Second, the presence of other helpers during the smartphone-enabled curriculum could have influenced results such as the additional attention paid to the student by the allied educator and curriculum designer. Third, the smartphone-enabled classes took the pre-test before which could have pre-conditioned them for the same post-test.

Going forward, for greater evidence of student's learning, more specific examples of individual students' learning are needed. The snapshot analysis provided did not provide an example snapshot partially due to the lack of access to students at this juncture. The processes between the activity facilitated by the smartphone and how the student learns will be examined in future. In addition, the team intends to routinize instructional activities to help the student focus on learning with the tool. Furthermore, a clearer pedagogical direction is needed to guide the process of learning such that key goals are met.

These results and analysis provide crucial feedback of the progress of the trial implementation of the smartphone-enabled Primary 3 English curriculum. The snapshot perspective, in particular, allows crucial junctures of the implementation to be analyzed for further discourse. In addition, from the results of the survey (the functional snapshot), the team, especially the teachers, were encouraged that what they were doing made an impact in their students' learning. The team is in the process of mobilizing other units for implementation in the classroom. Research is underway analyzing the various ways and levels the smartphone adoption has impacted the school.

John Dewey, an American philosopher and educator once said, "Education is a process of living and not a preparation for future living." As the team revises the curriculum, they are becoming more cognizant of identifying the processes of learning. They are learning new things everyday just as their students are gaining a better understanding and knowledge of the English language. Indeed, this is another fruitful dimension of "WE Learn".

Acknowledgements

Many thanks go to key sponsor, Qualcomm, Inc., the implementation team and extended members at the school. In addition, much appreciation goes to all the researchers and consultants involved.

References

Chan, T.-W., Roschelle, J., Hsi, S., Kinshuk, Sharples, M., Brown, T., . . . Hoppe, U. (2006). One-to-one technology-enhanced learning: an opportunity for global research collaboration. *Research and Practice of Technology Enhanced Learning, 1*(1), 3-29.

Chen, W.-P., Millard, D., & Wills, G. (2008). *A Four Dimensional Model of Formal and Informal Learning*. Paper presented at the The 16th International International Conference on Computers in Education. http://eprints.soton.ac.uk/266632/

Cochrane, T., & Bateman, R. (2010). Smartphones give you wings: pedagogical affordances of mobile Web 2.0. *Australasian Journal of Educational Technology, 26*(1), 1-14.

Gold, R. L. (1985). Roles in sociological field observations. *Social Forces, 36*, 217-223.

IDC. (2012). Smartphone Market Hits All-Time Quarterly High Due To Seasonal Strength and Wider Variety of Offerings, According to IDC Retrieved 27 May 2012, from http://www.idc.com/getdoc.jsp?containerId=prUS23299912

Kawulich, B. B. (2005). Participant Observation as a Data Collection Method *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* (Vol. 6).

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

Looi, C. K., Zhang, B., Chen, W., Seow, P., Chia, G., Norris, C., & Soloway, E. (2011). 1:1 mobile inquiry learning experience for primary science students: a study of learning effectiveness. *Journal of Computer Assisted Learning, 27*(3), 269-287. doi: 10.1111/j.1365-2729.2010.00390.x

MOE. (2008). Masterplan 3, from http://ictconnection.edumall.sg/cos/o.x?c=/ictconnection/pagetree&func=view&rid=665

Murray, A. R. (2006). *Discourse structure of software explanation: snapshot theory, cognitive patterns and grounded theory methods.* Ph.D. Dissertation, University of Ottawa, Ottawa.

Norris, C., Hossain, A., & Soloway, E. (2011). Using Smartphones as Essential Tools for Learning: A Call to Place Schools on the Right Side of the 21st Century. *Educational Technology, 51*(3), 18-25.

Patten, B., Arnedillo Sánchez, I., & Tangney, B. (2006). Designing collaborative, constructionist and contextual applications for handheld devices. *Computers & Education, 46*(3), 294-308. doi: 10.1016/j.compedu.2005.11.011

Sandberg, J., Maris, M., & de Geus, K. (2011). Mobile English learning: An evidence-based study with fifth graders. *Computers & Education, 57*(1), 1334-1347. doi: 10.1016/j.compedu.2011.01.015

Spikol, D., & Milrad, M. (2008). Physical activities and playful learning using mobile devices. *Resarch and Practice in Technology Enhanced Learning, 3*(3), 275–295.

Tan, S. C., Divaharan, S., Tan, L., & Cheah, H. M. (2011). Self-directed learning with ICT: Theory, Practice and Assessment. Retrieved from http://www.ictconnection.edumall.sg/ictconnection/slot/u200/mp3/monographs/self-directed%20learning%20with%20ict.pdf

Wong, L.-H., Chen, W., & Jan, M. (2011). How artefacts mediate small-group co-creation activities in a mobile-assisted seamless language learning environment? *Journal of Computer Assisted Learning*, no-no. doi: 10.1111/j.1365-2729.2011.00445.x

Wong, L.-H., & Looi, C.-K. (2011). What seams do we remove in mobile-assisted seamless learning? A critical review of the literature. *Computers & Education, 57*(4), 2364-2381. doi: 10.1016/j.compedu.2011.06.007

Zhang, B., Looi, C.-K., Seow, P., Chia, G., Wong, L.-H., Chen, W., . . . Norris, C. (2010). Deconstructing and reconstructing: Transforming primary science learning via a mobilized curriculum. *Computers & Education, 55*(4), 1504-1523. doi: 10.1016/j.compedu.2010.06.016

Zimmerman, B. J. (1989). Models of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research and practice* (pp. 1-25). New York: Springer-Verlag.

Android-based Mobile Assessment System

Mahtab DALIR[a*], Heiko RÖLKE[b], Björn BUCHAL[c]
[a, b, c] German Institute for International Educational Research, Frankfurt am Main, Germany
*dalir@dipf.de

Abstract: Mobile devices such as smartphones and tablet PCs have gained global popularity and are increasingly used in all areas of daily life, including learning and assessment activities. To support different kinds of learning and assessment content, a versatile and comprehensive system is desirable. In addition, other aspects of long-term assessments have to be covered such as security, data protection and adaptivity, e.g. to new schedules or items. In this paper, we present the concept and realization of an Android-based mobile assessment system especially designed for school usage. It has been used already in several research studies in elementary schools, e.g. to measure daily fluctuations of cognitive performance capabilities.

Keywords: Mobile learning, mobile assessment, ambulatory assessment, Android, QTI

Introduction

Modern daily life is more and more influenced by ubiquitous usage of smartphones and tablet PCs. This "not being tied to particular locations and times" [1] is not bound to specific domains or areas of application. The mobility is especially interesting in the areas of learning and assessment, particularly for, but not limited to, schoolchildren.

For certain kinds of studies it is important to be able to test on a recurring basis and more than once a day. In addition, the actual testing should be as smooth and non-interruptive as possible to avoid organisational problems and to ensure validity of the data. To cope with this general challenge, mobile technology can be a solution. Until recently, it was hard to find devices that were mobile, suitable for children and both affordable and powerful enough for testing purposes. As pointed out above, this has changed now. Using smartphones or tablets for assessment triggers a positive side-effect: Children are excited about using "cool" technology which might increase their motivation. Our group was asked for technical support for a psychometric study of daily fluctuations in children's cognitive performance in the school context. The general goal was to assess children's cognitive resources three times a day for several weeks within the school year. Approximately 120 children aged between 9 and 10 years were planned to be tested.

To make this possible, we designed a versatile and comprehensive mobile assessment (and learning) solution. Our overall system is targeted to the school context but not limited to it. We did not restrict our system in terms of types of content and aimed for open standards. Special emphasis was put on test security, autonomous test delivery, and automatic data synchronization.

The first section of this paper presents requirements for a mobile assessment system, followed by the system architecture in Section 2. Afterwards, in Section 3, we outline a case-study where our system has been used extensively. The paper concludes with a section on related work and an outlook.

1. System Requirements

As outlined in the introduction our aim is to support mobile assessments, especially in the school context. The analysis of this aim revealed three major requirement areas:

- Content
  - Wide range of item types
  - External production of items
  - Usage of standards
- Hardware and Systems
  - Inexpensive, yet reliable
  - Easy access to system internals
- Assessment Delivery
  - Autonomous (unproctored) testing, particularly outside of school (afternoon, week-ends)
  - Assessment security and data protection

We will give an overview on the requirements in this section. Possible strategies to cope with them are outlined here. Our solution(s) will be discussed in detail in extra sections.

The most important part of a generic assessment system is the content it delivers and the way it is produced. The system has to support a wide range of item types. At the same time, it should follow existing standards for content description and interoperability. This ensures that already existing content may be reused and newly created one may be used also for other purposes.

Our choice was to use the IMS Question & Test Interoperability Specification (QTI) standard [2] as a starting point. While this standard has certain deficits, it allows for reusing external (PC-based) tools as well as some assessment content. In addition, a later integration of our system into learning management systems like Moodle [3, 4] will be easier.

An important decision for any mobile application is the choice of the underlying hardware (and software) system. While it is in principle possible to develop hardware-independent mobile software – e.g., using browser technology – in reality this conflicts with the requirement of test security and autonomous testing.

So we decided to develop a native application instead of a web application and had to choose a platform. There are plenty of choices like Apple's iOS [5], Android [6], Samsung's Bada [7], Windows Mobile [8], and BlackBerry [9] to name but a few. As we aimed at a maximum of choices regarding the hardware and programming possibilities we chose the Android platform.

Assessment security was already introduced above as an important requirement. It means that it has to be as hard as possible to cheat while testing, to fake the results, to copy the items, or to gain access to other people's data. In addition, we wanted to enable unproctored testing, e.g., on afternoons or week-ends. Therefore we need to be able to lock down the mobile devices. All processes have to be completely under our control: assessment times, data synchronisation, updates and so on. We will describe our choices and the overall system design throughout the remaining sections of this paper.

2. System Architecture

To fulfil our requirements, we developed the system architecture illustrated in Figure 1. The two main applications, executed on the smartphone, are the "Assessment Controlling Application" and the "Assessment Rendering Application". Both applications are developed to be executed on Android based mobile devices. The Assessment Controlling Application is responsible for scheduling the actual assessments and for ensuring assessment security and data protection. It also enables the smartphones to connect to the Classroom Server and therefore also to the Central Server to get updates and export test results. The Assessment Rendering Application renders and illustrates the assessments. It also displays results and feedback to the user.

In addition to these applications, we have developed portable servers for each classroom to gather test results and to provide the necessary updates for the smartphones. Each of these classroom servers connects to the central server to synchronise the data. We will now describe these system parts in detail throughout the next sections.

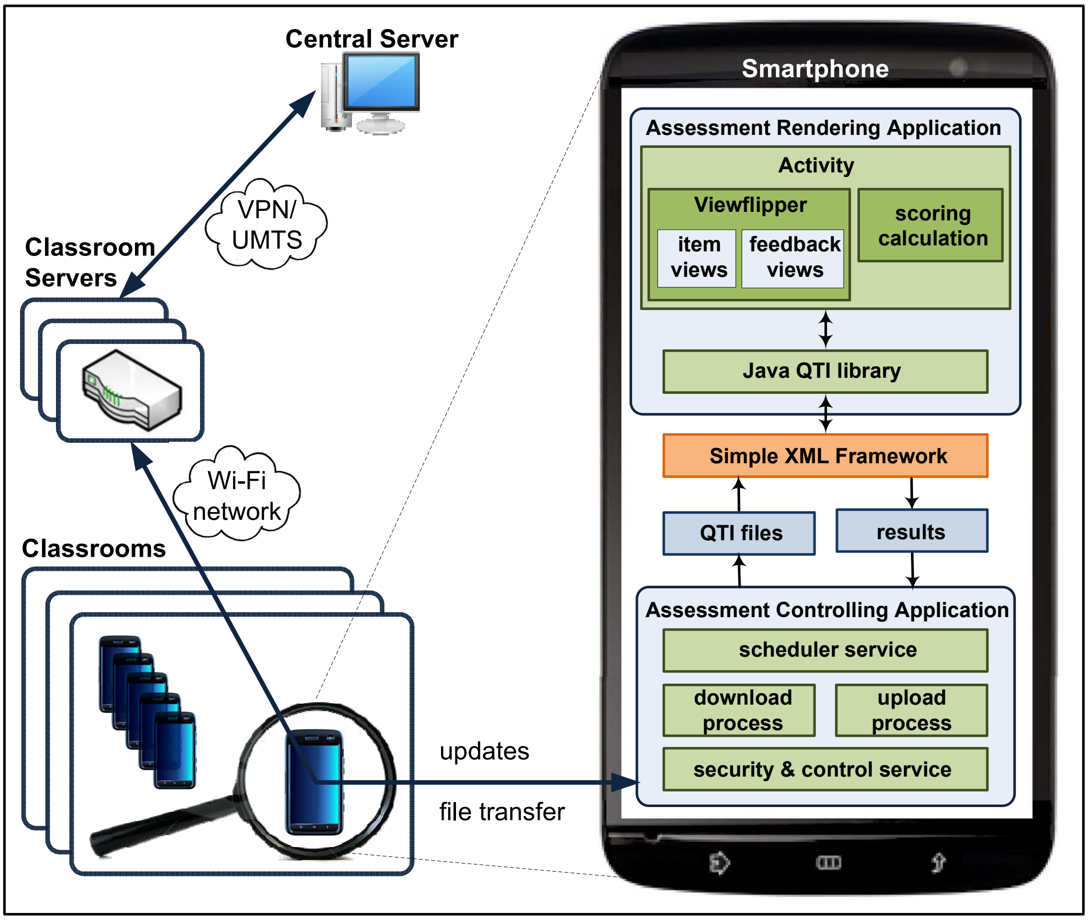


Fig. 1. The Mobile Assessment System

2.1 Assessment Rendering Application

Smartphones are built upon proven and existing software and technology and they are in principle compatible to usual PCs, but only up to a certain extent. There are some compatibility limitations such as screen- and memory size on the hardware side, and adaptations of the operating systems and programming languages that are directly supported on the software side. The content and the applications used for ambulatory assessment purposes therefore cannot be directly transformed from standard computerized assessment. Special design issues have to be considered. We deployed the Assessment Rendering Application upon Android system. The application delivers a framework to better support for offline ambulatory assessment and to minimize compatibility limitations.

2.1.1 *Assessment Tests and Items*

QTI is one of the most widely accepted specifications in the assessment and learning area. QTI defines a format for the representation of assessment content and results. It consists of a data model that defines the structure of questions, assessments and their results. Assessments can be built by combining several items (optionally combined in sections). Besides the actual item content, item descriptions specify whether the answers are used for response processing and whether feedback is provided within a test. QTI provides several templates for multiple-choice, matches, hot-spots, feedback etc [2].

2.1.2 *Android Platform*

Google's Android is a software stack for mobile devices. It is a combination of a free, open-source operating system for mobile devices, middleware and key applications. Android applications are written in Java as a programming language but executed by means of a custom virtual machine called Dalvik VM [10]. Dalvik is a JVM optimised for Android to cope with the mobile device limitations such as small memory, slow processor speed, and limited battery capacity. Every application runs in its own instance of the Dalvik VM [11]. Android applications consist of loosely coupled components and therefore are quite different from traditional (desktop) applications. The most important components of the Android SDK are Activities, Services, Content providers, Intents, and Widgets [12, 13]. The version we used for our implementation is Android 2.2.

2.1.3 *System Design of the Assessment Rendering Application*

The application is designed according to the Model-View-Controller (MVC) pattern [18]. This pattern separates the different aspects of the application and enables independent development, testing and maintenance. According to MVC pattern, the Activity components [17] are expected to take care of both the View and Controller tasks. In their role of a controller, they initialise the Model, provide the Accessors for getting data from the Model, and implement listeners. As a View they act like observers and notify automatically of any state changes.

Figure 2 illustrates the architecture of the application. It shows the main Activity of the application user interface, which includes a widget named Viewflipper [17]. A Viewflipper animates views that have been added to it. Views expose user interfaces for test items that handle screen layout and interaction with the user. To present the test items, elements such as widgets, buttons, images and texts are added to the hierarchical layouts within the respective view.

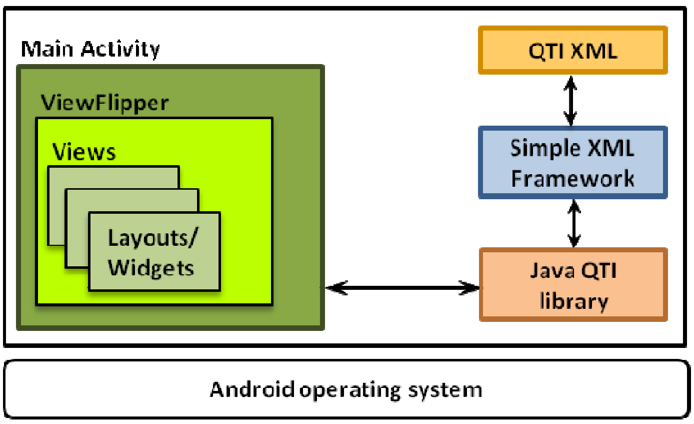


Fig. 2. Assessment Rendering Application

We use Simple [14], an XML serialization and configuration framework for Java, to process QTI XML. It supports the process of serialising Java objects to XML and de-serialising XML back to Java. Simple is provided for the Java platform and in comparison to the Android libraries available, it is easier to use and more flexible [15].

To enable interpretation and execution of the de-serialized files on a Java platform, we built a Java QTI library that reflects the QTI XML file structure. The Java QTI library serves as a domain model and illustrates both aggregations and lists of QTI tests and items. Differences to other existing libraries are described in the section on related work.

On Android, there are several ways to catch or handle the events generated by a user interaction with the touch screen of the smartphone [18]. For instance, touching a button is one possibility to generate an event within the user interface. A different possibility to capture the very same interaction (touching a predefined area of the screen) is to define hotspot points. We used the first method for implementing the widgets of the QTI interactive elements.

2.2 Assessment Controlling Application

This part of our Mobile Assessment System is a separate Android application that mainly works as a scheduler for executing instances of the Assessment Rendering Application, one for each assessment. Coordinating data transfer, managing security aspects and locking the phone are other jobs of this application.

An XML file describes the configuration and initialization settings. It contains information about the processes to be run and controls their behavior. A process can for example be an instance of the Assessment Rendering Application, a download- or an upload operation.

A scheduler file specifies the dates and time slots in which a process is to be run. There is also a timeout determined for test blocks. If the time is over, the test block is skipped. It is possible to define separate schedules and test materials for each user.

The Assessment Controlling Application establishes a wireless connection between the smartphones and the classroom servers. The connection is necessary to download new assessment content and software updates from the server to the phone and to upload the results of the assessment to the server. The smartphones need to be configured properly to be able to communicate with the classroom servers. For this purpose, there is a configuration file that contains information about the connection address, user name and password to access the server. It may also contain the user identifier. This id is used to distinguish between the uploaded files from different phones on the server. If no identifier is supplied, the device id is used which is unique for each phone.

As security plays a major role in unproctored testing, the Controlling System provides the essential security mechanisms to prevent gathering inaccurate data or faked results. An important element is changing the default home screen of the smartphone to our custom home screen. Unnecessary functionalities of the smartphones, which are not used for the assessments such as the camera, are permanently disabled. So the end-user is not able to manipulate the system.

2.3 Server Infrastructure

Besides implementing the client side of the system, developing the server infrastructure to support mobile assessment also posed a significant challenge. This includes an authentication system and a high security system for data synchronisation.

The classroom servers provide wireless LAN (WLAN) for the connection between the mobile devices and the classroom servers, based on the File Transfer Protocol (FTP). The connection between the classroom servers and the central server uses a point-to-point Virtual Private Network (VPN). It is not necessary to have a stationary internet connection in the classrooms as the classroom servers use mobile radio connections to connect to the central server. The smartphones are configured to connect to their classroom server on a regular basis to export the test results.

We decided to support two different connections between the classroom servers and the central server. The first (and preferred) possibility is to use the school's network and integrate the classroom server(s) into that network. Each classroom server establishes a connection to the central server via a VPN to ensure data security. The second networking possibility comes into play if the first connection fails or is not available in a certain school. The classroom servers are equipped with mobile broadband modems that support the advanced Universal Mobile Telecommunications System (UMTS) standard. This autonomous internet access provides the secondary VPN connection, which may also be run in parallel to the first.

The central server is equipped with two different network interfaces, one for each VPN connection. This is necessary to support both connections (UMTS and school network), because they are configured and rooted differently. Once at least one connection is established, it can be used not only for data synchronization or applying updates, it is also used to get secured remote access to the mobile server for administration or adjusting settings. For example, it might be necessary to delay the regular file transfer because of high network traffic caused by running backups.

The central server also provides a simple website. The website e.g. provides easy access for uploading updates or other files. These files will be transferred to the mobile server immediately and redistributed to the clients.

3. Case Study: FLUX

The project Assessment of Cognitive Performance FLUctuations in the School ConteXt (FLUX) deals with daily fluctuations in children's cognitive performance in the school context using mobile assessment techniques. The cognitive tasks such as working memory, processing speed, and updating as well as questions about mood, motivational aspects, and situational issues are being processed twice to four times a day for several weeks [19].

One idea behind the FLUX project is to assess children's cognitive performance in daily life. This so-called "ambulatory assessment" approach aims at increasing ecological validity of measures by assessing children in their naturalistic contexts. Another aim was to administer cognitive tasks and to assess children's answers and reaction times. After establishing tasks with satisfactory psychometric properties for this purpose, daily fluctuations in cognitive resources will be described and their role for so-called *fluid* intelligence and school achievement will be investigated at the group and the individual level applying multilevel as well as time series analyses [19]. The Dell Streak 5 was selected as a smartphone for the FLUX project.

The assessment tests and items in FLUX are defined in QTI format. The items are simple and the interactions are mostly choice, text entry, slider, hotspot, and graphic order. Figure 3 shows sample items that have been designed in FLUX. Please note that all rights on the images are with the FLUX project.

Item 1 is a single choice item. The participant has to decide whether the images shown on the left and on the right side are identical. The answer is given by pressing one of the buttons in the lower part of the screen. Doing so, the next item is shown automatically. In case of no answer the next item is shown after the timeout specified in the item definition file. Items 2 and 3 are multiple choice items, which illustrate *working memory tasks* [20]. The items assess the ability to simultaneously maintain and process information.

Fig. 3. Sample Items

Item 2 shows images placed in the grid at the beginning of the test. Afterwards, movements are shown. The test-taker has to remember the updated positions and place the images in the grid accordingly. In Item 3, the test-taker hast to add or subtract numbers displayed in random order.

The findings of the FLUX study offer interesting insights into the extent to which using mobile techniques might be applied for supporting ambulatory assessment. It also highlights a number of major challenges that this format raises for the design of such resources.

4. Conclusion and Outlook

We have presented an integrated approach to enable and support mobile assessment in schools. Special emphasis was put on the usage of open standards and on test security. Our solution has been tested in several smaller and in one larger study where more than 120 students are tested three times a day over several weeks.

We do not only provide software but integrate it into a generic approach consisting of a hardware/software system with a central server, multiple classroom servers and the smartphones (or tablets) used for the assessment. We aim for unsupervised assessment; the phones can be handed over to the students for the assessment period without manual intervention necessary.

Our future plans are manifold. We want to extend the support for additional item types. This may go beyond the QTI standard, so we have to investigate additional standards usable for our purpose. Assessments are an integral part of learning. We want to extend the possibilities of the system to better support e-learning activities. Parts of our software development, especially for the test security, have been thoroughly tested only on one type of smartphone (the Dell Streak). We plan to extend the basis of supported smartphones, especially to more recent models. However, having in mind the innovation speed in the mobile sector, this is an on-going effort without end. Moreover, we plan to include additional hardware support into our system. Most smartphones offer sensors like GPS or a camera that might be used for assessment purposes as well.

5. Related Work

As we mentioned above, important parts of our system are based on QTI XML. We have implemented an Android Java QTI library. A similar work called Mobile QTI playr [21] has been done at the University of Southampton. This application is an adapted version of a QTI engine [22]. The QTI engine uses a web interface to browse through an assessment. It transforms QTI XML files via XSLT [23] in two steps into XHTML files. An XHTML renderer displays these output files in a browser. In the *Mobile* QTI playr, the XHTML renderer has been redesigned and re-implemented for mobile devices. To display XHTML files, it is necessary to install a mobile web server, i-jetty [25], on each Mobile device [21]. The QTI library that we implemented uses an XML serialization framework, which is Android compatible, to de-serialize QTI XML directly to Java objects and to serialize the objects back from Java to XML. The items of the assessment are implemented in Java using the Android SDK. So they can be run directly as an App on any Android device.

Our design decision was mainly based on test security considerations. This is easier to accomplish having full control over the assessment application. In addition it is desirable to have as few processes run on the smartphone as possible so that full access to all resources can be ensured for the main process.

References

[1] A. Holzinger, A. Nischelwitzer, M. Meisenberger (2005). In proceeding of: Pervasive Computing and Communications Workshops: Mobile phones as a challenge for m-learning: examples for mobile interactive learning objects (MILOs)

[2] IMS Global Learning Consortium, Inc. (2011). IMS Question & Test Interoperability Specification, http://www.imsglobal.org/question/.

[3] Respondus Inc. (2011). Question Import/Export Format: Respondus QTI Importer, http://moodle.org/mod/data/view.php?d=13&rid=986.

[4] Moodle Trust (2011). IMS QTI 2.0 format, http://docs.moodle.org/20/en/IMS_QTI_2.0_format.

[5] www.apple.com/de/ios/

[6] http://www.android.com/

[7] www.bada.com/

[8] www.microsoft.com/windowsmobile

[9] ca.blackberry.com/apps-software/blackberry7/

[10] Reto Meier (2010). Professional Android 2 Application Developement (pp. 14).

[11] Android Open Source Project (2011). Application Fundamentals, http://developer.android.com/guide/index.html.

[12] Android Open Source Project (2011). What is Android?, http://developer.android.com/guide/index.html.

[13] Sayed Y. Hashimi, Satya Komatineni, Dave MacLean (2010). Pro Android 2 (pp. 31-33).

[14] Simple XML Serialization Inc. (2011). Tutorial, http://simple.sourceforge.net/.

[15] Fabrizio Chami (2011), Android XML Binding with Simple Framework Tutorial, http://simple.sourceforge.net/articles.php.

[16] Reenskaug, Trygve (2011). MVC XEROX PARC 1978-79, http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html.

[17] Android Open Source Project (2012). Reference, http://developer.android.com/reference/packages.html.

[18] Android Open Source Project (2012). Android 4.0 release Document, http://developer.android.com/guide/topics/ui/ui-events.html.

[19] DIPF (2011). FLUX Project, http://www.idea-frankfurt.eu/homepage/idea-projects/projekt-flux.

[20] Akira Miyake, Priti Shah (1999). Models of Working Memory.

[21] Pei Zhang, Gary Wills, Leser Gilbert (2010). IMS QTI Engine on Android to support Mobile Learning and Assessment. http://eprints.soton.ac.uk/271485/.

[22] http://wiki.qtitools.org/wiki/QTIEngine

[23] XSL Transformations (2012). http://www.w3.org/TR/xslt.

[25] code.google.com/p/i-jetty/

An empirical study of user acceptance of lifelong learning on the move

Xiaoqing GU[a*], Shirong FU[b]
[ab]*Department of Educational Information Technology, East China Normal University, China*
*xqgu@ses.ecnu.edu.cn

Abstract: Mobile learning is the ongoing project that aims at fulfilling learning requirements in informal settings for Shanghai citizens. In order to understand whether users like using the mobile learning materials, an empirical study was conducted in a simulated field situation that was based on an integrated model. Fifty volunteers were recruited from the target population to participate in this study. The results showed that perceived quality had a significant influence on user satisfaction, as well as on individual intention to adopt. The findings of this study will help mobile learning content providers to develop learning materials based on an understanding of the perceptions of potential adopters.

Keywords: user acceptance, information system success, mobile learning system, user satisfaction, simulated field study

1. Introduction

In a rapidly changing world, mobile and pervasive technologies provide a new way for lifelong learners to access learning at any time and anywhere. Lifelong learners are typically engaged in a variety of learning activities for the practical purpose of addressing immediate problems, gaining understanding, or practicing a skill as the needs emerge (Fischer & Konomi, 2007). One striking feature of mobile learning is that learners typically perform different activities that might compete for attention simultaneously (Lin et al., 2007). With the understanding that there are always other distractions competing for limited time and attention resources in mobile situations, the exploration of user behavior could prove helpful to designers who want to make learning materials for mobile users that better serve the everyday user's needs. Therfore this study explores user acceptance and factors that influence mobile learning content delivery for daily usage. The findings in this study will be used as a basis for further diffusion of the mobile learning system in the lifelong learning network in Shanghai. This study will provide new knowledge about user acceptance of mobile technology in learning and show how to better serve the needs of the lifelong learner on the move.

2. Research framework

We start this investigation by linking an Information System (IS) success model (DeLone & McLean, 2003) with a Technology Acceptance Model (TAM) (Davis, 1989), both of which are broadly used in IT research and practice fields. By using an integrated IS success and user acceptance model, this study will attempt to identify the factors that affect user intentions to use mobile learning in their lifelong learning.

In our research context, we need to know whether the mobile learning system has potential value to users, so success should be measured from the perspective of users and focus on the model to explain intention to use/reuse. The success of mobile learning is focused on the quality and satisfaction association, while quality is interpreted as user perceptions of usefulness, usability, and task fit (Figure 1). Meanwhile external variables were added as precedent factors that affect the perceived quality of the mobile learning system. Perceived quality of usefulness, usability, and task fit are the major determinants of satisfaction that further affect the intention to use and reuse.

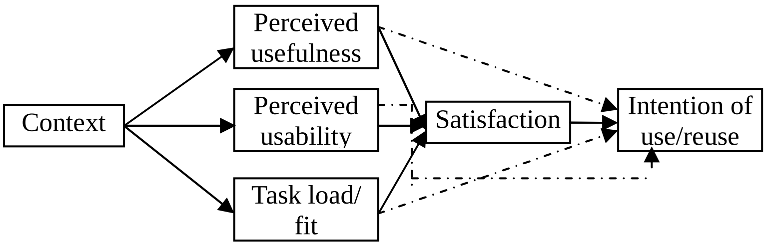


Figure 1 Research framework

External Quality

The success of the mobile learning system in contributing to individuals achieving their learning goals was the starting point of our research framework. Due to the difficulties of measuring the quality of a system, we propose to measure the quality using perceived usefulness, perceived usability, and task load as predictors of user satisfaction by integrating the measures from TAM and task-technology fit.

Lifelong learning on the move presumably takes place in various contexts, whenever the user can find available time. It is hypothesized that different usage contexts would have significantly different effects on the user's perception of usefulness, usability and task load for a learning product in use. Typical usage contexts are waiting and, on some occasions, walking. The contexts of waiting and walking are different, but we can project them onto three situations, according to previous studies: sitting, walking in a line, and freely walking along a path (Barnard et al. 2005; Lin, et al., 2007). The context of mobile learning was added to our framework as a form of 'facilitating condition'.

User satisfaction is the resulting attitude from the quality-value linkage of the IS success model (DeLone & McLean, 2003; Seddon, 1997; Rai et al. 2002), which is related to the intention to use in the updated IS success model (DeLone & McLean, 2003). The construct of satisfaction is a form of attitude toward behavior in the TAM, which is one of the predictors of behavioral intention (Davis et al., 1989). In our framework, user satisfaction was presumably directly affected by the quality of the mobile learning system, and this determined a user's adoption intention.

3. Methodology

The participants were widely recruited from the target lifelong learning population in Shanghai. An online questionnaire about attitudes towards obtaining information or resources with mobile phones, along with an invitation note to participate in our research was circulated on the popular social websites (such as renren.com), BBS, e-mail lists, and also in print form via leaflets in Shanghai. In the two weeks before the field study, 117 questionnaires were collected containing basic demographic information about the respondents, their attitudes towards obtaining resources with mobile phones, and a question about whether they would like to take part in the user experience study. The first 50 volunteers who ticked 'I'd like to' were invited to participate in the simulated field study.

The difficulties of obtaining user experience from real situations necessitated that we conducted a simulated field study, similar to Barnard et al (2005) and Lin et al. (2007), in which three mobility conditions were designed to correspond with the three typical mobile learning contexts chosen for this research: waiting (including sitting and standing), free walking (free of all obstacles), and walking through an obstacle course. Participants were asked to be seated in the waiting condition, while free walking involved walking along a path free of obstacles at a comfortable speed.

A testing protocol was designed for participants to follow in the experiment. Participants were asked to make a choice from a pool of available mobile learning resources, according to their own perception that the material could be useful then or in the near future. They were asked to enter one of the three different simulated contexts and situated themselves in any of them. Then to complete three tasks using any of the learning materials they selected, *i.e*., open, start learning, or quit on finish. After the tasks were finished, the participants were asked to finish a self-reporting survey.

The self-reporting survey was developed by adapting instruments from previous studies, as appropriate for this research context. Five measures were developed to assess the user intentions to use/reuse and the factors affecting this: *Perceived usefulness* (α=0.892), *Perceived usability* (α=0.914), *Task load/fit* (α=0.917), *Satisfaction* (α = 0.843) and *Intention to use/reuse* (α = 0.776). The 32 items for these five constructs were measured using a 7-point Likert scale labeled from ‘strongly agree = 7’ to ‘strongly disagree = 1’ (α = 0.956).

4. Results

A multiple regression analysis for the hypothesized model was performed to explore the relationships among the external factors, user perspectives of learning material quality, satisfaction, and intention of adoption. Altogether, the proposed model accounted for 52.8% of the variance in intention to use/reuse. Details have been elaborated in the following paragraphs of external factors and internal factors.

A MANOVA was used to further explore the influence of external variables on perceived quality of mobile learning products. With F as 2.757 (P =0.065), 7.976 (P = 0.000) and 5.119 (P=0.006), the results show that in different mobility contexts, users exhibited significant differences in their perception of usability and work load, but no difference were found on perception of usefulness.

Post hoc test comparisons showed that in terms of work load, there was a significantly better fit in the context of waiting, compared with free walking and walking through an obstacle course. There was no significant difference between the contexts of free walking and walking with obstacles. In terms of perceived usability, the results were significantly better in the context of waiting compared with walking through obstacles. There was no significant difference in perceived usability, either between waiting and free walking or between the two different walking contexts. Thus, the mobility context had significant effects on user perceptions of usability and task load.

The internal factors included the perceived usefulness, perceived usability, and workload fit were presumably affected by external factors, especially the mobility context, so this might have affected user satisfaction and thereby affect a user's intention of adoption. The Pearson correlation from the regression analysis showed that there were significant correlations between the internal factors of perceived usefulness, perceived usability, work load fit, and satisfaction. Regression analysis further supported casual effects on user satisfaction by perceived usefulness, perceived usability, and task load fit. With a coefficient of 0.7000 and a coefficient of determination of 0.487, these three factors can explain 48.7% of the variance in user satisfaction.

The correlation coefficients between the variables presumably determined the intention to use/reuse, which were also positively correlated. The casual effects were further supported in the stepwise regression analysis, where satisfaction explained 39.3% of the variance of user intentions of adoption, satisfaction and perceived usability explained 46.8% of the variance, while satisfaction, perceived usability, perceived usefulness explained 52.8% of the variance of user intentions of adoption.

The path coefficients in show that the standardized coefficients of perceived usefulness, task load fit, and perceived usability were 0.319 (t = 7.568, P = 0.000), 0.322 (t = 6.890, P =0 .000), and 0.208 (t = 5.191, P = 0.000), and of them were significantly correlated with user satisfaction. The standardized coefficients of satisfaction for perceived usability and perceived usefulness were 0.307 (t = 7.107, P = 0.000), 0.292 (t = 7.791, P = 0.000), and 0.304 (t = 7.569, P = 0.000), and all of them were significantly correlated with user intentions to use/reuse. Task fit had a significant effect on user satisfaction and stepwise regression on intention, but the effect of task fit on user intention to use was not significant.

5. Discussion and Conclusions

The results indicate that perceived usefulness had a significant influence on user satisfaction and individual intention of adoption. This means that the majority of users were concerned with the possible usefulness of the learning material. Another measure of perceived quality of a mobile learning system, perceived usability, was found to have a significant impact on user satisfaction and individual intention of usage. This means that most of the users tend to adopt the mobile learning if it results in a satisfactory experience. Thus, the content provider of lifelong learning on the move should pay special attention on the usefulness and usability of the learning materials. As expected, we found that task fit, the construct that was separated from perceived ease of use, had a significant impact on user satisfaction. Although contrary to our expectation that task fit would indirectly influence user intention of adoption, this finding still suggests that the majority of users were concerned about the time needed to finish learning on the move and the complexity of the learning task in a contextual situation full of distractions. Therefore, the content provider should strive to design the learning tasks to fit the complexity likely to occur in mobile learning situations.

Meanwhile, external factors had a significant impact on the user's perceived quality toward the mobile learning materials. We found that perceived usability and workload fit were affected by different mobility contexts. On the onther hand, differences in mobility contexts did not have a significant impact on perceived usefulness. This means that the potential usefulness of the mobile learning materials was context-independent, while perceived quality of usability, subjective attractiveness, and task load fit were related to the different contexts of usage. Thus, mobile learning providers should take the mobility context into consideration when they design learning material and suitable learning tasks for lifelong learners on the move.

Acknowledgements

Source of Funding: Research on mobile learning resource construction in the lifelong learning context, National Social Science Foundation (China), No. BCA080038. The authors would also like to thank Betty Collis for her constructive comments on an earlier version of the manuscript.

References

Barnard, L., Yi, J., Jacko, J. A. & Sears, A. (2005). An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human-Computer Studies*, 62(2005). 487–520.

Davis, F.D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*. 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*, *35*(8), 982-1003.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems, 19*(4), 9-30.

Fischer, G., & Konomi, S. (2007). Innovative socio-technical environments in support of distributed intelligence and lifelong learning. *Journal of Computer Assisted Learning, 23*(4), 338-350.

Lin, M., Goldman, R., Price, K., Sears, A., & Jacko, J. (2007). How do people tap when walking? An empirical investigation of nomadic data entry. *International Journal of Human-Computer Studies, 65*(9), 759-769.

Seddon, P. B. (1997). A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research, 8*(3), 240.

Yi, C., Liao, P., Huang, C., & Hwang, I. (2010). Acceptance of mobile learning: A respecification and validation of information system success. *International Journal of Behavioral, Cognitive, Education and Psychological Sciences, 2*(1), 55.

A Method for Determining Classroom Seating Arrangements by Using a Genetic Algorithm

Kazuhiro Shin-ike[a*], Hitoshi Iima[b]
[a]*Department of Electrical and Computer Engineering, MNCT, Japan*
[b]*Department of Information Science, Kyoto Institute of Technology, Japan*

Abstract: This study exploits a method for determining the classroom seating arrangements by using a genetic algorithm. In the genetic algorithm, we propose a procedure named a matrix crossover. Experiments are carried out based on the computational result. It is found from the experiments that satisfaction of each student for our proposed method is higher than the traditional classroom seating arrangement and each student is actually comfortable in the classroom because of the relations between other students.

Keywords: Genetic algorithm, A matrix crossover, Classroom seating arrangement, *Desk rows*, Coaching interview

Introduction

In Japan, classroom seating arrangements are adopted *desk rows* [1] that are placed students' desks like a matrix. In our earlier paper [2], we proposed a method to determine the optimal classroom seating arrangements by using a genetic algorithm. The optimization of two students, however, is not considered in the case where there is an aisle between them and so on. In this study we propose a method for determining the classroom seating arrangements considering relationships between a student and the other students sitting around the student. In order to determine the optimal classroom seating arrangements, a genetic algorithm [3] is applied on the basis of the questionnaire results and the analysis of the observation of behaviors between students.

Based on our proposed classroom seating arrangements, we carried out experiments. We compare our method with a traditional one that is determined by using students' ID numbers or by the homeroom teacher and students' intention. It is found that each student's satisfaction with our method is higher than that with the traditional one. Actually, most students answer in the questionnaire that they feel more comfortable in the classroom in the case of using our method. This proposed method can determine the classroom seating arrangements by a simple process in a short time.

1. A Method for Determining the Classroom Seating Arrangements

1.1 A genetic algorithm to determine classroom seating arrangements

The problem of determining the optimal classroom seating arrangements is a kind of combinatorial optimization problem, and it is well known that it is difficult to solve the problem [4]. Our optimization problem is to determine the combination of classroom seating arrangements in such a way that the minimum of fitness rates of all students is as large as possible. In order to determine the optimal classroom seating arrangements, a genetic algorithm is applied on the basis of the questionnaire results and the analysis of the observation of how the students behave in the classroom.

We have p students $\{i; i = 1, 2, \ldots, p\}$, where p is the number of students. The solution x for the problem is the classroom seating arrangements for the students. The evaluation value W_i, which student i sets on his or her present seat, is described by

$$W_i = \sum_{m=1}^{5} \beta_m a_{im} \tag{1}$$

where a_{im} is one of the results of questions which student i answered, β_m is a weight. The evaluation value W_{ij_k}, which shows behaviors between students i and j_k, is described by

$$W_{ij_k} = \max\{C_i, P_i, S_i, A_i\} + \max\{C_{j_k}, P_{j_k}, S_{j_k}, A_{j_k}\} \tag{2}$$

where C_i, P_i, S_i and A_i show four kinds of personalities student i has and C_{j_k}, P_{j_k}, S_{j_k} and A_{j_k} show four kinds of personalities student j_k has. The evaluation value W, which shows behaviors between student i and the other students sitting around him or her, is described by

$$\mathrm{W} = \sum_{k=1}^{n} \alpha_k W_{ij_k} \tag{3}$$

where n is the number of the students sitting around student i, and α_k is a weight. The objective function is defined by

$$\max \quad Z = \min \quad (W + \gamma W_i) \tag{4}$$

where γ is a weight.

The flow of a genetic algorithm for solving our optimization problem is as follows.

1. Generate randomly the initial population which consists of plural candidate solutions. Set g $\leftarrow$ 1.
2. Pick up two candidate solutions x_1 and x_2 randomly from the current population, and remove them from the current population.
3. Generate two new candidate solutions x_3 and x_4 from x_1 and x_2 according to a procedure called the matrix crossover, which is proposed by us. In this crossover, these new candidate solutions are generated in such a way that the classroom seating arrangements of x_1 and x_2 are partially inherited.
4. Generate a new candidate solution x_5 by changing a part of x_1. Similarly, generate another new candidate solution x_6 by changing a part of x_2. They are generated by a procedure called the mutation.
5. Select two best candidate solutions from the six candidate solutions $\{x_1, x_2 \ldots x_6\}$ and add them to the population.
6. If $g = G$, terminate this algorithm and output the best candidate solution as the answer. If not, set $g \leftarrow g + 1$ and return to Step 2. The parameter G is given in advance.

1.2 Encoding of chromosomes, crossover and mutation

Encoding of chromosomes is one of the problems in the genetic algorithm. Permutation encoding can be used because student ID numbers are used to encode chromosomes. In the permutation encoding, each chromosome is an m rows and n columns matrix where student ID numbers are used.

In this study, we propose a new crossover named a matrix crossover. Figure 1 shows parents P_1 and P_2. $P_1(p,q)$ is a submatrix of P_1. Figure 2 shows the first child (C_1) that is made by using a procedure shown as follows: First of all, a submatrix is randomly selected in the first parent (P_1) shown by the m by n matrix; Secondly, the chromosome of P_1 except the submatrix is copied to the first child (C_1) ; Finally, when the number which C_1 except

$C_1(p,q)$ does not have is found after P_2 is scanned from the first row and the first column through the m th row and the n th column, it is added to the elements in $C_1(p,q)$. This procedure is repeated until the elements of $C_1(p,q)$ is filled with figures. The second child (C_2) is also made by the same procedure as C_1.

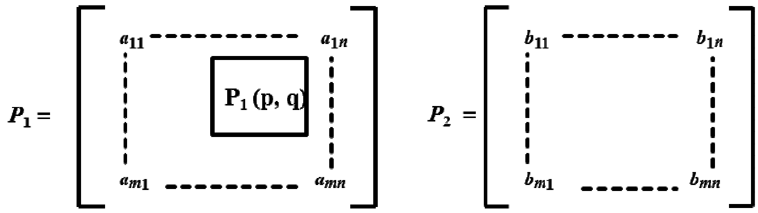


Fig. 1 Parent P_1 and P_2

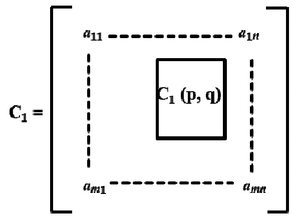


Fig. 2 Child C_1 obtained by the matrix crossover

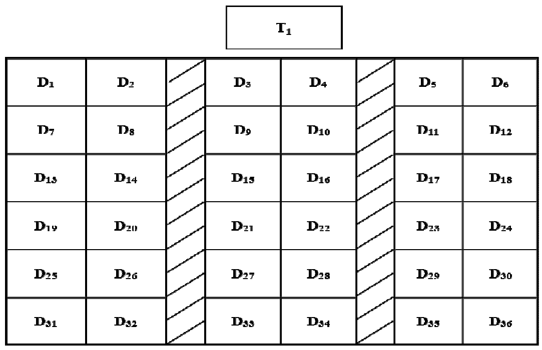


Fig. 3 Desk rows in the classroom

I D	FREQ	HEAR	VIEW	INTE	SEAT
i	a_{i1}	a_{i2}	a_{i3}	a_{i4}	a_{i5}
j_k	$a_{j_{k1}}$	$a_{j_{k2}}$	$a_{j_{k3}}$	$a_{j_{k4}}$	$a_{j_{k5}}$

ID: Student ID number
FREQ: Frequency of conversation
HEAR: Hearing range
VIEW: Viewing distance
INTE: Interest in the class
SEAT: First-choice seat location

Table 1 Fitness values of students i and j for determining classroom seating arrangements

Mutation plays an important part of the genetic search to prevent the population from stagnating at any local optima. In this study, we adopt order changing as a mutation in permutation encoding. Two numbers belonging to one chromosome are selected randomly and exchanged. This order changing is performed in the other chromosomes.

2. Simulations and Experiments

2.1 Problem statement for determining classroom seating arrangements

In higher education, each student takes his or her seat thinking about other students who sit around his or her seat [5]. Figure 3 shows *desk rows*, that is a traditional classroom seating arrangement. In this figure, D_1, D_2, …, D_{35} and D_{36} show desk numbers and slash marks show aisles in the classroom. T_1 shows a teacher's desk and a blackboard is behind the desk. An optimization of the classroom seating arrangements is carried out between a student and the other students who sit around the student.

Table 1 shows fitness values that students i and j_k set on their own seats. In this table, a_{i_1} shows an index as to whether student i often has a non-productive talk. a_{i_2} and a_{i_3} show indexes as to whether the student i can hear the teacher's voice and as to whether student i can see the letters which the teacher writes on the blackboard, respectively. a_{i_4} shows an index of the student i's interest in the class and a_{i_5} shows an index as to which seat student i wants to sit in. $a_{j_{k_1}}$, $a_{j_{k_2}}$, $a_{j_{k_3}}$, $a_{j_{k_4}}$ and $a_{j_{k_5}}$ show the indexes of student j_k.. Each fitness value is translated on a scale of zero to one.

Table 2 Weight values between personalities of students i and j_k obtained from coaching interview

α_k	C_i	P_i	S_i	A_i
C_{j_k}	0.10	0.50	0.75	0.25
P_{j_k}	0.50	0.50	0.75	0.25
S_{j_k}	0.75	0.75	1.00	0.75
A_{j_k}	0.25	0.25	1.00	1.00

C: Controller type
P: Promoter type
S: Supporter type
A: Analyzer type

Table 3 The computational result for the classroom seating arrangements

FG	10,000	100,000
PT	116 (s)	1002 (s)
Z	2.40	2.63
γ	1.0	1.0o

FG: Final generation
PT : Processing time
Z : Values of objective function

Table 2 shows a weight α_k found in equation (3) that is employed to express a chemistry between students i and j_k. In this equation, α_k is a weight of whether they get on well together or not. α_k is determined based on personalities of students i and j_k obtained from coaching interview. Personalities are divided into four types that are a controller type, a promoter type, a supporter type and an analyzer type. A maximum value of C_i, P_i, S_i and A_i determines a personality of student i and a maximum value of Cj_k., Pj_k, Sj_k. and Aj_k determines a personality of student j_k., respectively.

2.2 Computational result

In order to simulate the classroom seating arrangement, thirty six students are selected. First of all, we carry out a questionnaire to investigate the students' feeling on seating arrangements. Secondly, on the basis of the questionnaire, the classroom seating arrangements are calculated by using the genetic algorithm. In our genetic algorithm, there are three parameters: γ shown in equation (4), the population size (*PS*) and the final generation (*FG*). The genetic algorithm is performed ten times with various random seeds.

Table 3 shows theses parameters, processing time and values of objective function obtained from the calculation results. It is confirmed from this table that the genetic algorithm which *FG* is one hundred thousand is better than the other, because the value of objective function is higher than the one in the case where *FG* is ten thousand.

2.3 Experiment based on the computational result

On the basis of computational results, experiments are carried out. First of all, the students take classes sitting the seats which are decided by using only students' IDs shown in Fig. 3. Secondly, students take classes sitting the seats obtained from the computational result. Finally, we compare the results from the first and the second.

Figure 4 shows the classroom seating arrangement obtained from the genetic algorithm. In this figure, each decimal number shows each student's ID number. Students sit in the assigned seats and take a class for approximately one hundred minutes. After the class, they fill in a questionnaire related to the classroom seating arrangement.

Table 4 shows the result of the questionnaire for the classroom seating arrangement in the case where *FG* is one hundred thousand. In this table, the first column shows numbers of the questions and, in the first row, A, B and C show positive alternatives and D and E show negative alternatives of each question. It is found from the first question that 54.0% of the students have some students around them, whom they can talk about their questions. In the second question, 73.0 % of the students are able to sit in and near the seat which they want to. In the third question, since 92.0 % of the students choose the answer A, B or C, it is thought that they are able to hear a teacher's voice more or less. In the fourth question, 88.0 % of students reply that they are able to see the letters which the teacher writes on the blackboard. In the fifth question, 81.0 % of students reply that they prefer their seats obtained from the computational result to their seats which are determined in numerical order?

T_1

7	30		29	3		23	27
10	20		11	16		17	8
22	5		2	28		9	1
25	34		31	6		35	15
21	26		12	36		32	18
33	24		13	4		19	14

Fig. 4 The classroom seating arrangement based on the proposed genetic algorithm (FG = 100,000)

Table 4 The result of the questionnaire for the classroom seating arrangement (FG = 100,000)

Item	A (%)	B (%)	C (%)	D (%)	E (%)
1	54.0	46.0	0	0	0
2	11.0	62.0	27.0	0	0
3	34.0	27.0	31.0	8.0	0
4	34.0	19.0	35.0	8.0	4.0
5	19.0	27.0	35.0	11.0	8.0

3. Conclusion

In this paper, we propose the classroom seating arrangements to satisfy the requests of students. A genetic algorithm is applied to find the best classroom seating arrangement. Based on the results of the questionnaire, the classroom seating arrangements are calculated by using the genetic algorithm and we carried out experiments at a national college of technology in Japan. Students take classes based on the classroom seating arrangements obtained from the genetic algorithm.

We compare our proposed classroom seating arrangement with the traditional one that is determined by using only students' IDs or by the students and the homeroom teacher's intentions. From the result of the questionnaire filled in after the class, it is found that satisfaction of each student is higher in our method than the traditional one. In addition, this method can determine the classroom seating arrangements by a simple process in a short time. It is confirmed that the proposed method is effective.

References

[1] Buckalew, L. W., Daly, J. D., Coffield, K. E. (1986). Relationship of initial class attendance and seating location to academic performance in psychology classes, *Bulleting of the Psychonomic Society*, Vol. 24(1), pp.63-64.

[2] Kazuhiro Shinike, Hitoshi Iima. (2011). A Method for Determining Classroom Seating Arrangements by Using a Genetic Algorithm. *SICE Annual Conference* 2011, pp. 161-166. Tokyo, Japan, Septenber 13-18. .

[3] D. E. Goldberg. (1989). Genetic Algorithm in Search, *Optimization and Machine Learning*, Addison - Wesley.

[4] H. H. Hous and T. Stützle. (2005). "Stochastic Local Search: Foundations and Applications", *Morgan Kaufmann*.

[5[Adams and D. Zuckerman, "The effect conditions on personal space requirements", T*he journal of general psychology*, Vol. 115-4, pp. 335-340, 1991.

Sustaining Student Retention of Computer Programming with mLearning Apps

Dr. EM Nalaka EDIRISINGHE
Lecturer, Temasek Polytechnic, Singapore
nalaka@tp.edu.sg

Abstract: The learning of computer programming fundamentals remains a challenge for novice learners across the world. Many different approaches have been attempted from more engaging languages to media computation. One area that has not been addressed yet is how to help students sustain their retention of this knowledge and apply it successfully in their current and future modules and careers. In addition, with the proliferation of smart-phones today's learners are demanding more access to immediate information to help them in their tasks. In this paper, we describe an approach leveraging upon mLearning strategies to help students sustain and apply their knowledge of fundamental programming concepts in the foundation modules. This paper shares the processes and lessons learned from a pilot project of developing customized, native mLearning applications on multiple platforms to address this gap. The paper describes the strategy and experiences of developing the mLearning apps, which included a multiple programming language reference and customized self-tests, and adopting them in the teaching of programming subjects. It discusses the students' and faculty's experiences and their implications.

Keywords: Mobile learning, technology, education, computer programming.

Introduction

Learning fundamental programming for novice learners during their first year of studies at higher education institutions remains a challenge. The attrition rate in such subjects and programs is quite high. Many efforts over the years have been made at tackling this problem through the use of a variety of tools and techniques with different levels of success. However, an area not often studied relates to how students successfully retain this initial burst of knowledge after their first fundamental programming module. The retention of this knowledge throughout their subsequent years of study remains a critical challenge. It is common to find students unable to recollect fundamental concepts learnt in their foundation years, particularly in the final year projects and subsequent careers. At the same time, today's higher education landscape has seen dramatic shifts in terms of teaching and learning, including the easy access to information, the increasing availability of mobile devices amongst learners, and the increasing demand for anywhere, anytime learning. As such, mobile learning and applications provide one platform upon which to tackle this issue of sustaining student retention of fundamental programming concepts. Thus, this paper intends to explore, on a pilot basis, the adoption of mLearning apps during the first year of studies to sustain retention of programming knowledge.

1. Literature Review

1.1 Learning Programming

Students learning programming for the first time face multiple challenges. Resnick (2009) indicated that novice learners faced challenges, such as mastering the syntax of the language and understanding the complexities of programming environments in which they must learn to write programming solutions. Other challenges include the lack of motivation and the inherent difficulty faculty face in engaging student through text-based programming languages (Kelleher, C. & Pausch, R., 2005). Many solutions have been adopted in the past through storytelling and tools, such as Alice, game-based learning and media computation (Kelleher, C. & Pausch, R. 2007). However, there has been inadequate work done into exploring how to facilitate student retention of their knowledge of programming using a specific mobile platform and strategy.

1.2 Mobile Learning

The proliferation of mobile technologies into our everyday lives continues at a rapid pace. Now more than ever learners at a younger age are using smart-phones that provide them with access to a world beyond the confines of their classroom. A recent study of students in higher education in China identified that 85.7% of the pool of mobile users actually access universities via their smart-phones (CNNIC 2010). As such, Rajasingham (2011) highlights that we, as educators, need to take into account the changing landscape and design for the needs of the new mobile savvy audience. One way the educational community has been meeting the needs of such an audience is through mobile learning. Mobile learning is defined as any form of learning that is conducted while the learner is on the go (Sharples et la, 2005). Conole (2008) stated that mobile technologies continue to be used in developing innovative pedagogical scenarios for learners. In the context of learning fundamental programming, different studies have been done to motivate students by allowing them to build mobile applications (Mahmoud, 2008, Spertus et al, 2010). Further, studies have shown it is useful to incorporate mobile learning as a form of blended learning (Shen, 2008). However, insufficient studies have been done on actually deploying customized mobile learning applications in a blended setting to support the student retention of fundamental programming concepts.

2. Background

Temasek Polytechnic is a tertiary institution with over 21 years of experience in conducting information technology related diplomas. All students who enroll in these diplomas every year take up a fundamental programming module. The objective of this module is to teach novice learners programming fundamentals. After this introductory programming module, students continue to take more advanced modules in the areas of object-oriented programming, algorithms and web development, all of which require a strong retention of the fundamental programming knowledge.

3. Subject

3.1 Rationale

Historically teaching the fundamental programming module at our institution has faced many challenges in terms of student motivation and engagement. In the past this gap had been addressed through the use of different tools such as Alice and flow-charting. However, in recent years faculty members teaching advanced modules had highlighted that student retention of their fundamental programming concepts, such as writing iterative constructs or declaring arrays, was very poor. Faculty continued to face students who had succeeded in their early programming modules but could not sustain their mastery of the knowledge in advanced modules. As such, the teaching team decided to develop native mobile applications on the iOS and Android platforms to help students sustain their retention of fundamental programming constructs.

3.2 Mobile Strategy

In developing the native mobile applications the teaching team adopted a five principle mobile strategy. First, the mobile applications would be focused on providing students with "just-in-time" instruction and access to information. The rationale for selection of the "just-in-time" approach was that the learners would not only be using these applications in their first year of studies but during their subsequent years and careers. Second, the mobile applications would not be a substitute for existing courseware and e-learning content available on the institution's learning management system. Third, the mobile applications would follow an outcomes-based approach to learning. As such, the mobile applications would not simply be a repository of course content or PowerPoint notes. Instead, the mobile applications would give information to achieve certain learning outcomes. This was done in alignment to previous research that indicated that effective learning requires a student-centered, outcome-focused approach (Field, 2005). Fourth, the mobile applications would provide a simple user interface with minimal clicks to access the relevant content. Fifth, each mobile application included differentiated instructional self-tests, which were differentiated by content. Based on these five principles, two mobile applications were developed for the iPhone and Android markets. The mobile applications consisted of two modules: a reference module in multiple programming languages and a self-test module as shown in Figure 1.

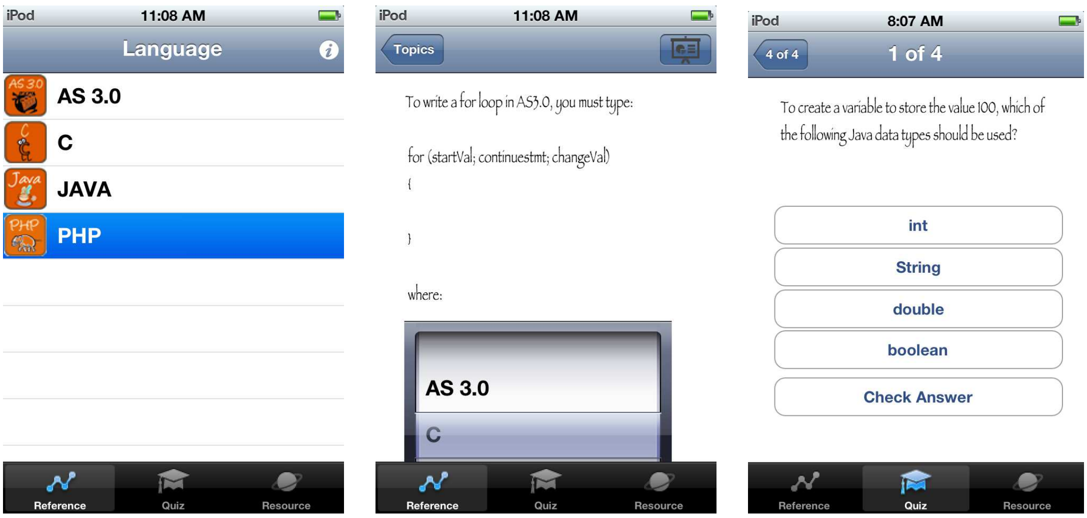


Figure 1: Example Screenshots of the iPhone mLearning Apps

The reference module listed key outcomes students needed to remember, such as declaring variables, writing loops, declaring arrays, identifying even and odd numbers and searching an array. To help students the reference module provided necessary key content for multiple programming languages, such as Java, C, PHP and ActionScript, that students would encounter in their advanced programming modules. The self-test module was customized to help learners with different abilities remember the key content. Simple learning analytics was integrated as part of each self-test. The self-test module provided differentiated content for two levels of users. The differentiated content was determined through a pre-quiz assessment. Based on the scores achieved in the pre-quiz learners with different abilities would be sent to one of two possible paths. Each path would be customized with the necessary scaffolding and feedback to assist learners to master the content. As such the multiple paths provided learners with a degree of customization of their learning content.

4. Methodology

Since the project was a pilot study the methodology used was qualitative in nature. This pilot focused on the faculty and student populations. The faculty population consisted of those teaching advanced modules and the student population consisted of those taking advanced programming-related modules. The sample size consisted of two (2) faculty members and twenty (20) students. The faculty members and students were selected using simple random sampling. The instruments used in the study included the faculty members' journals, interviews with the students and interviews with the faculty members. The mobile applications were deployed over a period of 3 months and all journal data and interviews were collected and conducted at the end of this phase. The data analysis consisted of content analysis of the textual data and went through 4 distinct stages. In the first stage, the textual data was separated between the student and the faculty members. In the second stage, a partial set of the textual data was grouped together to identify common themes amongst each of the two groups in the sample population. In the third stage, coding was performed to derive inductive categories. In the fourth stage, the inductive categories were ranked based on the number of occurrences falling into each category from all the interview and journal data. As such, these inductive categories formed the results described in the next section.

5. Results

The results of this study can be categorized into two areas: the student perspective and the faculty perspectives.

5.1 Student Perspective

In terms of student perspectives, five key results were noted. First, the students found the mobile applications very helpful in reminding them of key content that they had just learnt when they needed it (i.e. just-in-time). This result was in alignment with previous studies that showed that mobile learning was a step towards making the educational process "just-in-time, just-enough, and just for me" (Peters, 2007). Second, the students liked the convenience with which to find the content without the need to research on-line or lookup their old lecture notes or textbooks. This implied that the mobile learning applications provided students with a time-saving element, thereby increasing their efficiency and productivity, which was aligned to previous studies related to offering students the flexibility to determine the conditions of their own learning (Field, 2005). Third, the students found the quantity of the content just right for their needs. This implied that the amount of content provided for each topic was a potential factor in students using their mobile learning applications. Fourth, the students found the multiple programming languages very useful in helping them navigate through new programming languages in the future. This implied that students may potentially continue using these apps beyond their first year as they provided a wide range of programming languages that may be relevant in their subsequent years. Fifth, the students found the customized self-tests helped aid them in managing the pace of their own learning. This implied that learners do view personalized learning experiences as a potential factor to their buy-in of mobile learning applications.

5.2 Faculty Perspectives

In terms of faculty perspectives, three key results were noted. First, the teaching staff highlighted a very high adoption rate by students. It was observed that students were quick to use the mobile applications even inside the classroom when they needed to refer to something. Second, the teaching staff indicated that students were very comfortable with the mobile applications as it was considered just another app in their smart-phone. Third, the teaching staff indicated that the outcomes-based approach was very suitable for mobile learning applications as it helped promote the process of "just-in-time" learning. This implied that the development of content needed to be carefully done without simply transferring existing courseware or e-learning content onto a mobile learning application.

6. Conclusion

In conclusion, this pilot study potentially suggests that mobile applications for learning based on a specific strategy customized to the learners' needs can help in student retention of fundamental programming knowledge at the initial stage. However, further studies need to be done to measure student adoption over subsequent years and their effectiveness.

References

[1] Conole, G., de Laat, M. Dillon, T. & Darby, J. (2008). Disruptive technologies, pedagogical innovations : What's new ? Findings from an in-depth study of students' use and perception of technology. *Computers & Education*, 50, 511-523.

[2] CNNIC (2010). 26th Statistics Report for Internet Development in China.

[3] Field R. (2005). Favourable conditions for effective and efficient learning in a blended face-to-face / online method. Proceedings of ASCILITE 2005. Retrieved January 12, 2006, from http://www.ascilite.org.au/conferences/brisbane05/proceedings.shtml.

[4] Kelleher, C. & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computer Surveys*, 37(2), 83-137.

[5] Kelleher, C. & Pausch, R. (2007). Using storytelling to motivate programming. *Communications of the ACM*, 50(7), 64.

[6] Mahmoud, Q. (2008). Integrating mobile devices into the computer science curriculum. Paper presented at the Frontiers in Education Conference, 2008. FIE 2008. 38th Annual.

[7] Peters, K. (2007). m-Learning: Positioning educators for a mobile, connected future. *International Journal Of Research in Open and Distance Learning*, 8(2), 1-17.

[8] Rajasingham, L. (2011). Will Mobile Learning Bring a Paradigm Shift in Higher Education. *Education Research International*, Volume 2011.

[9] Resnick, M., Maloney, J. Monroy-Hernandez, A. Rusk, N. Eastmond, E. Brennan K., et al. (2009). Scratch: programming for all. *Communications of the ACM*, 52(11), 60-67.

[10] Sharples, M. Taylor, J. & Vavoula, G. (200). Towards a Theory of Mobile Learning. Paper presented at the mLearn 2005, Cape Town, South Africa.

[11] Shen, R., Wang, M. & Pan. X (2008). Increasing interactivity in blended classrooms through a cutting-edge mobile learning system. *British Journal of Educational Technology*, 39(6), 1073-1086.

[12] Spertus, E., Chang, M.L., Gestwicki, P. & Wolber, D. (2010). Novel approaches to CS 0 with app inventor for android. Paper presented at the Proceedings of the 41st ACM technical symposium on Computer Science Education.

How to increase ubiquitous experiential learning

Mengmeng Li, Hiroaki Ogata, Bin Hou, Noriko Uosaki
The University of Tokushima, Japan

Abstract: This paper introduces a mechanism in a ubiquitous learning log system, which is designed to guide learners to participate in the learning activities recommended by the system. In order to provoke learners' interests on the knowledge, the recommended knowledge is related to both what learners are studying and the learners' current learning contexts. And to help learners grasp the knowledge, in the second step the mechanism provides learners with relevant learning activities to guide learners to use the recommended knowledge.

Keywords: Ubiquitous learning log, location based recommendation, learning experience, language learning

Introduction

The fact that a lot of note taking software such as evernote is popular with smartphone users proves that taking notes with digital devices is a becoming common way [1]. And besides the informal notes (like reminder, to-do list) [2], [3], knowledge or learning experience can also be tracked as formal notes, such as English vocabulary [4]. This kind of formal notes is called ubiquitous learning log in this study. In order to make use of learners' learning logs, we proposed two directions to deal with them. On one hand the ubiquitous learning log can be used to help learners to review and recall what they have learned; on the other hand it can be treated as learning source to be recommended to other learners. We have developed a system called SCROLL (System for Capturing and Reminding of Learning Log) to implement this proposal. But the way to recommend specified learner with other learners' ubiquitous learning log does not perform as well as we expected [5]. Therefore, in this paper we propose a new way to improve the recommendation.

Another goal of this paper is to introduce a way to induce learner to use the recommended knowledge. We do this for two reasons. Firstly, according to the learning by experience theory, learning through experience can reinforce learners' efficiency of mastering the obtained knowledge [6]. Secondly, the recommended knowledge should be linked from the digital world with the real world. And it may motivate learners' interests compared with only presenting knowledge in the digital devices [7]. Consequently, a mechanism about how to induce learners to participate in activities in real world using the recommended knowledge is also described in this paper.

The rest of this paper is organized as follows. In section 1, we describe the previous work and give a brief introduction about SCROLL. Section 2 presents a new way of recommending ubiquitous learning log based on learning context. Then, the mechanism to induce learner to use the recommended knowledge is described. Finally, we give the conclusion and the future work of this study.

1. Previous work

The ubiquitous learning log is defined as a recorded form of knowledge or learning experience acquired in our daily lives and it serves as memory storage for notable or important knowledge to review, to remind and to reflect. SCOLL is a system to help learners to record and recall their ubiquitous learning log. Until now, it mainly focuses on the language learning field, especially on assisting international students to study Japanese in Japan. Therefore, in this case the ubiquitous learning log represents the Japanese knowledge gained by the oversea students in their daily lives in different kinds of situations, such as shopping in the market, seeing a doctor in the hospital, having a haircut in a barbershop, visiting the museum and so on.

In order to recommend specified learners with other learners' learning logs, the paper [5] also presented one way of recommendation, which can notify a learner with other learners' learning logs near him. And if he selects his interested learning logs, the system will show him a path from the current location to the selected learning logs on the map. Even though the recommendation has considered the factors such as the languages that the learner studies, his mother language and his proficiency level of the learning language, the learners' response rate of such recommendation is quite low (only 25% [5]). We find two reasons, which can account for the fact. One is that such recommendation way is not capable of arousing learners' interests or curiosity to study other learners' learning logs and even to find out them via the help of the navigator function. The other one is that only the knowledge recommendation does not motivate learners to explore more about the knowledge and does not help learners to memorize or master the knowledge. Therefore, in order to improve the meaningful recommendation rate for learners and to induce learners to use the knowledge, we propose a new way, which includes two parts: providing the learners with the learning logs relevant to the learning activity at hand and inducing the learners to use the recommended knowledge. The following two sections introduce the two parts separately.

2. Related learning log recommendation

To improve the quality of the recommendation, the system will recommend the knowledge more related with what the learners are studying. It means that the recommendation occurs after the learner saved his learning log and then the system will search the other learners' learning logs related to the saved one in two directions: context-related and concept-related. The following sections will introduce the two directions in detail.

2.1 Similar context based learning logs

In the study [5], we introduced a kind of learning log called location-based learning log. This kind of learning log is regarded as the knowledge that can be recalled by the location as a retrieval cue. According to the theory of encoding specificity, the place where we learned can be encoded as a retrieval cue initially and it is effective to activate a stored memory [5]. Therefore, when a learner (called learner A) comes near to the place where he learned something, the system can remind him of his old knowledge. What's more, if another learner (learner B), who has the same learning needs with learner A, also comes to the place where learner A learned, the system will recommend learner A's knowledge for learner B. However, because a learning log is only combined with only one place, the probability of such recommendation is relatively low. Therefore, we propose another concept: similar context based learning log. It means that a learning log is not only related to the place where the learner learned it, but also related to the similar context. For example, if we learned the

Japanese names of vegetables in a supermarket, it is also meaningful to remind us of the knowledge in another supermarket. Consequently, when a learning log is created by a learner, he is recommended to designate what kinds of contexts the learning logs are related to. The learners are motivated to do this because the system can also help them to recall what they have learned based on the similar contexts. And the automatic context detection function, which is implemented by making use of learners' GPS data and Google Places API (Application Programming Interface), makes it convenient for learners to choose the context of the learning log. In addition, the contexts have been predefined including school, hospital, super-market and so on.

As a result, the new way of recommendation will focus on both the contexts that learning logs are belonged and the learners' current learning context. For example, if a learner learned "豆腐 (tofu)" in the supermarket context, the system will provide the learner with the other learning logs in the supermarket context.

2.2 Related learning logs

After searched the learning logs in similar context, the system will focus on the learning logs which are most relevant to. To judge how related two learning logs are, the system utilizes the tag of the learning log to calculate the two learning logs' distance. For example, in Figure 1 tofu and celery share a same tag: cooking, so the distance between tofu and celery is 1. Similarly, the distance between celery and tomato is also 1 and the distance between tofu and tomato is 2. In this study, only the learning logs, whose distances from the original learning log are less than 3, are recommended for the learner.

Figure 1. Distances among learning logs

3. Induce learner to use the knowledge

According to the experiential learning theory, knowledge can be created from the combination of grasping and transforming experience [8]. Thereby, to help learners to grasp and master other learners' learning logs, the system is designed to induce learners to use the recommended knowledge. And to induce the learners, a concept called task is proposed. What are tasks, who provides the tasks and how the learners participated in the tasks are discussed in the next two sections.

3.1 What are tasks and who provides tasks?

Tasks are referred as the activities that the knowledge can be used. They are related to the learning contexts talked in section III like school, hospital, post office and so on. For instance, if the system recommends a learner a Japanese word "トマト(tomato)" in a supermarket, the learner can talk with the staffs in the supermarket using the word "トマト (tomato)", such as asking its price, location, recipe and so on. And it has been proved that by talking with the Japanese native speaker using the recommended word, learner can master the word well [9]. The activity of asking about the information is a kind of the so-called task.

Basically, the learners who saved the learning log are responsible for providing what kinds of tasks the knowledge can be utilized. And one learning log can be used in several tasks. Moreover, the system provides some predefined tasks in different contexts in order to reduce the learners' burden of designating tasks when they save their learning logs. Table 1 shows part of the predefined tasks in different contexts. What's more, the tasks can be defined by the learner and designated by the administrator of the system.

Table 1. The predefined tasks and their contexts

Context	Predefined Tasks
Post office	Ask about the postage Buy a piece of stamp or an envelope
Restaurant	Ask about the recommended menu
Bank	Ask about the interest rate, currencies deposit money
Hotel	Ask about the check in time Ask about the campaign
Airport	Ask about the flight
Station	Ask about the route, the price Ask about how to buy a ticket

3.2 How learners participated in tasks?

Because one learning log has several tasks, the system assigns the appropriate task for a learner based on the difficulty of the task and the learner's ability. For example, asking the price of the production is easy for learners to finish while asking about the recipe of the vegetables is quite difficult for most learners. And when the learners received the recommended the learning log and the task, they are also asked to provide feedback for the system. For example, they are asked to take the photos of the object if they are asked to inquire the location of it. And if they are asked to learn about the place of the production, they need to accomplish this information on the system. Only providing the feedback can prove that they have really used the knowledge. And if the learner meets new problems when he carries out the tasks, he can record them in photos, videos, audios or texts and upload them to the system in order to ask for help. Such accumulated data is also meaningful for the other learners.

4. Conclusion and future work

This paper describes a mechanism of recommending learners with ubiquitous learning logs which are relevant to learners' learning contexts. It considers both the context-based correlation and concept-based correlation. Moreover, the mechanism is also designed to induce learners to use the recommended knowledge in the real world. The tasks related to the contexts are offered by the system. After finished the tasks, the learners are asked to provide the feedback and questions. To motivate learners participate in the activities are help them to master the knowledge is the final goal of this mechanism. As for the future work, we will implement the system and conduct an experiment to evaluate this mechanism. Both the usability of the system and the learners' response will be evaluated.

Acknowledgements

This research work was supported by JST PRESTO, and the Grant-in-Aid for Scientific Research No. 21650225 from the Ministry of Education, Science, Sports, and Culture in Japan.

References

[1] M. Trautschold and G. Mazo, "Eliminate Your Paper Notes," *iPad 2 Made Simple*, pp. 583–600, 2011.
[2] L. Dai, W. G. Lutters, and C. Bower, "Why use memo for all?: restructuring mobile applications to support informal note taking," in *CHI'05 extended abstracts on Human factors in computing systems*, 2005, pp. 1320–1323.
[3] M. Lin, W. G. Lutters, and T. S. Kim, "Understanding the micronote lifecycle: improving mobile support for informal note taking," *Proceedings of the SIGCHI conference on Human factors in computing systems 2004*, vol. 6, pp. 687–694, 2004.
[4] H. Ogata, M. Li, H. Bin, N. Uosaki, M. E.-B. Moushir, and Y. Yano, "SCROLL: Supporting to Share and Reuse Ubiquitous Learning Log in the Context of Language Learning," *Proc. of mLearn 2010*, pp. 69–82, Oct. 2010.
[5] M. Li, H. Ogata, B. Hou, N. Uosaki, and Y. Yano, "Personalization in Context-aware Ubiquitous Learning-Log System," in *2012 IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education*, pp. 41–48.
[6] H. Janta, P. Lugosi, L. Brown, and A. Ladkin, "Migrant networks, language learning and tourism employment," *Tourism Management*, vol. 33, no. 2, pp. 431–439, 2012.
[7] G. J. Hwang, C. C. Tsai, and S. J. . Yang, "Criteria, strategies and research issues of context-aware ubiquitous learning," *Learning*, vol. 11, no. 2, pp. 81–91.
[8] D. A. Kolb and others, *Experiential learning: Experience as the source of learning and development*, vol. 1. Prentice-Hall Englewood Cliffs, NJ, 1984.
[9] H. Ogata, G. L. Hui, C. Yin, T. Ueda, Y. Oishi, and Y. Yano, "LOCH: supporting mobile language learning outside classrooms," *International Journal of Mobile Learning and Organisation*, vol. 2, no. 3, pp. 271–282, 2008.

Evaluation of Mobile Learning: A Cognitive Style Approach

Li-Ping CHANG, Pei-Ren HUANG and Sherry Y. CHEN*
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*sherry@cl.ncu.edu.tw

Abstract: Mobile devices bring a lot benefits to student learning, including flexibility, convenience and ubiquity. On the other hand, students have various characteristics, among which cognitive styles play an important role. This study aims to investigate how students react differently to mobile device from a cognitive style perspective. The results suggest that students in the mobile device scenario had more engagement and performed better than those in the desktop computer scenario. Furthermore, Holists who made more movements and Serialists who frequently used the Keyword Search could achieve good performance in the desktop computer scenario. On the other hand, the students performed similarly in the mobile device scenario though Serialists who made more repeated visits and browsed more pages could have better performance.

Keywords: Cognitive styles, mobile device, technology-based learning tool

1. Introduction

The advancement of wireless communication technologies has recently provided an opportunity for educators to use new technology-based learning tools. Among various technology-based learning tools, mobile devices (MDs) are being widely applied to support student learning (Hein & Irvine, 1999). This is due to the fact that MDs offer ubiquitous information access (Zhang, 2007). More specifically, the MDs are portable so geographical access barriers can be overcome (Gulati, 2008). Thus, educational practice can be performed any places with MDs (Cavus, 2011).On the other hand, the screen size of a MD is small (Kukulska-Hulme, 2007). Accordingly, the MDs may not suit to everone. In particular, diversities exist among students, in terms of their knowledge, skills, and needs (Chen & Macredie, 2004). Therefore, there is a need to examine relationships between individual differences and the use of MDs.

Among a variety of individual differences, previous research found that cognitive styles greatly affect student learning (Chen & Liu, 2011) because it refers to a person's information processing habits, capturing an individual's preferred mode of perceiving, thinking, remembering, and problem solving (Messick, 1976). In this vein, the study reported in this paper aims to examine students' different reactions to the DCs and the MDs from a cognitive style perspective. To this end, two research questions are investigated: (a) how students react differently to the DCs and the MDs; (b) how cognitive styles affect their reactions to the DCs and the MDs. In order to obtain a complete understanding, both learning behavior and learning performance are applied to find answers for the abovementioned two research questions. Answers to these two questions are sought by using a data mining approach to analyze students' learning patterns because data mining has been successfully applied to examine students' learning behavior (Chen & Liu, 2008).

2. Methodology Design

44 students from a university took part in our empirical study. 21 students were assigned to a DC scenario, in which students interacted with the WBL system via a DC. On the other hand, 23 students were allocated to a MD scenario, in which students interacted with a Web-based learning (WBL) system via a MD, i.e., an iPad. Regardless of the DC scenario or MD scenario, the WBL system gave the lecture of "Interaction Design" and included eight sections. The system provided two kinds of navigation tools. One is Keyword Search, which allows students to locate specific information based on their particular needs. The other one is Hierarchical Map, which provides a global picture of the subject content. Their interactions with the WBL system were recorded in log files. Furthermore, all of the participants were initially required to take a SPQ, which is an 18-item inventory for categorizing students as Holists or Serialists (Ford, 1985). According to the results of the SPQ, there were 26 Holists and 18 Serialists and no intermediate students. Subsequently, all participants needed to take the pre-test to identify their preliminary understanding of the subject content. In the next stage, they were required to complete practical tasks by interacting with the WBL system. More specifically, they needed to complete the tasks by finding information from the WBL system. Finally, the participants were requested to take the post-test to evaluate their learning performance. Both of the post-test and pre-test included 20 multiple-choice questions.

Data analyses were conducted using traditional statistical and data mining techniques. The former was applied to determine whether there are differences between the DC scenario and the MD scenario. The latter was employed to produce clusters of students that shared similar learning behavior, and subsequently the corresponding cognitive styles and learning performance for each cluster were identified. Among various data mining techniques, K-means was used to create clusters for this study because our recent studies (e.g., Chen & Liu, 2011) found that K-means is a useful tool to cluster students' behavior.

3. Results and Discussions

3.1 Learning Behavior

After carefully examining students' learning behavior showed in their log files, we found that the hierarchical map was rarely used, regardless of the DC scenario (Mean = 8.73; SD = 16.46) or the MD scenario (Mean = 11.67; SD = 33.17). Thus, the frequencies of the use of the hierarchical map were excluded. In other words, only five attributes are considered as the inputs of the K-means algorithm: (1) the frequencies of using Keyword Search, (2) the frequencies of making movements, (3) the frequencies of repeated visits, (4) the number of pages browsed, and (5) the time spent for completing the tasks. As showed in Tables 1 and 2, students used fewer keywords ($t=5.129$; $p<.01$), made fewer movements ($t=3.031$; $p<.05$), had fewer repeated visits ($t=4.962$; $p<.05$), browse fewer pages ($t=4.987$; $p<.001$) and spent less time for completing the tasks ($t=3.987$; $p<.05$) in the DC scenario than those in the MD scenario. These findings imply that students in the former had more engagements than those in the latter.

3.1.1 The Desktop Computer Scenario

As showed in Table 1, students' learning behavior in the DC scenario are grouped based on the following trends:

- **C 1:** Students had the lowest frequencies of using the Keyword Search, made the fewest movements, made the fewest repeated visits and browsed the fewest pages.
- **C 2:** Students spent the least task time among the three clusters. However, they made the most movements, the most repeated visits and browsed the most pages.
- **C 3:** Students spent the most task time and had the highest frequencies of using the Keyword Search.

After checking the corresponding cognitive style for each cluster, we found that the distribution of Holists and Serialists in each cluster is similar. In other words, cognitive styles did not affect students' learning behavior in the DC scenario. A possible reason is that DCs have been the mainstay of the computing world for more than 20 years (Masters & Ellaway, 2008). Nowadays, most students are familiar with DCs so cognitive styles have no effects on students' learning behavior in the DC scenario.

Table 1. The mean and standard deviation of each attribute in the DC scenario

Attributes		**Overall**	**C 1**	**C 2**	**C 3**
Task Time	Mean	2974.52	2918	2844.71	3258
	SD	546.79	616.28	420.67	576.82
Keyword Search	Mean	34.52	18.89	44	49.4
	SD	20.13	9.12	22.44	10.97
Movement Made	Mean	95	27.67	174.71	104.6
	SD	68.75	7.23	32.63	22.24
Repeated Visits	Mean	19.38	5	35	23.4
	SD	14.96	2.35	9.73	7.37
Pages Browsed	Mean	75.62	22.67	139.71	81.2
	SD	54.34	5.22	25.06	15.19

3.1.2 The Mobile Device Scenario

As showed in Table 2, the trends of students' learning behavior in the MD scenario are:

- **C1:** Students spent the least task time and had the lowest frequencies of using the Keyword Search of the three clusters. Moreover, they made the fewest movements, made the fewest repeated visits and browsed the fewest pages.
- **C2:** Students had the highest frequencies of using the Keyword Search, regardless of the DC scenario or MD scenario.
- **C3:** Students spent the most task time, made the most movements, made the most repeated visits and browsed the most pages among the three clusters.

Cluster 2 (N=9, 39%) and Cluster 3 (N=9, 39%) are the two major clusters in the MD scenario. After identifying the corresponding cognitive style of each cluster, we found that most Holists (N =6, 42.86%) appeared in Cluster 3. As mentioned earlier, students in Cluster 3 made the most movements. These findings reveal that Holists tended to make a lot of movements. This may be due to the fact that Holists prefer to get an overview so they tend to get a global picture by making many movements. Conversely, most Serialists (N =5, 55.56%) emerged in Cluster 2, where they frequently used the Keyword Search. A possible reason is that Serialists tends to focus on procedural details when processing information in a learning context. On the other hand, the Keyword Search, which can facilitate students to locate specific information, is useful for Serialists to get particular details (Pask, 1976).

Table 2. The mean and standard deviation of each attribute in the MD scenario

Attributes		Overall	C 1	C 2	C 3
Task Time	Mean	3606.83	2564	3799.33	3993.67
	SD	733.88	366.4	647.38	324.27
Keyword Search	Mean	44.74	34.2	51.44	43.89
	SD	12.11	12.62	6.56	12.64
Movement Made	Mean	176.91	131.4	166.78	212.33
	SD	39.55	8.59	18.57	33.15
Repeated Visits	Mean	35.43	25.8	33.33	42.89
	SD	8.9	3.77	5.7	7.27
Pages Browsed	Mean	141.48	105.6	133.44	169.44
	SD	32.26	8.08	14.89	29.16

3.2 Learning Performance

In general, we found that students in the MD scenario (Mean = 14.5 SD=1.2) performed better than those in the DC scenario (Mean = 10.1 SD=0.7). As mentioned in Section 3.1, students with the MDs had more engagements than those with the DCs. These findings suggest that students can benefit from such engagements to get better performance. Furthermore, we also examined how different cognitive style groups performed differently in the DC and MD scenarios. Regarding the DC scenario (Figure 1), students in Cluster 3 got the highest gain score (Mean = 11.6; SD= 2.88) while those in Cluster 1 got the lowest gain score (Mean = 9.67; SD= 3.61). As described in Table 1, students in Cluster 3 most frequently used the Keyword Search while those in Cluster 1 least frequently used the Keyword Search. These findings suggest that frequently using the Keyword Search can help students explore a variety of concepts so that they achieve good performance. Moreover, we also found that Holists got higher gain scores than Serailsts in each cluster (Figure 1). In other words, Holists performed better than Serailsts in the DC scenario.

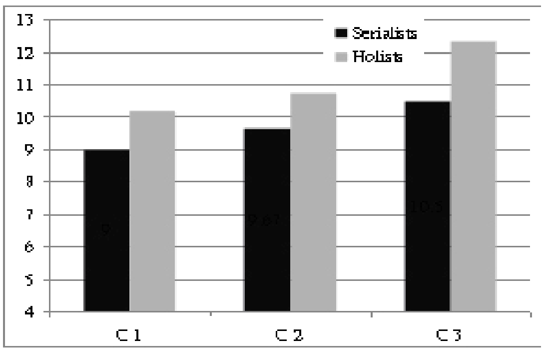


Figure 1. The performance of DC scenario

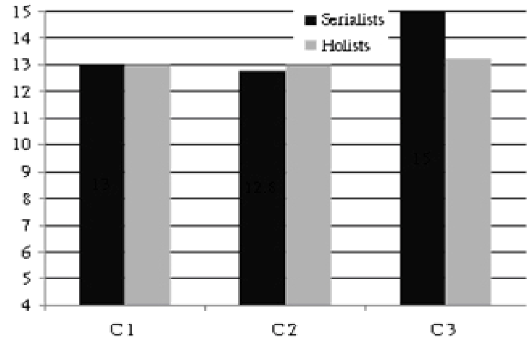


Figure 2. The performance of MD scenario

Regarding the MD scenario (Figure 2), students in each cluster got similar gain scores in the MD scenario. Thus, students' learning behavior is not associated with their learning performance in MD scenario. We, however, found that Serialists in Cluster 3 got the highest gain score. As mentioned earlier, student in Cluster 3 made the most repeated visits and browsed the most pages. These results reveal that making more repeated visits and browsing more pages can support Serialists to obtain better learning performance in the MD scenario. Unlike Serialists, Holists demonstrated similar performance in each cluster. In other words, Holists' learning behavior does not affect their learning performance.

4. Conclusions

Two research questions are examined in this study. The answer to the first research question is that students in the MD scenario had more engagement and performed better than those in the DC scenario. The answer to the second research question is that Holists who made more movements and Serialists who frequently used the Keyword Search could achieve good performance in the DC scenario. On the other hand, the students performed similarly in the MD scenario though Serialists who made more repeated visits and browsed more pages had better performance. The present study shows fruitful results but there are several limitations. Firstly, this study was only a small-scale sample. Further research needs to be undertaken with a larger sample to provide additional evidence. Another limitation of this study is that only cognitive styles were investigated. Thus, it is necessary to consider other human factors, such as gender difference and prior knowledge, in the future. Such evidence can not only be helpful to promote the use of MDs, but also is useful to incorporate personalization into ubiquitous learning environments.

Acknowledgements

The work presented in this paper is funded by National Science Council, Taiwan (NSC 101-2511-S-008 -010-MY3, NSC 99-2511-S-008-003-MY2, NSC 100-2631-S-008-004).

References

[1] Cavus, N. (2011). Investigating mobile devices and LMS integration in higher education: Student perspectives. *Procedia Computer Science*, *3*, 1469-1474.

[2] Chen, S. Y., & Liu, X. (2008). An integrated approach for modeling learning patterns of students in Web-based instruction: A cognitive style perspective. *ACM Transactions on Computer-Human Interaction*, *15*(1), Article 1.

[3] Chen, S. Y., & Macredie, R. D. (2004). Cognitive modeling of student learning in Web-based instructional programs. *International Journal of Human-Computer Interaction*, *17*(3), 375-402.

[4] Chen, S. Y., & Liu, X. (2011). Mining students' learning patterns and performance in Web-based instruction: A cognitive style approach. *Interactive Learning Environments*, *19*(2), 179-192.

[5] Chen, S. Y., & Macredie, R. D. (2002). Cognitive styles and hypermedia navigation: development of a learning model. *Journal of the American Society for Information Science and Technology*, *53*(1), 3-15.

[6] Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology*, *16*(1), 65-79.

[7] Gulati, S. (2008). Technology-enhanced learning in developing nations: A review. *International Review of Research in Open and Distance Learning*, *9*(1), 1-16.

[8] Hein, T. L., & Irvine, S. E. (1999). Technology as a teaching learning tool: Assessing student understanding in the introductory physics lab. *Annual conference of the American Society for Engineering Education, Charlotte*, North Carolina, 1999, Session 2380.

[9] Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning and instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

[10] Kukulska-Hulme, A. (2007). Mobile usability in educational contexts: What have we learnt?. *International Review of Research in Open and Distance Learning*, *8*(2), 1-16.

[11] Masters, K. & Ellaway, R. (2008). e-Learning in medical education Guide 32 Part 2: Technology, management and design. *Medical Teacher*, *30*(5), 474-489.

[12] Messick, S. (1976). *Individuality in learning*. San Francisco: Jossey-Bass.

[13] Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology*, *46*(2), 128-148.

[14] Zhang, D. (2007). Web content adaptation for mobile handheld devices. *Communications of the ACM*, *50*(2), 75-79.

Integrating Collaborative and Mobile Technologies for Fostering Learning about Negotiation Styles

Dan Kohen-Vacs[ab*],Miky Ronen[a],Orna Bar-Ness [a], Marcelo Milrad[b] & Arianit Kurti[b]
[a]*Holon Institute of Technology, ISRAEL*
[b]*CeLeKT, Linnaeus University, Sweden*
*mrkohen@hit.ac.il

Abstract: This paper presents and discusses our efforts aiming to expand the potential of an existing CSCL environment through the integration with additional software applications in order to support a cross context learning activity. We used a CSCL environment for supporting asynchronous types of interactions, mobile devices for face-to-face interaction and a dedicated web application for self-assessment. We present the design and implementation of a scripted learning activity that deals with negotiation styles and describe the integration of different software applications that supported the students' interactions along the various activity phases. The results indicate the potential and benefits of the integrative approach using collaborative and mobile technologies in order to support and enhance a wide range of pedagogical activities.

Keywords: Collaborative learning, script, negotiation, collaborative and mobile technologies, systems integration

1. Introduction

A CSCL (Computer Supported Collaborative Learning) script describes a well-defined instructional strategy organized into a collaborative pedagogical activity and supported by the use of information and communication technologies (ICT) [1]. According to Dillenbourg and Jerman [2] a CSCL script addresses five main attributes of the collaborative learning process: the task that students have to perform, the composition of the group, the way that the task is distributed within and among groups, the mode of interaction and the timing of the phase. A script may also include students' interactions with learning resources originating from different sources like teacher's materials or with emerging learning objects (ELOs) contributed by their peers [3, 4]. Learning activities may involve distinct learning contexts in which the learners' trajectory should go through [5]. This multiplicity of contexts may influence the activity trajectory and challenges the activity planners to find ways in order to provide a seamless learning experience [6]. Seamless learning implies that learners can learn whenever they are curious in a variety of scenarios and that they can switch between these learning settings easily and quickly using their portable device as a mediator [7]. The need for supporting a seamless learning experience becomes prominent when considering Goodyear's claim [8] addressing two perceptible changes in the field of educational research. The first is a shift in our sense of the spaces and contexts in which education takes place, as different learning activities are becoming more commonly distributed across a variety of contexts. The second change is a wider understanding with the conception of educational praxis, acknowledging the growing importance of design. The current areas of focus addressed in this paper are grounded in these two major changes: 1) The design of an interactive learning environments that include CSCL scripts and different software tools to support a seamless learning experience 2) Exploring how the proposed approach can be integrated into everyday educational practices in order to become a sustainable part of the learning environment. Addressing these two aspects calls for an integrated approach for learning design based on advancing the current socio-technological configurations available in educational settings nowadays. In this paper, we describe our current efforts to design and enact a particular learning activity dealing with negotiation styles and its implementation with university students. The topic of negotiation styles is traditionally taught in a regular face-to-face session performed in a regular classroom [9]. However, this topic presents many opportunities to be taught in different ways that implements advanced learning approaches enabling students to cope with real life negotiation opportunities [10]. We present the pedagogical requirements and goals and how these have been transformed into actual learning tasks supported by different ICT solutions. The actual implementation of these goals relies on the use of various CSCL scripts and software applications in order to ensure a seamless learning experience.

2. Learning about negotiation styles

2.1 Integration of cross context learning with CeLS

The interactive learning environment that was used to support this activity was Collaborative e-Learning Structures (CeLS) [11]. CeLS enables teachers to design and enact online collaborative activities using various pedagogical approaches such as collaborative problem solving, peers' products assessment, competition, jigsaw and combinations of the above. CeLS was originally developed for asynchronous activities performed via stationary or laptop computers. However, a CeLS script can also include notations that address interactions to be performed with other communication technologies like mobile phones or dedicated applications [12]. In the next section we describe such integration designed to provide and support a seamless learning experience.

2.2 Activity script

The learning activity was designed for support undergraduate and graduate courses dealing with negotiation and conflict management. The main goal of this activity is to familiarize learners with the concept of negotiation styles, to develop their ability to argue according to a certain style and to identify a person's style according to the arguments he expresses during a negotiation. The activity was planned to be implemented after the teachers' introduction of Rahim's model [13] defining 5 types of negotioation styles. Table 1 summarizes the structure of the activity and its interrelated pedagogical, technological and implementation aspects.

2.3 Technological integration and implementation

The activity script enactment is supported by integrated technologies. CeLS environment serves as a main technological platform for the enactment of the activity script and its different phases. Two stand-alone applications, namely SMS-HIT and NeSI [12] have been developed to support the different phases, as described in Table 1. These applications are used as extensions that enable to expand some of CeLS ´capabilities. SMS-HIT and NeSI have their dedicated run time engines that generate interaction pages supported by a variety of end user devices according to predefined activity properties. The user interactions with both applications are stored in their dedicated SQL servers using XML data format. This information is retrieved following to a CeLS request. The retrieval process is performed by a dedicated middleware application that analyses the CeLS request and performs a data migration to the CeLS database. This process is performed according to the CeLS script rules using the phase and the building block identifiers [12]. Figure 1 illustrates the script implementation supported by integrated technologies.

Table 1: Description of the script pedagogical, technological and implementation aspects

Phase	(1) Self diagnosis	(2) Learner's reaction according to style	(3) Identifying statements styles	(4) Negotiation style diagnosis	(5) Summary and debriefing
Students' Task	Identify your primary and secondary negotiation styles according to Rahim's model [13]	An employer-employee conflict scenario is presented. Express an argument reflecting your negotiation style from both positions	Students are presented with several peers' statements. Identify the negotiation style of each statement.	Diagnosis of personal negotiation styles using a validated questionnaire [13].	Class discussion based on a comparative representation of students' inputs contributed during the activity phases.
Instructional Goal	Understand the negotiation concepts by relating student's implementation to personal situation as a motivational strategy	Develop student's ability to argue during a negotiation process according to one of the styles	Foster student's understanding of the negotiation concepts and their ability to relate statements to negotiation styles	Compare the declared personal negotiation style with the profile diagnosed by an objective measuring tool	Summarize and reflect upon the activity phases
Type of interaction	Face to face	Asynchronously			Face to face
Place of interaction	Classroom	Anywhere			Classroom
Duration	1 class session	2-3 days	2-3 days	1 class session	1 class session
System	SMS-HIT PRS	CeLS		NeSI	CeLS
Interaction device	Mobile	Stationary or laptop			Teacher's computer
Data format and storage	SMS-HIT SQL server database	XML structure stored in CeLS SQL server database		NeSI SQL server database	CeLS SQL server database
Integration methods	Integration of SMS-HIT data into CeLS database	Integration of SMS-HIT data into CeLS database by data fetcher using an XML data structure	No need of integration at this phase	Integration of NeSI data into CeLS database by data fetcher using an XML data structure	Reuse of integrated data from phases 1 and 4

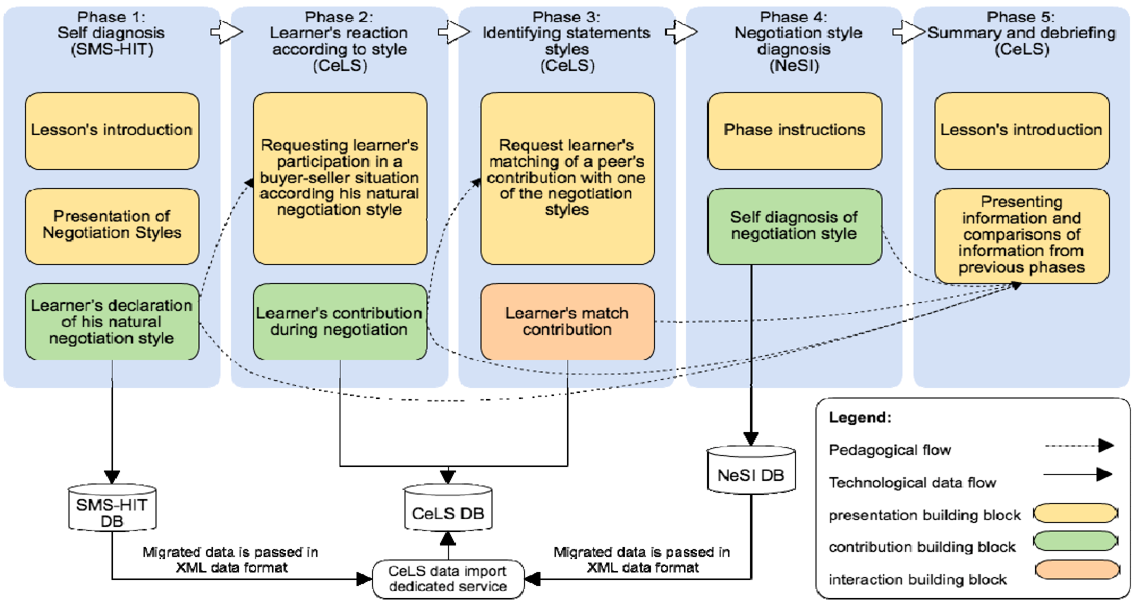


Figure 1: The script for supporting learning process for negotiations styles

3. Evaluation of the learning activity

The learning activity was tested with 25 students in a graduate course that took place during the spring semester of 2012. The evaluation was based on in-depth data analysis of all the students' contribution during the activity phases using the data stored in the CeLS database, observations of the class sessions and open interviews with the students and the teacher. We present a glimpse to the analysis of students' contributions. Most students (64%) managed to perfectly identify their own style (phase 1). Most students (60%) managed to partially match their peers' statements (phase 2) with the declared style while only 20% managed to achieve a perfect match. A further data examination reveals that all the statements that were perfectly matched belonged to students that have perfectly identified their own style. This finding is not surprising since students who were not fully aware of their own negotiation style may not have express it clearly enough and as a result peers could not definitively associate the argument with a style.

During the interviews students expressed their personal impressions about the experience, described their level of engagement and evaluated the usefulness of the activity for promoting their understanding of the new concepts. Most students mentioned the structure of the activity as an aspect that provided an intriguing way to acquire and practice theoretical skills. They also appreciated the use of the SMS-HIT application during the face-to-face session for providing meaningful interactivity by using the mobile technologies that are available to everyone. Learners were asked about the added value of the transferability and reuse of the knowledge acquired from one phase into the following phases along the activity. They considered these aspects as a key factor that enabled knowledge acquisition, peer learning, learning in action during negotiation practices and finally synthesis of new insights.

The teacher had 14 years of previous experience of teaching this subject and trying to conduct similar activities without any kind of ICT support. She considered the activity structure and its supportive software applications as convenient means that enabled her to conduct cross context activities that were not possible before. The major challenge mentioned by the teacher was the reliance on learners' cooperation because of the interdependencies between the phases.

4. Concluding remarks

Pedagogical activities may comprise of a variety of ways in which students interact through the use of different technological means along the learning activity phases. We have briefly presented our current efforts aiming to expand the potential of an existing CSCL environment through the integration with additional software applications in order to support a cross context learning activity. In our case, we used a CSCL environment for asynchronous types of interactions, mobile devices for face-to-face interaction and a dedicated web environment for self-assessments. We also presented the relevancy and implications for design of the mobile seamless learning dimensions along the activity phases that are related to the following aspects: students' shift between personal and social learning, phases that occur across time and location, student's accessibility to ubiquitous learning materials along the phases, involvement of physical and digital learning environment, shift between learning tasks along the phases and finally, knowledge synthesis that occurs along the activity. We have also illustrated the kind of ICT support that can be provided in order to facilitate a smooth transition between the different phases, thus supporting a seamless learning experience. One of the major challenges of today's education is no longer about finding the best ways for knowledge delivery, but rather designing, developing and delivering learning environments, digital tools and activities for learners to construct knowledge by engaging and inspiring them to learn. These aspects have inspired the efforts described in this paper and we will continue working in this direction. Our forthcoming efforts will deal with the refinement of our approach and the development of different software applications (with a focus on mobile devices) and their integration within other subject domains. One of our objectives is to further understand the potential and challenges involved in the integration of software applications and user generated content designed for supporting learning across contexts using collaborative scripts.

Acknowledgements

We thank Hagai Reuveni and Shavit Cohen for their efforts during the design and development of the CeLS, Noga Lavi, Oren Ben-Aharon and Adi Silagy for their efforts to develop the SMS-HIT and Natali Asher and Limor Abraham for the development of the NeSI.

References

[1] Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In Kirschner P. (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 61-91). Heerlen, The Netherlands: Open University of the Netherland

[2] Dillenbourg, P. & Jermann, P. (2010). Technology for Classroom Orchestration, in: M. S. Khine & I. M. Saleh (Eds) *New Science of Learning: Cognition, Computers and Collaboration in Education* (pp. 525 – 552). Springer.

[3] Hoppe, H. U., Pinkwart, N., Oelinger, M., Zeini S., Verdejo, F., Barros, B., & Mayorga, J. I. (2005). Building bridges within learning communities through ontologies and "thematic objects". Proceedings of the *International Conference on Computer Supported Collaborative Learning, International Society of the Learning Sciences*, pp. 211-220.

[4] Ogata, H., Li, M., Hou, B., Uosaki, N., El-Bishouty, M., & Yano, Y. (2011). SCROLL: Supporting to Share and Reuse Ubiquitous Learning Log in the Context of Language Learning. *Research and Practice in Technology Enhanced Learning*, Vol. 6, No. 2, pp. 69–82.

[5] Vavoula, G., Sharples, M., Rudman, P., Meek, J., & Lonsdale, P. (2009). Myartspace: Design and evaluation of support for learning with multimedia phones between classrooms and museums. *Computers and Education*, 53(2), pp. 286-299.

[6] Chan, T.-W., Roschelle, J., His, S., Kinshuk, Sharples, M., Brown, T., Patton, C., Cherniavsky, J., Pea, R., Norris, C., Soloway, E., Balacheff, N., Scardamalia, M., Dillenbourg, P., Looi, C. K., Milrad, M., &Hoppe, U. (2006). One-to-one technology-enhanced learning: An opportunity for global research collaboration. *Research and Practice in Technology Enhanced Learning*, 1(1), pp. 3-29.

[7] Spikol, D., Milrad, M. Svensson, M., Pettersson, O. and Persson, M. (2008). Mobile Collaboration Tools and Systems to Support Ubiquitous Learning. Proceedings of the *Fourth Annual Conference on Collaboration Technologies 2008*, Wakayama, Japan August 30-31, 2008.

[8] Wong, L.-H, & Looi, C.-K. (2011). What seams do we remove in mobile-assisted seamless learning? A critical review of the literature. *Computers & Education*, 57(4), pp. 2364–2381.

[9] Core, M., Traum, D., Lane, H.C., Swartout, W., Gratch, J., Lent, M. van, Marsella, S. (2006). Teaching Negotiation Skills through Practice and Reflection with Virtual Humans. *SIMULATION*, Vol. 82, No. 11, pp. 685-701.

[10] Schneider, Andrea Kupfer. (2000). Building a Pedagogy of Problem-Solving: Learning to Choose Among ADR Processes. *Harvard Negotiation Law Review*, Vol. 5, 2000.

[11] Ronen, M. & Kohen-Vacs, D. (2010). Modeling, enacting sharing and reusing online collaborative pedagogy with CeLS. In: Persico, D. & Pozzi, F. (Eds.), *Techniques for Fostering Collaboration in Online Learning Communities: Theoretical and Practical Perspectives*, IGI Global, pp. 319-339.

[12] Kohen-Vacs, D., Kurti, A., Milrad, M. and Ronen, M. (2012). Systems Integration Challenges for Supporting cross Context Collaborative Pedagogical Scenarios, submitted to the 18th *CRIWG Conference on Collaboration and Technology*, Raesfeld, Germany.

[13] Rahim, M. (1983). A Measure of Styles of Handling Interpersonal Conflict. *Academy of Management Journal*, vol. 26, no. 2, pp. 368 – 376.

Using Mobile Digital Storytelling to Support Learning about Cultural Heritage

Susanna NORDMARK[*] & Marcelo MILRAD
School of Computer Science, Mathematics and Physics, Linnaeus University, Sweden
*susanna.nordmark@lnu.se

Abstract: This paper presents our on-going research related to the use of mobile digital storytelling (mDS) as means for promoting innovative ways of collaboratively learning about cultural heritage. In a joint effort we have worked closely together with our fellow stakeholders – a local museum and a primary school – on actively engaging a group of Swedish school children and their teachers in using mobile digital storytelling as facilitator for introducing the concept of seamless learning. The outcomes of these efforts provided us with valuable insights on how to proceed further in our work of designing versatile mDS applications supporting collaborative learning in authentic contexts.

Keywords: Mobile Digital Storytelling, Mobile Seamless Learning, Collaboration, Participation, Cultural Heritage

Introduction

The objective for the study reported in this paper was to continue our investigations into the development of mobile applications and systems and their integration with pedagogical design approaches for supporting the flow of seamless learning [8]. In this paper we describe and discuss our efforts while developing a seamless and integrated flow of formal learning activities supported by a designated application for mobile digital storytelling (mDS) that we have developed. For this particular case, we wanted to provide and support teachers and young learners with new educational practices and ways of individual and group interaction as well as group dynamics in the specific topic of cultural heritage. In our previous studies [9, 10], we have identified some of the essential components and conditions that may serve for developing an mDS framework, i.e. a complete toolkit for introducing and supporting the methodology and technology in various educational settings. One of the objectives of the current study was to gain additional knowledge and insights useful for developing such a framework.

The concept of mobile seamless learning (MSL) suggests that learners could learn ubiquitously, i.e. regardless of where, when and how, and that they can switch effortlessly between these various learning scenarios by using a mobile device as mediator [8]. Wong and Looi [14] have identified the following ten different dimensions applied to MSL: 1) formal and informal learning, 2) personalized and social learning, 3) learning across time, 4) learning across location, 5) ubiquitous access to learning resources, 6) physical and digital worlds, 7) multiple device type, multiple learning tasks, 8) switching between multiple learning tasks, 9) knowledge synthesis, and 10) multiple learning models. In section II we illustrate in more detail how we integrated the MSL dimensions with each phase of the study. Today's wide-ranging use of mobile devices offers a unique potential and support for learning interaction regardless of context, well in tune with the MSL dimensions above. Keeping in mind Goodyear's [6] declarations on two major changes in contemporary educational research: our altered understanding of when and where learning actually takes place, and our heightened perception of the importance of learning design, the present study is inspired by a design-based research approach [5], or a "design intervention", where the research impacts on and leads to upward spirals of revision, in close relation to user needs and experiences, and to earlier findings [9, 10].

1. Using Mobile Digital Storytelling for Learning: Motivation and Study Rationale

Applying digital storytelling in learning contexts is a relatively new field of research within Technology Enhanced Learning (TEL). Adding a mobile dimension to the concept of digital storytelling for learning might provide added opportunities for supporting the experience of seamless learning [14] - i.e. a time and context independent learning model - than using a solution of a more stationary kind. How factors of digital software and web 2.0 social media may be changing patterns of stories are discussed in [1], in the sense that stories are now open-ended, branching, participatory, and unpredictable, and may be revealing new directions for how we tell narratives, which is also discussed in [11 & 12]. Lombardo and Damiano [7] have worked on a system supporting contextually aware storytelling units, using methods addressing both interactivity and movement to adapt the flow of stories as to how people actually move around in exhibition areas in a museum. In this case, the system acted as the storyteller, not the user, delivering its stories depending on where the visitors chose to go. Callaway et al. [3] present a mobile museum system delivering slightly different dramatic stories to participants of smaller visiting groups, encouraging the group to share their experiences through animated discussions. Druin, Bederson and Quinn [4] have developed a mobile application called StoryKit, intended at intergenerational storytelling. Bidwell, Reitmeier, Marsden & Hansen [2] have developed StoryBank, a mobile tool aimed at cross-cultural storytelling. Although these efforts have explored how digital storytelling can be used to support different aspects of formal and informal learning, none of them have explicitly investigated how to utilize mDS to specifically support and provide a mobile seamless learning experience.

Our purpose for developing an mDS application of our own is twofold: 1) to provide users with a number of functionalities for mDS that are not supported by other existing applications [9], and 2) to offer an important component that will be part of a complete toolkit supporting the entire learning experience of mDS: from the workflow of the learning activities, via relevant technical features and scaffolding tools, to a web service for sharing, storing and retrieving stories, together with a collaborative editing feature for continued work. Furthermore, we are aiming at creating a cross-platform application which is not contextually restricted or bound to a certain subject or activity, but rather dynamically adaptive to whatever purpose it is used for. The research efforts connected to this study involve different aspects of TEL design and implementation, innovative educational practices together with usability and sustainability, closely connected to the framework of Mobile Seamless Learning described above. We are predominantly exploring what factors may transform TEL designs into functional, sustainable technologies in combination with approaches supporting innovative designs for teaching and learning in the 21st century.

2. Design and Development of the Learning Activity

2.1 Methodological Approach and Implementation

The following section describes the flow and content of the learning activity as well as the settings in which it took place, together with a brief presentation and overview of the technology designed and used for the activities.

The study was designed as a series of five interrelated intervention phases, actively involving 4 teachers and 53 children in the ages of 7–8 from a local primary school together with staff from the museum of Kulturparken Småland, and the authors of this paper. It was conducted over a total time period of three weeks during May 2012, and cautiously adapted to the guidelines of the national curriculum. Consequently, the main learning objective for the children was to achieve a deeper understanding about the notions of time and space by exploring, studying and sharing historical sites and events in their school neighborhood. Divided into 18 groups of 3 pupils in each, the children initially took a guided tour with a museum curator to learn about some of the historical features in their school vicinity. The tour content was augmented using iPod Touch with a guide application developed by us, providing a tour map with five designated points of interest (POI), each revealing a series of historical images, accompanied by sound where applicable. During the tour, the groups used mind maps for writing down a minimum of four keywords at each POI. The mind maps were primarily intended for memory support when later selecting the story theme and creating the script for the voiceover recordings. At the POIs, the children took turns taking notes and using the iPods for photographing objects of interest, later serving as digital content for creating their stories. Thus, each group member took an individual and active part in the gathering of materials and the story creation process. However, before venturing into the concluding mDS stage, each group had to examine and discuss all their images and keywords, agree on what their story should tell, and which images and script would best support their chosen theme.

After the day of the tour and story creation, the children met the researchers at three further occasions. The purpose of those meetings was to conduct a discussion concerning the use and perceived benefits and disadvantages of the mobile applications, and to accomplish a reflective discussion about the stories, their content and the overall story creation process. Fig. 1 below outlines and describes the activity phases, their settings and content:

Figure 1. Study Activity Outline

	Theme and setting	Researcher & museum staff activities	Teacher and student activities	MSL dimensions
Phase 1	**Introduction,** 60 min, @ classroom	Staff presentation, activity and assignment introduction | Mobile device introduction and software overview, mDS instruction. | "Tour rules": when to photograph, when to open POI's, when to use mind maps, etc.	Presentation, discussion, questions	1-2, 5, 8 - 10
Phase 2	**Inspirational tour and mDS activity,** 5 hours per day, @ outdoors	Division into workgroups | Repetition of software and storytelling instruction, repetition of tour rules, hand-out of mobile devices and mind maps | Run tour | Manage hands-on activities at museum grounds | Run storytelling activity.	Take tour: Photograph | Mind-map key words and take notes | Participate in hands-on activities | Accomplish mDS assignment: a) Negotiate and decide story theme, b) Select which photos to use, c) Storyboard and script, d) Run mDS app: add and arrange photos, add story title, record voiceover, preview, edit, finish. | Hand in device and materials.	1-5, 8-10
Phase 3	**Concluding activity,** 2.5 hours, @school yard	Activity introduction | Run activity | Finish up: what happens next time we meet?	Collaboratively discuss, plan and create a physical time line of tour and stories, using a rope divided into time sections combined with images from tour app + images of own house + images from stories.	1-5, 9-10
Phase 4	**Application evaluation,** 30 min per group, @ school	Activity introduction | Initiate and moderate usability discussion using app screen shots with smiley indicators.	[Groups A - I focus on mDS app Groups J - R focus on guide app] In groups discuss and answer a series of questions relating to the applications used during the tour and the storytelling activity.	1-5, 8-10
Phase 5	**Reflective discussion on story content and work process,** 40 min per group, @ school	Activity introduction | Initiate and moderate discussion using app screen shots with smiley indicators + post-it notes.	Watch all stories | With focus on own story: Name 4 group views on the following components: image use, sound quality, theme relevance, and overall impression. | With focus on all stories: a) If you were to recommend your friends to watch these stories, how would you describe them? What would you say? b) Discuss the following: If your group were to create a new story, would you do anything differently? If so, what? If not, why? c) If you were to help some new group members create a mobile digital story, what would you tell them? What do you think is important to know and consider?	1-5, 8-10

Inspired by the ideas and theoretical concepts described in the previous sections, we strived to address and incorporate relevant criteria and aspects of MSL in all phases, as presented in Fig.1 above. After completing all phases, the authors lastly returned to the school to meet with the teachers and museum staff to discuss issues related to the story creation and production, the collaborative work process and the general views of the activity.

2.2 The mDS application

In this subsection we give a brief description of the mDS application that we have designed and developed. Our intention is to provide built-in support and encouragement of the pedagogical ideas we established and implemented through our previous trials, together with stable and relevant mDS functionality. Some of the specific functionalities we have developed that we see as crucial for the mDS workflow, and that are not or only partly included in other mDS apps available, are e.g.: 1) an image-by-image voiceover recording for each of the included images, allowing the user to smoothly edit and redo the recordings for each step, 2) the possibility to add text to images, and to add, design and edit title and credit pages, and 3) the possibility to add two separate soundtracks for voiceover and background music or sound effects. From a technical perspective, we chose to use HTML 5 and JavaScript in order to support a mobile application that partly has access to the native device features (camera, voice recording, etc.), and that can run across platforms. For this particular study, the application code has been deployed on the iPod Touch using PhoneGap and Xcode. Each specific mobile digital story generated by this application contains images, sounds and transitions stored with a unique ID identifier, allowing storing them in a well-defined data structure. In the current implementation we used JSON (JavaScript object notation) for the data structure as it provides a lightweight data-interchange format that is becoming a "standard" for web development. This approach allows for content reuse and the potential integration of the digital materials generated for the different stories with other collaborative web applications that will be developed in our future work.

3. Study Outcomes

In this section we describe the outcomes of the learning activities described above. The data we collected during the different phases mainly consists of three kinds of video: the children's stories, the observation films from the activity, and the videos from the evaluation sessions. Also, we collected the mind maps, the scripts and the notes from each group, along with notes from the simplified usability test and the reflective discussions.

The stories created by the 18 groups covered four main themes covering various aspects related to the neighborhood's development through time and space: the old dairy/brewery, the plague grave, the girls' school, and the railway restaurant. As it turned out, most groups chose to focus their stories on the plague grave, and only one of the total 18 groups chose to tell their story about the school for girls. To get additional information on how the children perceived handling the mDS application and experienced the work method, a simplified usability test and a follow up discussion concluded the study (see Fig. 1). In phase 4, we used screenshots of all nine steps of the mDS process, applying smiley icons to help the children relate their discussions to their perceptions and experiences of the app [13]. The children's young age had fuelled concerns on their prospects of handling the varied and complex assignments as intended. We therefore piloted careful discussions with the teachers and museum staff to try to facilitate as much as possible. However, the groups managed very well to handle both the technology and most of the complications that occurred, in general by trying out each other's suggestions. The overall familiarity with technology actually proved to be much more advanced than expected, which greatly influenced their capability of coping. Most technical incidents occurred in relation to the voiceover recording, which is not an easy task for any age group, and many of the children agreed that the interface for the voiceover recording could be improved. Several groups also reported that when problems occurred that none of the group members could solve, they rather started all over again than persisted in trying to solve that specific difficulty. However, this "start-all-over-approach" resulted in two stories with no or extremely distorted sound, and one story not showing any of the selected images. Most of the "incomprehensible" issues occurred in relation to the overall mDS process, i.e. the how, what, when and who. Therefore, a vital lesson learned to remember well when venturing into the next design iteration, is to make all process steps and stages of the mDS workflow as clear and self-regulated as possible, also so since very few of the children paid attention to the built-in reminders and instructions available.

4. Discussion and Concluding Remarks

This study has integrated mDS as a facilitator for encouraging a mobile seamless learning experience when learning about cultural heritage. The study was conducted in May 2012 as a stage in further developing a framework and toolkit for mobile digital storytelling for educational purposes. The outcomes supplied us with valuable feedback on how to proceed with the next iteration of design and development. One of the most interesting lessons learned from working with children as young as these was they did not encounter as many problems in handling the technology, as with coping with the mDS methodology, which after concluding the last two study phases stood out as the greatest obstacle for the children to overcome. Most children also found it challenging to account for their decisions and choices throughout the process. Few of them had a clear sense of why a specific theme was finally chosen, or how they had gone about selecting the images to go with that same theme. Conversely, it was much easier for them to discuss how they went about creating the script for the voiceover recording, probably so due to the mind maps in which they gathered keywords from each tour POI. Contrariwise, the general concern of the teachers was focused on handling and mastering the technology rather than understanding the mDS workflow, which none of them reported to struggle with. All teachers agreed on a great need to learn more about mobile units and the applications used before daring to run similar activities on their own.

For the reported technical issues there are now well-advanced redesigns awaiting final implementation. One of the highest priorities is to get the voiceover interface stabilized and easier to handle, along with designing a contextualized help option assisting each step of the mDS process. Also, two separate workflows will be implemented: a basic alternative covering essential features only, and an advanced ditto providing the user with a more substantial toolbox. Part of the next iteration of the full mDS toolkit development will also be to carefully test the consequences of the suggestions of improvement derived from this particular study, along with some of the before planned alterations, e.g. options for sharing. The qualities of MSL have inspired the efforts described in this paper and we will continue exploring and developing learning activities that encourage learners to collaboratively construct knowledge in engaging and inspiring ways.

References

[1] Alexander, B. & Levine (2008). Web 2.0 storytelling: Emergence of a new genre. *Educause review*, November/December, pp. 40-56.

[2] Bidwell, N.J. & Reitmeier, T., et al (2010). Designing with Mobile Digital Storytelling in Rural Africa. *Proceedings of CHI 10*, p. 1593 – 1602.

[3] Callaway, C. et al. (2012). Mobile Drama in an Instrumented Museum: Inducing Group Conversation via Coordinated Narratives. *New Review of Hypermedia and Multimedia*, p 37-61.

[4] Druin, A., Bederson, B.B., and Quinn, A. (2009). Designing intergenerational mobile storytelling. *Proceedings of the 8th International Conference on Interaction Design and Children, ACM New York, NY, USA* 325-328.

[5] Ejersbo, L et al. (2008). Balancing product design and theoretical insights. In A. E. Kelly, R. A. Lesh, J. Y. Baek (Eds.), *The handbook of design research methods in education: Innovations in science, technology, engineering and mathematics learning and teaching.* New York: Routledge.

[6] Goodyear, P. (2011). Emerging Methodological Challenges. Book chapter in Markauskaite, Freebody & Irwin (Eds), *Methodological Choice and Design*, Volume 9, Part 4, pp. 253-266. Springer Netherlands

[7] Lombardo, V. & Damiano, R. (2012). Storytelling on mobile devices for cultural heritage. *New Review of Hypermedia and Multimedia*. p. 11-35.

[8] Milrad, M., Wong, L.-H., Sharples, M., Hwang, G.-J., Looi, C.-K., Ogata, H. (submitted) Seamless Learning: An International Perspective on Next Generation Technology Enhanced Learning. Under review for Z. L. Berge & L. Y. Muilenburg (eds.) *Handbook of Mobile Learning*. New York: Routledge.

[9] Nordmark, S., & Milrad, M. (2012). Mobile Digital Storytelling for Promoting Creative Collaborative Learning. *Proceedings of the Seventh IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education*, WMUTE 2012, Takamatsu, Japan, March 27th-30th 2012.

[10] Nordmark, S., Frolunde, L. & Milrad, M. (2010). Promoting the Creation of Digital Stories by Teacher Students: A Pilot Project on Curriculum Development in a Swedish Teacher Education Program. In D. Gibson & B. Dodge (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010* (pp. 1153-1160). Chesapeake, VA: AACE

[11] Nack, F. (2010). Add to the real, *IEEE Multimedia*, Volume 17, Issue 1.

[12] Pachler, N., Bachmair B., & Cook, J. (2010). *Mobile learning; Structures, agency, practices.* New York, Dordrecht, Heidelberg, London: Springer.

[13] Read, J.C. (2008). Validating the Fun Toolkit: an instrument for measuring children's opinion of technology. *Cognition*, *Technology and Work*, **10**(2): p. 119-128.

[14] Wong, L.H. & Looi, C.K. (2011). What seams do we remove in mobile assisted seamless learning? A critical review of the literature. *Computers & Education*, Volume 57, Issue 4, pp. 2364-2381.

The Functions of Smart Classroom in Smart Learning Age

Ronghuai Huang[a*], Yongbin Hu[b] ,Junfeng Yang[c]& Guangde Xiao[d]
[a]*R&D Center for Knowledge Engineering, Beijing Normal University, China*
[b]*R&D Center for Knowledge Engineering, Beijing Normal University, China*
[c]*R&D Center for Knowledge Engineering, Beijing Normal University, China*
[d]*R&D Center for Knowledge Engineering, Beijing Normal University, China*
* huangrh@bnu.edu.cn

Abstract: Reconstructing classroom and creating new type of classroom for students and teachers is an inevitable trend, and integrating sensor technology, artificial intelligence, rich media technology and communication technology into classroom is an inevitable choice. We propose a SMART model of smart classroom which characterized by showing, manageable, accessible, interactive and testing. We analyze the characteristics of three typical smart classrooms: "high definition" smart classroom, "deep experience" smart classroom and "rich interactivity" Smart classroom.

Keywords: smart learning environment; smart classroom; SMART model

1. Instruction

With the development of ICT in education in China, the majority of the teachers in class transform from the original "blackboard + chalk" teaching mode to the "computer + projection" teaching mode. But the teaching reform only remain in teaching-show status, the current teaching mode has not changed (Huang, 2011). If we consider from the classroom environment perspective, we could say there are some main predicaments in the nowaday's classroom environment.

First, the serialized presentations hinder students' understanding of the content. Second, learning material is intermittent displayed, which separates the backward and forward linkages of the teaching content. Third, multimedia console is fixed In the front of the classroom, which limits the teachers' performance. Fourth, a unified and fixed seating layout, is not conducive to a variety of teaching activities. Fifth, the equipment of the net classroom can't meet the needs of the students to explore.Sixth, a certain gap exits between the electronic whiteboard teaching application and the expected deeply interactive teaching.

The classroom is a physical environment which should provide support for the implementation of the curriculum. In the information age, new kind of classroom should be effective to present teaching contents, convenient to acquire learning resources, able to promote classroom interaction, with contextual awareness and environmental management, which may be called smat classroom.

2. Concept Model of Smart Classroom

A smart classroom relates to the optimization of teaching content presentation, convenient access of learning resources, deeply interactivity of teaching and learning, contextual awareness and detection, classroom layout and management etc.. It may be summarized as Showing, Manageable, Accessible, Real-time Interactive and Testing, which nicknams "S.M.A.R.T". The five dimensions just embody the wisdom of a smart classroom feature, which can be referred to as "SMART" concept model, as shown in figure 1.

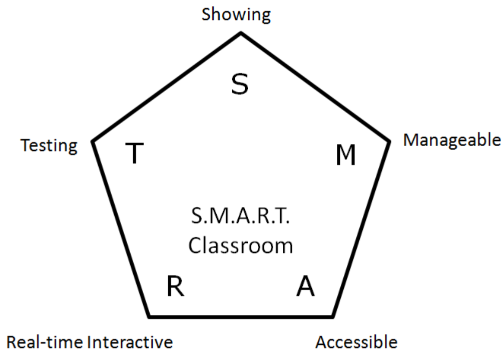


Figure 1 concept model smart classroom

2.1 Showing

Showing dimension represents teaching information presentation capabilities of the classroom, which requires not only showing the contents can be clearly visible, but also showing content suitable for learners 'cognitive characteristics, to the learning process to enhance the learners' learning materialsunderstanding and processing. The existing research shows that: multi-screen display can reduce the cognitive load and improve learners' achievement, Colvin (2007) pointed out that the multi-screen better than single-screen in the improvement of learners' achievements; Chen Changsheng (2011) developed a theory of dual channel dual teaching platform, and put forward a twin-track teaching mode. Displaying the teaching content by multiple-screen, smart classroom can effectively overcome students' thinking discontinuity problem which causes by the single screen. The Smart classroom can be equipped with light sensors to judge the intensity of the ambient light intelligently, even can automatically control the curtains opening and closing and opening and the intensity of lighting, which maintaiing the brightness of the screen is suitable.

2.2 Manageable

Manageable dimension represents diverse layouts and the convenience of management of the Smart classroom. The equipment, systems, resources of Smart classroom should be easy managed, including layout of the classroom management, equipment management, physical environment management, electrical safety management, network management etc. In terms of seating arrangement, rows are the typical environment for a teacher-centered classroom and/or individual learning. However, this layout severely limits the interaction between students, resulting in the students' passive learning (Su Hong, 2003). The layout of the Smart classroom should be flexible, diverse, supporting a variety of teaching and learning activities, give full consideration to the placement of various devices to improve the space utilization efficiency, the design of desks and chairs should consider the material, structure, color and other factors. Furthermore, the desks and chairs must be applied ergonomic principles consistent with adolescents' body scale (Sun Shanshan, 2011).

2.3 Accessible

Accessible dimension represents convenience of resources acquisition and equipment access in the Smart classroom, which involves resource selection, content distribution and access speed. Chen Shijian (2003) pointed out that the rich network of learning resources is conducive to independent learning, interactive cooperative learning, personalized learning, the implementation of educational socialization. In the selection of resources, the Smart classroom should be able to provide a wealth of teaching resources to support teaching and learning activities, computers, tablet PCs, smart phones, PDAs, wireless projectors, interactive whiteboards and other equipment can be flexible to facilitate access and supportinteract with the resource in the teaching process, the operation and re-generated. In content distribution, curriculum, lesson plans, teaching content, teaching tools should be able to facilitate distributed learning terminal. The speed of resource access and terminal accesss should not affect the teaching and learning activities.

2.4 Real-time Interactive

Real-time Interactive dimension represents the ability to support the teaching interaction and human-computer interaction of the Smart classroom, whichi involves convenient operation, smooth interaction and interactive tracking. In convenient operation, the Smart classroom should be able to support the man-machine natural interaction, interactive equipment and interface with a simple, full-featured, clear navigation, consistent with the operating habits and characteristics, touch, visual and voice interaction can improve the mouseman-machine interactive experience of the keyboard, the interaction tends to be more natural. In smooth interaction, the Smart classroom hardware should meet the interactive needs of the multi-terminal, and a large amount of data. In interactive tracking, smart classroom shoud record and store the basic data among teacher student and computer, so as to support the decision-making of teachers and students' self-assessment. The Horizon Report predicts that learning analysis technology will be popularized in 4-5 years (Johnson, 2011). The smart classroom should support teaching interaction comprehensively. Keep abreast of the topic of student interaction, students' difficulties and problems and to guide or help them would be smart classrooms' important function. Record interactive process and timely analyze the data obtained is the main way to achieve this function.

2.5 Testing

Testing dimension represents perception of the physical environment and learning behavior in Smart classroom. The physical environment factors,including air, temperature, light, sound, color, odor etc, affect the physical and mental activities of teachers and students (Li Bingde, 1991). With he development and popularization of sensor technology, a variety of sensors can be used in Smart classroom to detect indoor noise, light, temperature, odor and other parameters timely, , automatically adjusts the blinds, lamps, air conditioning, fresh air system equipment in the light of default ideal parameters,which maintain sound, light, temperature, air regulator suitable for students' physical and mental health status in the classroom.

3. Three Typical Smart Classroom

According to 'SMART' model, Managable and Testing dimensions are the common requirements of the smart classroom. Managable dimension requires Smart classroom should be achieved for all equipment, system, resource monitoring and management. Testing dimension includes two aspects. On the one hand, by monitoring the indoor air, temperature, light, sound, color, odor and other factors, Testing become reasonable. On the other hand, teachers can use classroom recording and broadcasting system records the teaching process and the use portable computing devices to record interactive process and to monitor learning outcomes, thus completing the tracking of the learning process.

The tradictional classroom, originally supports imparting knowledge, unable to meet the actual needs of the classroom teaching nowdays. The redesign of traditional classrooms, multimedia classroom and networked classrooms is urgently needed.

We propose three types of smart classroom, whici is high definition smart classroom, deep experience smart classroom and rich interactivity smart classroom.

"High definition" smart classroom is mainly used in transfer-accept teaching mode, the background of the teaching model is directly related to meaningful to accept learning theory which proposed by David Ausubel's, a well-known American educational psychologist. The theory suggests that the learning of students happens by acceptance learning, rather than discovery learning. In other words, the students master the knowledge and experience of their predecessors.by teachers' instruction and presentation. However, this acceptance learning should be meaningful rather than mechanical. "Deep experience" smart classroom is is mainly used in inquiry teaching mode. The mode is under the guidance of teachers in the teaching process, students explore co-operation to autonomy, characterized by learning teaching content knowledge, independent learning, group cooperation and exchange in-depth inquiry and, in order to achieve the requirements of a teaching model curriculum standards on cognitive goals and emotional goals.

Strong interaction smart classroom is mainly used in the collaborative learning mode. The group's collaborative learning is a group activities as the main teaching activities, mutual co-operation between the students, the interaction of teaching and learning activities which the power source .

4. Conclusion

The study of smart classroom equipment and its application model is an inevitable requirement for the development of educational information to a higher stage. This study if of great significant for the promotion of the digital school construction, elimination of the difficulties faced by the multi-media teaching, enhancing of students' innovative ability. The ultimate objective of the study is to change learning and teaching methods in the information age.

References

[1] Huang Ronghuai (2011). ICT in Education Promotes Current Educational Change: Challenges and Opportunities [J]. Chinese educational technology, (1):36-40.

[2] Colvin, J., Tobler, N., & Anderson, J.A. (2007). Productivity and multi-screen computer displays. Rocky Mountain Communication Review, 2(1), 31–53.

[3] Chen Changsheng, Liu Sannvya, Wang Hong,&Chen Zengzhao (2011). Dual track teaching modal based on dual coding theory [J]. China Education Info, (3):52-55.

[4] Su Hong (2003). Basic education curriculum reform and school cultural reconstruction [J]. Curriculum, Teaching Material and Method, (7):10-14.

[5] Sun Shanshan (2011). On school desks and chairs design [J]. Art Design, (8):134-135.

[6] Chen Shijian, &Wang Chong (2003). The meaning, function and type of e-learning resources in the internet [J]. E-education Research, (10):50-54.

[7] Johnson, L., S. Adams and K. Haywood (2011). The NMC Horizon Report: 2011 K[EB/OL].

[2012-2-20] http://wp.nmc.org/horizon2011/.
[8] Li Bingde (1991).Theory of instruction [M]. Beijing: People's Education Press:266-267.

Development of teaching material in tablet PC based on computer graphics by quantum chemistry calculation - Walden's inversion

Akira IKUO [a*], Nozomi NISHITANI [a], Yusuke YOSHINAGA [a] & Haruo OGAWA [a]
[a] *Department of Chemistry, Tokyo Gakugei University, JAPAN*
* ikuo@u-gakugei.ac.jp

Abstract: We developed computer graphics (CG) teaching material (TM) for university student, concerned about reaction with drastic change of the structure of reactants in following reaction as a model of Walden's inversion; $OH^- + CH_3Cl \rightarrow CH_3OH + Cl^-$. The CG-TM could demonstrate the nature of the reaction by space filling model, ball-and-stick model, and the reaction profile that can provide image of energy change during the reaction. The CG-TM enabled to load with note PC, tablet PC, and smart phone. Practice to the first year or the second year student of Tokyo Gakugei University was conducted. Surveys revealed that the CG-TM in tablet PC was sufficiently effective to provide information about the nature of the Walden's inversion.

Keywords: Teaching material, CG, Visualization, Tablet computer, Quantum chemistry calculation, Walden's inversion, Structure change, Space filling model

Introduction

Visualization of computer graphics (CG) gives us great help to realize not only images of molecular properties but also those of molecular behavior on dynamical reaction mechanism. It is our aim to produce CG teaching material (CG-TM), which provides realizable images of the nature of chemical reaction [1]. Walden's inversion is one of typical reactions in organic chemistry, and the reaction is often adopted in TM on the curriculum of the university, including some appropriate schemes [2]. The schemes should be developed for student to acquire more realizable images of the nature of the reaction. We developed CG-TM for university student, concerned about reaction with drastic change of the structure of reactants in the reaction (Eq. 1) as a model of Walden's inversion.

$$OH^- + CH_3Cl \rightarrow CH_3OH + Cl^- \qquad \text{(Eq. 1)}$$

Production of CG-TM by quantum chemistry calculation and practice of the CG-TM in tablet computer, which is more effective to provide image than that with PC and the projector [3], to the university student are reported.

1. Procedure

1.1 Quantum chemical calculation

The semi-empirical molecular orbital calculation software MOPAC with PM5 Hamiltonians [4] in CAChe Work System for Windows (ver. 6.01, FUJITSU, Inc.) was used to find the transition state and the reaction path of chemical reactions. The structure of the reactants on the transition state was searched by use of the optimized map with the PM5 energies (One Label) in MOPAC with the model described in above. The optimized structure of the transition state was verified by the observation of a single absorption peak in the imaginary number by the use of the program Force in MOPAC for vibration analysis. If the peak was observed, Intrinsic Reaction Coordinate (IRC) [5] calculation was done and the reaction path was confirmed.

1.2 Production of CG teaching material

A movie of the reaction path was produced by the software DIRECTOR (ver. 8.5J, Macromedia, Inc.) following the display of the bond order of the structure of the reactants in each reaction stage, which was drawn by the CAChe. It was confirmed that the Cast members were arranged on the stage and the molecular models of reactants moves smoothly. The ball was arranged on the reaction profile and the movement of the ball and the reactants was confirmed. The movie file was converted to the Quick Time movie by the Quick Time PRO (ver. 7.66, Apple, Inc.) and was saved to iPad (Apple, Inc.) and iPhone 4 (Apple, Inc.), by using the iTunes (Apple, Inc.).

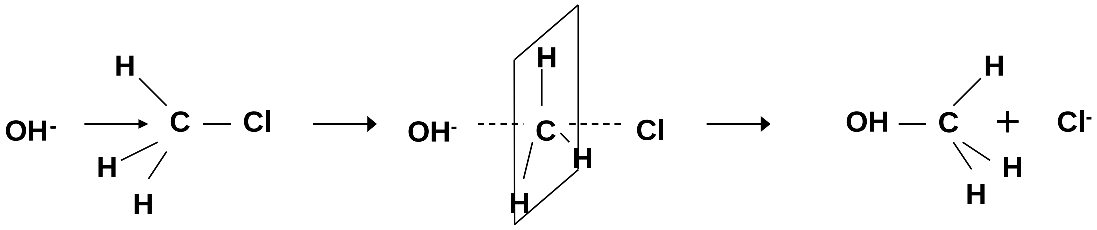
1.3 Practice of teaching material

Teaching material was practiced on three classes, the first year student of teacher training course for elementary school and the second year student of natural environmental science course, of “Chemistry laboratory” in the fall semester of the 2010 fiscal year at Tokyo Gakugei University. Total numbers of students were 103. All practice was carried out by one instructor and one teaching assistant. Teaching material used for the trial was the CG movie of ball –and-stick model shown by the tablet PC. Procedure of the practice is described as follows. Preliminary survey, which consisted questions to ask them to draw image of Walden’s inversion from formula and questions about information of students, were conducted. Then, usage of the tablet was explained by instructor with projection of the CG movie. After the explanation, 9 tablet PCs were distributed to students and asked them to watch CG movie in the tablet as shown in the figure 1. Students started to touch it and watched the movie over and over enthusiastically. One tablet PC was shared by two to four students depends on size of class. Time duration for usage of the tablet PC was about three minutes for each student. During this time period students were quiet which suggests that they were concentrated on the subject. After about 10 minutes, students were asked to start answer the posteriori survey that consisted question to ask them to draw image of Walden’s inversion from the movie. Some students were still watching the movie even when they were answering the questionnaire.

Figure 1 Practice of teaching material

2. Results and discussion

2.1 CG teaching material



Scheme 1 Images of Walden's inversion

The S_N2 reaction of chloromethane and hydroxide [2] is shown in the scheme 1. Carbon atom at center to which halogen attaches is attacked by the nucleophile, hydroxide, from a position away by 180 degrees and methyl alcohol forms. Therefore, the transition state was searched from the reactants where the bond angle of O-C-Cl was adjusted to 180 °. A single absorption peak in the negative region was found at -399.46 cm^{-1} from the vibrational analysis. This result indicates vibrational mode due to the decrease of potential energy for the direction of only one path via a true transition state at the saddle point.

The reaction path from the reactants to the products via the transition state was searched by the intrinsic reaction coordinate (IRC) calculation [5] in MOPAC. The inter-atomic distances, *d*, and the heat of formation obtained by the calculation are listed in the scheme 2 and in the table. These values were in good agreement with the literature values [6, 7]. Therefore, it was confirmed that the reaction path and the molecular configurations obtained by the calculation were appropriate.

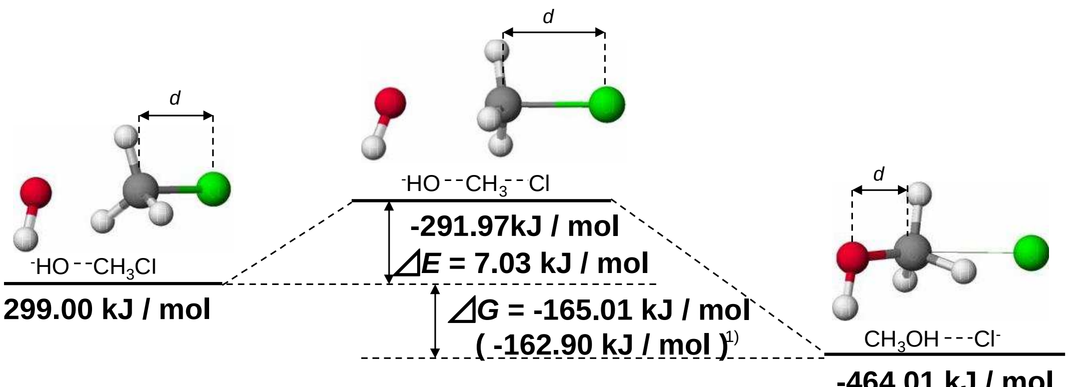


Table. Inter-atomic distance calculated with PM5

Sample	*d* / Å		
	Calculation	Experimental	Subtraction
$^-$HO-- CH_3Cl	1.87	1.78 [2)]	0.09
$^-$HO --CH_3Cl-- Cl^-	2.16		
CH_3OH --Cl^-	1.41	1.43 [3)]	-0.02

Scheme 2 The structures of the initial state, the transition state and the final state of formation methyl alcohol and energy

The literature values are indicated in the parentheses.

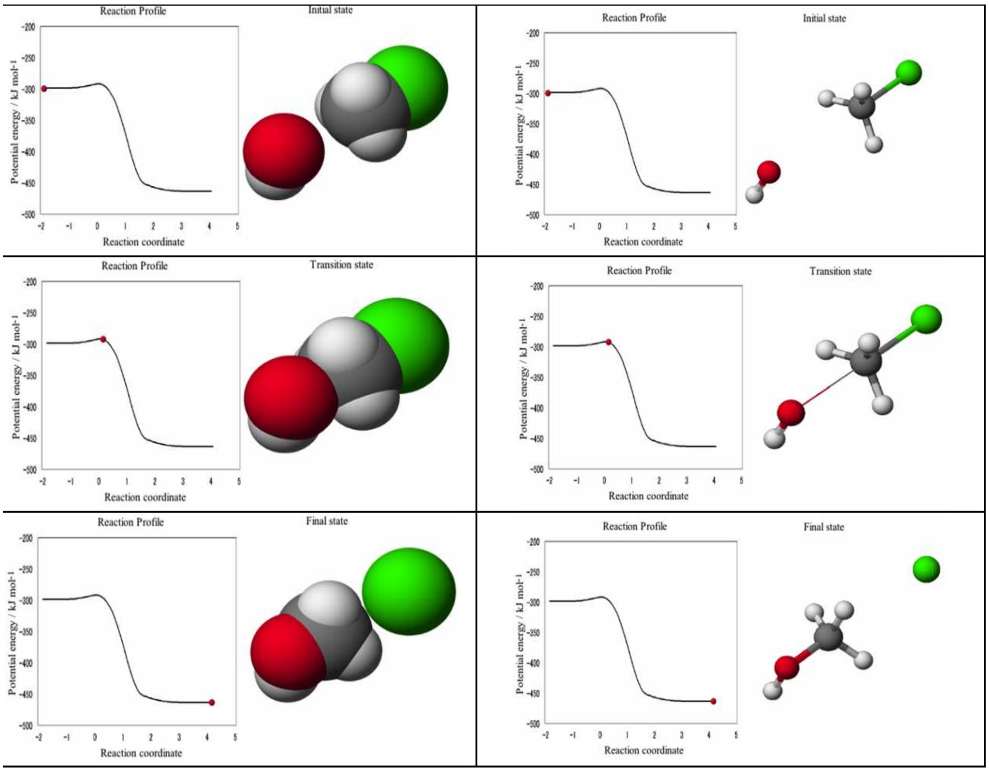
1) Nihonkagakukai (1984). *Kagakubinran kisohen* 3rd Ed., Maruzen, Ⅱ314; 2) Weast, R.C., (1982). *CRC Handbook of Chemistry and Physics* 63rd Ed., CRC Press, Inc., F180-181; 3) Shida, S., (1981). Kagakujiten, Morikitasyuppan, p.1251.

2.2 Teaching material

Selected picture of CG movies are shown in the figure 2. The CG shows the reaction profile, which demonstrates the degree of the reaction progress by the ball indicating potential energy vs. reaction, coordinate. Movies were made by using not only the space filling model which shows realistic shape but also the ball-and-stick model which shows change in molecular configuration easily. A student is expected to obtain the image of an umbrella reverse like motion in Walden's inversion. In the space filling, the existence probability of the electron is 90 %. In the ball-and-stick, the thickness of stick changes by bond order.

When the CG is touched by learner, the Quick Time control bar appears and the red ball can move by learner's choice. This manual control feature provides "Hands-on" feeling to learner. This CG teaching material could provide not only images of energy change during reaction but also images of dynamical structure change during chemical reaction.

Figure 2 Selected picture of CG movies
Reaction profile and image of reactants in space filling and ball-and-stick model



2.3 Evaluation

The result of the questionnaires was summarized. The answer judged to be able to acquire the image of Walden's inversion (the image to which an umbrella reverses) was follows; the image obtained from the reaction formula was 24% and from the CG teaching material was 51%. The number of CG teaching material was better than that of the reaction formula. Students were abele to obtain the image of Walden's inversion from the CG teaching material.

3. Conclusion

In this work, the change in the molecular configuration in Walden's inversion made visible from the quantum chemistry calculation. The CG teaching material enabled to load with note PC, tablet PC, and smart phone. Surveys revealed that the teaching material in tablet PC was sufficiently effective to provide information about the nature of the Walden's inversion.

Acknowledgements

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) (22500803).

References

[1] a: Ikuo, A., Ikarashi, Y., Shishido, T. & Ogawa, H. (2006). USER-FRIENDLY CG VISUALIZATION WITH ANIMATION OF CHEMICAL REACTION: ESTERIFICATION OF ACETIC ACID AND ETHYL ALCOHOL AND SURVEY OF TEXTBOOKS OF HIGH SCHOOL CHEMISTRY. *Journal of Science Education in Japan*, 30 (4), 210-215; b: Ikuo A., Nagashima H., Yoshinaga Y., & Ogawa H. (2009). Calculation of potential energy in the reaction of "$I + H_2 \rightarrow HI + H$", and its visualization. *The Chemical Education Journal (CEJ)*, Registration #13-2.

[2] For example: McMurry, J., (2001). *John McMurry Organic chemistry*. Tokyokagakudojin, p.272 (in Japanese).

[3] Ikuo, A., Nagashima, H., Yoshinaga, Y., & Ogawa, H., (2012). Development and Practice of Teaching Material in Tablet Computer Based on Computer Graphics by Quantum chemistry Calculation - Reaction of $I + H_2 \rightarrow HI + H$ -. In *Proceedings of the Seventh IEEE International Conference on Wireless, Mobile and Ubiquitous Technology in Education* (pp. 82-86). Takamatsu, Kagawa, Japan.

[4] a: Stewart, J. J. P. (1989). *J Comp. Chem*., 10, 209–220; b: Stewart, J. J. P. (1989). *J. Comp. Chem*., 10, 221–264; c: Stewart, J. J. P. (1991). *J. Comp. Chem*., 12, 320–341.

[5] Fukui, K. (1970). *J. Phys. Chem*., 74, 4161-4163.

[6] Nihonkagakukai (1984). *Kagakubinran kisohen* 3rd Ed., Maruzen, Ⅱ314.

[7] Weast, R. C., (1982). *CRC Handbook of Chemistry and Physics* 63rd Ed., CRC Press, Inc., F180-181.

Nurturing self-regulation by mathematical inquiry in a one-to-one TEL environment

Håkan SOLLERVALL
Halmstad University, Sweden
hakan.sollervall@hh.se

Abstract: We design and evaluate a curriculum-based mathematical learning activity involving secondary students' geometrical constructions, mathematical modeling and algebraic validation of hypotheses based on hands-on explorations with the interactive geometry application GeoGebra available on individual laptops. We argue that guided inquiries in a technology-enhanced learning environment that invites blending of interactive technologies and traditional resources may be an efficient means for developing self-regulatory skills.

Keywords: Self-regulation, inquiry-based learning, TEL, mathematics, geometry, algebra.

1. Introduction

We have designed and implemented a mathematical learning activity addressing coordinate geometry and specifically the distance formula. The activity involves guided inquiries and is intended for use in upper secondary school in Sweden, where mathematics teaching is dominated by solving routine textbook tasks by following prescribed rules. The 38 students in our study can be regarded as novices in the inquiry context. Besides the distance formula as a specific learning objective, the activity aims to develop their self-regulatory skills and inspire them and their teachers to pursue further development of these skills. The blending of one-to-one technologies and traditional resources offer invite the students to make meaningful mathematical connections between geometric and algebraic representations.

2. Theoretical background

Current research in psychology [1] suggests a relation between overt behavior and learning gains, where interactive-constructive behavior (exploring, investigating and generating in collaboration with peers) is more favorable than just being active in the classroom, for example, when solving routine problems in a textbook. However, inquiry-based learning [2] (as well as problem-based, learner-centered, discovery, experiential, and constructive learning) is often implemented with minimal guidance from the teacher and has in such cases been shown to be less effective for novice learners than guidance specifically designed to support the their cognitive processing [3]. The successful implementation of inquiry-based learning activities requires careful design considerations of guidance, in terms of embedded features of the activity and carefully planned strategies for teacher guidance and peer scaffolding during the implemented activity [5]. Inquiry-based learning challenges students' self-regulation regarding cognition, motivation, behavior, and context, in corresponding phases of self-regulation: forethought, planning, and activation (cognition); monitoring (motivation); control (behavior); reaction and reflection (context) [4].

3. The mathematical learning activity

In the first task the students work on a clear screen in order to stimulate attending only to geometric relations. The task concerns points equidistant to two given points: *The distance from the point to P should be equal to the distance from the point to Q.* They are instructed to place several such points on the screen (Fig. 1, left pane) and answer the question:

What geometric figure do you get from all the points that satisfy the condition?

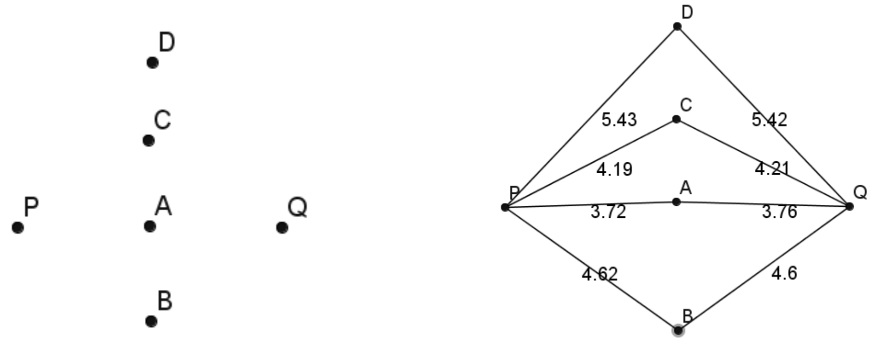


Figure 1: Constructions in GeoGebra according to the instructions for the first task.

Next, the students are invited to check the placement of their points by having GeoGebra measure the distances from each point to P and Q, respectively (Fig. 1, right pane).

The second task addresses the same geometric condition and the same question, but now in a coordinate system. The students are instructed to keep their constructs from the first task, show grid and axes, and place the points at P = (0,4) and Q = (2,0) respectively (Fig. 2, left pane). A correct reconstruction is illustrated in Figure 2 (right pane).

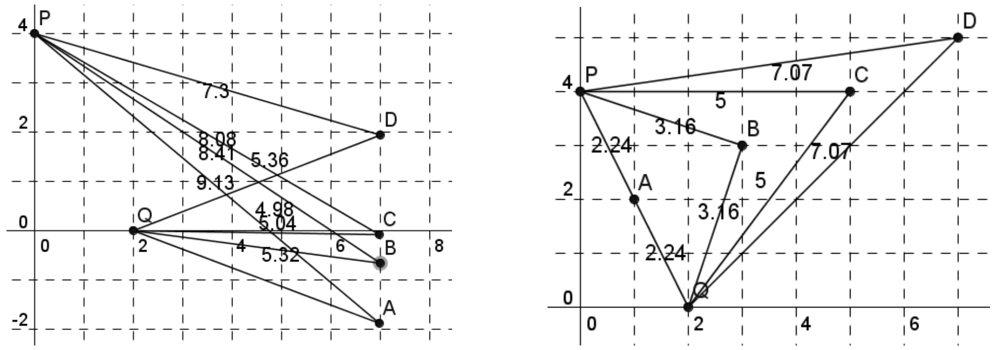


Figure 2: A possible initial setup (left pane) and a correct construction (right pane).

On the next page of instructions the students are asked to work with pen and paper to find the equation of the straight line.

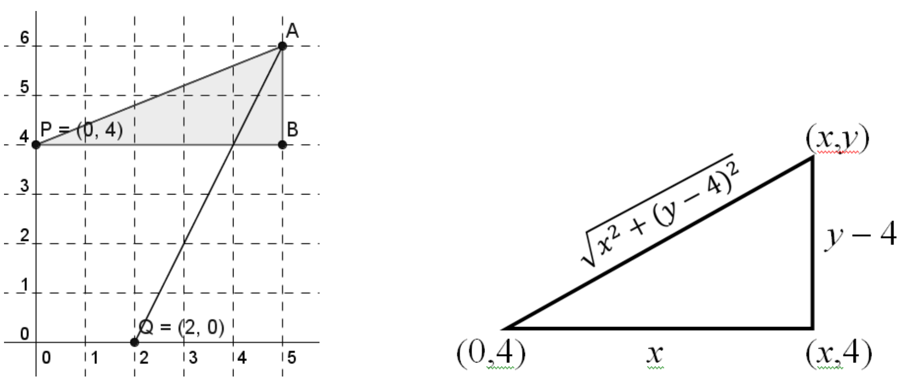


Figure 3: Continuation of task 2 in Geogebra (left pane) and on paper (right pane).

So far, the students have been acting on definitions by interpreting and representing them, generalizing from a few points to all points, justifying geometric constructs, and acting on a generalization of a visual representation by representing it algebraically. The continuation of the second task is more challenging. The students are asked to *prove algebraically that all the points (x,y) that satisfy the condition lie on a straight line*. They are instructed to implement the setup shown in Figure 3 (left pane).

They draw the triangle PAB in GeoGebra (Fig. 3, left pane) and work with pen and paper to find an algebraic expression for the hypotenuse (Fig. 3, right pane). They are also asked to find an expression for the distance between A and Q and simplify the equation:

$$\sqrt{(x-0)^2+(y-4)^2}=\sqrt{(x-2)^2+(y-0)^2} \quad \text{yields the equation} \quad y=\frac{1}{2}x+\frac{3}{2}$$

The students who finish readily recognize that the last equation defines a straight line.

4. Main results and concluding remarks

In task 1, several groups initially answered 'a cross' or 'a triangle', even after they had constructed the line segments that allowed them to measure distances. The picture (Fig. 1) was prioritized before the written condition in the students' reflections, which may be considered as a misinterpretation of the available contextual elements.

In the final part of task 2, the students made several algebraic mistakes such as 1) replacing the square of a binomial with the sum of the squares of the two terms and 2) canceling square roots and squares term by term. They did not activate any strategies for self-control and had to be told by the teacher to check their incorrect calculations.

Several students who got stuck while working with the tasks showed signs of lacking motivation. However, they recovered quickly when receiving content-oriented guidance.

In a classroom environment where students mainly solve problems in a textbook, most mistakes can be easily adjusted by checking answers or asking peers or the teacher. In such an environment, individual students are not responsible for monitoring, control, and reflection on their work. Mistakes do not propagate and do not affect future work, so they are not stimulated to develop strategies for self-regulation. In comparison, even a minor mistake in an inquiry can have fatal effect on its continuation and may cause the students to fail in achieving the intended learning objectives. Sensitively guided inquiries can serve to nurture not only mathematical learning but also self-regulation. The 38 participating students in our study show good promise, they just need to be offered more frequent opportunities to engage in carefully designed and meaningful mathematical inquiries.

References

[1] Chi, M. (2011). Differentiating four levels of engagement with learning materials: The ICAP hypothesis. *Presentation at the 19th International Conference on Computers in Education*. Chiang Mai, Thailand.

[2] Edelson, D., Gordin, D., & Pea, R. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *The Journal of the Learning Sciences*, *8*(3&4), 391-450.

[3] Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.

[4] Schunk, D. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist, 40*, 85-94.

[5] Wong, L-S., Looi, C.-K., Boticki, I., & Sun, J. (2011). Improving the scaffolds of a mobile-assisted Chinese character forming game with the SCAPE framework. *Proceedings of the 19th International Conference on Computers in Education.* Chiang Mai, Thailand.

Examination of Effective Information Presentation Using an AR Textbook

Norio SETOZAKI[a*], Tsutomu IWASAKI[b] & Yusuke MORITA[a]
[a]*Faculty of Human Sciences, Waseda University, Japan*
[b]*CYBERNET SYSTEMS CO., LTD., Japan*
*setozaki@aoni.waseda.jp

Abstract: In this study, an AR textbook for astronomy education using a tablet device was developed and three of its content elements were evaluated. The results of all the subjective evaluation questions of the AR CG Model Content were assessed as positive. The free description results clarified the content's improvement needs.

Keywords: Augmented Reality, Information presentation, Mobile learning

Introduction

Various fields of education have studied Augmented Reality (AR), which can synthetically present virtual objects in real environments [1]. Research has also examined the practical use of AR equipment such as AR textbooks that overlay the lecture video onto the paper textbook [2] and the effect of information presentation by AR equipment [3]. However, considering the preparation and installation of equipment, AR equipment using a PC connected to a head mounted display and a web camera is not yet practical for use in schools.

In contrast, tablet device usage is increasing, and its practical application in education is expected [4]. AR technology became comparatively easy to use with the attachment of the camera to the tablet device. This facilitates effective learning by viewing and listening to digital content without requiring the students to be aware of computer mediation in contrast to using the AR PC equipment. In this study, an AR textbook using a tablet device was developed for astronomy education and three of its content elements were evaluated to identify improvements needed for useful AR textbooks.

1. Procedure

1.1 AR Textbook for Astronomy Education

The outline of the "AR Textbook for Astronomy Education" is shown in Figure 1. This textbook was developed using an AR browser for a tablet device (junaio) and contains three types of content: Video Content, AR Video Content, and AR Computer Graphics (CG) Model Content. These contents are displayed when the camera attached to the tablet device recognizes each page of a paper textbook.

Video Content displays a lecture video on the tablet's full screen. Once the camera recognizes the image marker on a paper textbook, the lecture video is played to the end. Therefore, the camera need not continuously monitored the paper textbook.

AR Video Content allows the lecture video to be overlaid onto a paper textbook. The lecture video is attached to a 3D virtual object as a video texture and therefore, the paper textbook must always be continuously monitored by the camera.

AR CG Model Content allows CG models to be overlaid onto a paper textbook. The 3D model used in this study displayed the animation of the earth's rotation and the moon's orbit. Sunlight's shadow effect was overlaid on the 3D model of the earth and the moon. The paper textbook must be continuously monitored by the camera as for AR Video Content.

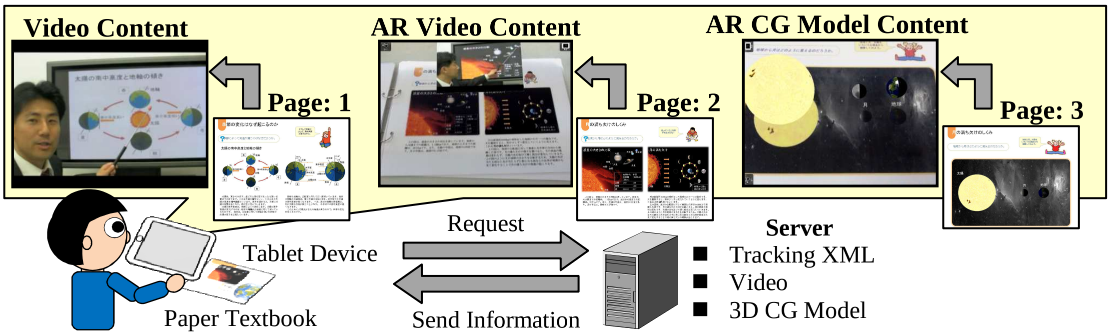


Figure 1: AR Textbook for Astronomy Education

1.2 Subjective Assessment by Survey

85 undergraduate university students participated in the survey. After viewing all three types of content, the students responded to statements in five question categories: Interest, Understanding, Motivation, Usefulness, and Obviousness. The students responded by selecting from the responses: Strongly agree, Agree, Disagree, and Strongly disagree. The positive and negative responses were totaled for each item and were compared using Fisher's exact test. Students also wrote free description responses about the improvement needs of each content type. The free description replies were classified by improvement type within content category and totaled.

2. Results and Discussion

Table 1 provides the analysis results of Fisher's exact test. Table 2 presents the results of the responses suggesting needed content improvements.

For Video Content, the highest number of negative answers were given to the question items on Obviousness. The free description comment analysis in Table 2 indicates the two most frequent responses relating to the negative rating. Of these, the highest number of participants suggested improving the lecture content so that students do not get bored. The next highest number of participants found the lecture video's image quality poor. We assume that students had difficulty reading the characters displayed on the lecture video slides. Other improvement needs noted, in descending order of total comments, are: Additional functions, Sound quality, How to use a tablet device, Download time, and Video play time.

For AR Video Content, the question items on Obviousness also received the most negative evaluations. More than half of the recommendations for improvements were about the "Stability of the display." In AR Video Content, the camera attached to the tablet device needs to continuously monitor the paper textbook, therefore, as compared with Video Content, the display of the lecture video was unstable. The paper textbook's composition needs to be enhanced with an image marker to increase the recognition performance. Other improvement needs noted, in descending order of total comments, are: Image quality, Additional functions, How to use a tablet device, Instruction, Contents of the lecture, Sound quality, and Video play time.

The results of the subjective evaluation revealed that the AR CG Model Content is the most useful. However, it too needs improvements in the Obviousness category for more useful content in three major areas: Additional functions, 3D model quality, and Stability of the display. Other improvement needs noted, in descending order of total comments, are: How to present the content, How to use a tablet device, and Instruction.

Table 1: Results of subjective assessment

Categories	Video Content			AR Video Content			AR CG Model Content		
	Positive	Negative	Fisher's exact test	Positive	Negative	Fisher's exact test	Positive	Negative	Fisher's exact test
Interest	77	8	**	82	3	**	84	1	**
Understanding	70	15	**	56	29	**	81	4	**
Motivation	72	13	**	72	13	**	83	2	**
Usefulness	75	10	**	65	20	**	81	4	**
Obviousness	52	33	†	19	66	**	77	8	**

**: p < .01, *: p < .05, † : .05 < p <.10

Table 2: Improvement needs for each content type

Video Content		AR Video Content		AR CG Model Content	
Category	Number of answers	Category	Number of answers	Category	Number of answers
Contents of the lecture	15	Stability of the display	34	Additional functions	10
Image quality	13	Video Size	10	3D model quality	10
Additional functions	5	Image quality	5	Stability of the display	9
Sound quality	5	Additional functions	4	How to present the content	4
How to use a tablet device	4	How to use a tablet device	4	How to use a tablet device	2
Download time	2	Instruction	2	Instruction	2
Video play time	1	Contents of the lecture	1	**Total**	**37**
Total	**45**	Sound quality	1		
		Video play time	1		
		Total	**62**		

3. Conclusion

In this study, an AR textbook for astronomy education was developed and three of its content elements were evaluated. Among the content types, the subjective evaluation rated the AR CG Model Content the most highly positive in all the question categories. The free description results clarified the content's improvement needs, suggesting the specific factors of each category that need the greatest improvement.

Future research will focus on improving AR textbooks based on the data obtained in this study. Examining effective practical uses of AR textbooks in elementary schools or junior high schools will be another future focus area.

References

[1] Yuen, S., Yaoyuneyong, G., Johonson, E. (2011). Augmented Reality: An Overview and Five Directions for AR in Education. *Journal of Educational Technology Development and Exchange*, 4(1), 119–140.

[2] Morita, Y., Fujishima, H., Setozak, N., Iwasaki, T. (2011). Development and Evaluation of an AR Textbook for Astronomy Education. *Japan Jounal of Educational Technology*, 35 (Suppl.), 81–84.

[3] Setozaki, N., Iwasaki, T., Morita, Y. (2011). Effectiveness of Information Presentation using Augmented Reality Teaching Equipment for Statue Appreciation, *Proceedings of the 19th International Conference on Computers in Education*, 522–524.

[4] Savilla, B. (2010). Integrating the iPod Touch in K-12 Education: Visions and Vices. *Computers in the Schools*, 27(2), 121–131.

Exploration of the influence of instructional material annotation on "teaching and learning" by teachers and students

Kuo-Hung CHAO, Kuo-En CHANG, Yao-Ting SUNG, Stefan CHAO
PhD. Student, National Taiwan Normal University, Taiwan
Professor, National Taiwan Normal University, Taiwan
Professor, National Taiwan Normal University, Taiwan
Instructor, Taoyuan Innovation Institute of Technology ,Taiwan
stefanckh@gmail.com

Abstract: The grasp of main ideas as a learning strategy and activity can have various benefits on learners in different instructional objectives, such as reading comprehension, memory, and discerning the main points of reading content. A necessary condition for grasping main ideas is determining the importance of sentences [2]. Thus, this study attempts to explore the interaction of a reciprocal teaching digital annotation automatic matching system with teachers and learners, and how it affects learning activities. This study hopes to assist teachers in using computer systems to quickly assess how student groups are marking important points, and uses the computer to automatically evaluate the similarity between class markings and teacher markings, so that the teacher can have a general idea of how students have different markings from the teacher. This is used to understand students' learning conditions and the distribution of their markings for later instructional discussion of important points, so that learners can receive better learning guidance in a limited period of time. The concept of the system in this study can allow the teachers and teaching assistants to provide timely assistance to learners in the teaching and learning process, and provide suitable support, with the objective of "internalizing" the development of learner ability development. The whole system concept proceeds in a digital learning environment, and serve as a learning motivation for learners and a direction for instructors to improve their teaching, so that they can achieve the purpose of offering individualized teaching.

Keywords: digital annotation, reciprocal teaching, instructional reflection

1. Introduction

Annotation is a common support behavior in learning. Based on the "meta cognition" theory and "reciprocal teaching" in psychology, this study attempts to explore the benefits of annotation on personal learning and the instructional process for teachers. In addition, it seeks to understand the current annotation systems and how they might be improved in the context of a mature environment of electronic instructional materials.

- *Changes in learning environment and methods – electronic textbooks*
- *Paper digitization – electronic annotations*
- *Theoretical basis – metacognition and learning*
- *System prototype – reciprocal teaching*

Most systems for marking main points focus on individual learners, given them simple feedback of right and wrong, but there is no convenient method to quickly understand the distribution and ratios of incorrect or correct readings of instructional materials for all the learners, which would make it inconvenient for teachers to carry out instructional discussion and have a reference for instructional strategies. Thus, the objective of this study is to explore a digital annotation automatic matching system designed based on reciprocal teaching, and whether this has a positive or negative effect on interaction between teachers and students in the process of instruction.

2. Research Structure

2.1 Connection between annotation and learning process

In sum, annotations can serve to remind one of the important points, assist in memory, and record personal reflections. The ability of annotation and finding main points directly affects the quality and quantity of memory. Too much information that is too complex is also detrimental to memory. They can have good long-term memory after they accept signals and have effectively absorbed it.

2.2 Connection between annotations and instructors

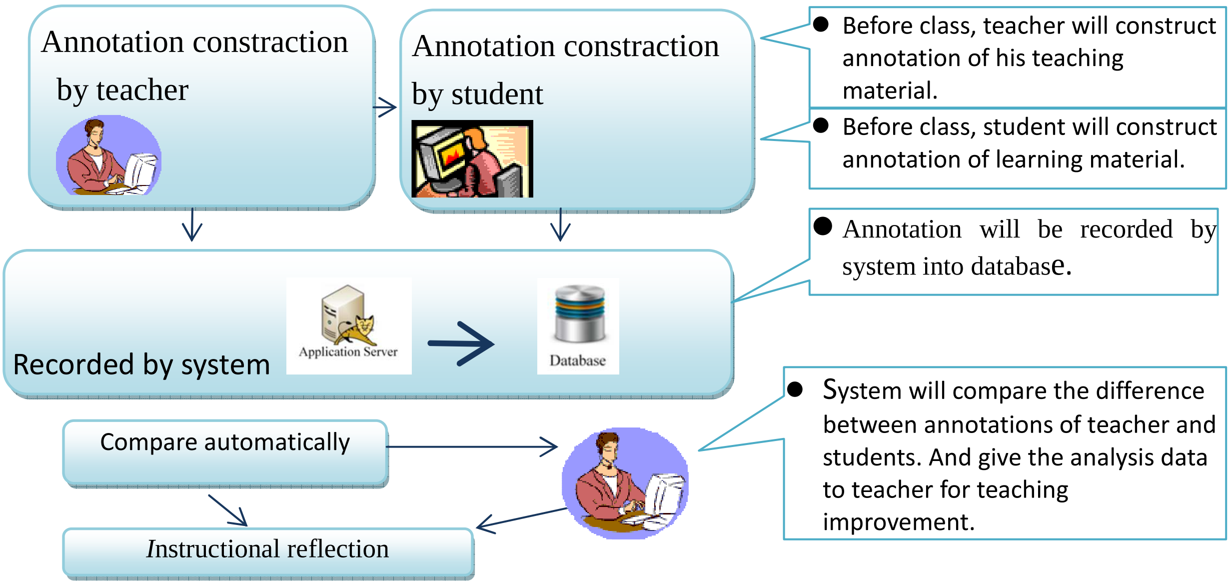
Instructors can use the main points made by standard markers as standards, and use the automatic matching system to compare student annotations one by one to calculate the length of overlaps. After observing the distribution and the ratios, teachers can find specific annotators based on instructional priority for specific annotations, and engage in question instructional discussion and view all annotations and overall meaning, in order to modify instructional progress and depth, and to use instructional reflection to achieve the objective of customized instruction. It is hoped that this can give students the opportunity to learn self-supervision and thinking, and cultivate student techniques in meta cognition.

3. Research Method

This study mixes qualitative research and quantitative research, differing in the subjects and stages. In terms of teacher instruction, the first stage is an attitude questionnaire survey, collecting information about the types of subjects taught by teachers, the length of their teaching experience, whether they use annotations, and whether they agree with modifying instructional models based on students. The second stage carries out actual work on the system after the automatic annotation system is complete. After the course concludes, another attitude questionnaire survey will be administered, as well as individual in-depth interviews, which is used to understand whether this method benefits instruction and whether the system meets requirements in instruction.

A blended study of qualitative and quantitative is simultaneously conducted for students. This study evaluates whether the annotation system gave them learning assistance, whether this type of instruction and learning produce positive or negative effects, and the psychological effects produced.

4. System prototype



5. Conclusions and Suggestions

This study designs an online reciprocal teaching annotation system to help students read electronic and digital instructional materials in real time. It is further hoped that this can help students to use system assistance tools to guide students in their active understanding of the meaning of instructional materials to enhance reading comprehension abilities, while giving students the opportunity to learn self-supervision and thinking, and cultivate student techniques in meta cognition. In addition, the teachers can track the learning process to obtain student behavioral data to understand which difficulties they are encountering in reading, as well as the support of which systems would enhance their reading comprehension in the process of prediction, clarification, posing questions, and finding important points.

References

[1]. Ovsiannikov, I. A., Arbib, M. A., & Mcneill, T. H. (1999). Annotation technology. International Journal of Human-Computer Studies, 50, 329-362.
[2]. Rich, P., & Tripp, T. (2010). Choosing the right video annotation tool for the job: A conceptual framework. *Proceedings of Society for Information Technology & Teacher Education International Conference 2010*, 1171-1178.
[3]. Wolfe, J. L. &Neuwirth, C. M. (2001). From the margins to the center: The future of annotation. *Journal of Business and Technical Communication*, 15, 333-371.
[4]. (15)Yao-ting Song, Hui-tseHou, Guo-an Chang. Incorporating Information Technology in Instruction: Borrowing from American Experiences, Reflecting on Taiwanese Development. *Bulletin of Educational Research*, Vol.51,pp. 31-62, 2005.

Training of Differential Diagnosis of Dissociative Disorders with Virtual Reality Simulations. Effectiveness and Usability.

Jose GUTIERREZ-MALDONADO*, Ángel AGUILAR-ALONSO & Marta FERRER-GARCIA
University of Barcelona, Spain
*jgutierrezm@ub.edu

Abstract: A training program of differential diagnosis skills was developed to enhance the learning of the psychopathological exploration of Dissociative Disorders using Virtual Reality (VR) based simulations. The sample of the study consisted of 60 psychology students (University of Barcelona). Comparisons between the experimental (VR training) and the control group (traditional role-playing training) showed that students trained with the VR system obtained better scores than students trained with the traditional method. The usability was assessed with the *Software Usability Measurement Inventory* (SUMI). Results suggested that simulated interviews are a friendly and motivating tool to train diagnostic abilities in psychology students.

Keywords: diagnostic interview, psychology, psychopathological exploration, Dissociative disorders, virtual reality, artificial intelligence, simulations.

Introduction

Virtual Reality (VR) is widely used to train health-care professionals [1]. This technology provides trainees simulations of real life situations where they can learn by doing in a safe educational context. A virtual environment allows students to change their point of view by observing a phenomenon from different perspectives, enhancing the recall of objects and their localization [2,3]. Similarly, students can play different roles in a social virtual environment which enhances their social skills. The main purpose of this project was to develop a new method to improve the acquisition of psychopathological exploration skills, in psychology students, for the diagnosis of Dissociative Disorders [4], by developing a computerized resource based on VR.

Methods

Sample

Sixty under-graduated students participated in the study. Mean age was 21.30 (SD= 2.20) and most of them were female (71.3%).

Instruments

A software based on virtual reality and artificial intelligence, *Simulated Interviews*, was developed to enhance skills learning. Virtual environments were developed with 3D studio. Agents were modelled with Poser and Character Studio. Interaction and navigation were programmed with Virtools Dev. Voices were directly recorded from actors. The knowledge base resulted from a data base matching question classes and answer classes according to the DSM-IV hierarchical system of differential diagnosis (Diagnostic and Statistical Manual of Mental Disorders) [4] for the diagnostic group considered for the training (Dissociative Disorders). The system included 3D models with realistic textures and illumination, and avatars that play the role of virtual patients whose facial expression matches the verbal contents according to the psychopathology simulated (Dissociative Disorders). A diagnostic interview skills test was also used to evaluate the psychopathological exploration and differential diagnostic skills acquired. The final score was calculated taking into account the correct answers converted on a 10 point- scale. The usability was assessed with the Software Usability Measurement Inventory (SUMI) [5]. The SUMI is a 50-item questionnaire that measures five aspects of user satisfaction: affect (do users like the program), helpfulness, learnability, efficiency, and control.

Procedure

Participants were randomly assigned to one of the following conditions: 30 students who received psychopathological exploration skills training using the simulated interviews (experimental group) and 30 students who received psychopathological exploration skills training using the traditional method of role playing (control group).Two professors were available for each group.

Students in the experimental group were requested to assist to the laboratory for two consecutive sessions of 50 minutes with a ten minutes pause. Every student received from the professor in charge a basic explanation of the main characteristics of the Dissociative Disorders. Later, students had to interact with the virtual simulations of four patients who displayed the disorders. The same procedure was applied to the students in the control group, however, instead of interacting with the simulated interviews training program they received a traditional training based on role playing in which the professor played the role of the patients and every student had to perform the interviews to identify the correct diagnosis. Finally, the effects of the training program on the students learning was compared to the traditional teaching method, thus, students in both groups were required to do the diagnostic interview skills test. Students in the experimental group also assessed the usability of the software with SUMI inventory.

Results and Discussion

Effectiveness

After confirming the homogeneity of both groups in age and gender, a t-test for independent samples was conducted to evaluate the differences between the scores obtained by students in the experimental and control group in the diagnostic interview skills test. Students who were trained with the simulated interviews obtained better scores than the students trained with the traditional method of role-playing ($t=3,89$, $p< 0,01$). The group that received the training program by means of simulated interviews reached higher scores in the diagnostic interview skills test than the group that received traditional training based on role playing. This difference was statistically significant. Given that both groups were homogenous in age and gender, the differences found in the final scores obtained in the diagnostic interview skills test can be attributed to the different training programs applied to the experimental and control group.

Usability

Only SUMI's items that were applicable to our software were considered for data analysis. We selected 15 items listed in Table 1. Participants showed a good level of general satisfaction with the application. Most participants (90%) would recommend it to their colleagues.

Table 1. Answer frequencies

	Frequency (%)		
	Agree	Undecided	Disagree
2. I would recommend this software to my colleagues	90	8,3	1,7
3. The instructions and prompts are helpful	91,7	8,3	0
5. Learning to operate this software initially is full of problems	0	5	95
7. I enjoy my sessions with this software	80	18,3	1,7
12. Working with this software is satisfying	81,7	16,7	1,7
13. The way that system information is presented is clear and understandable	95	5	0
17. Working with this software is mentally stimulating	75	20	5
19. I feel in command of this software when I am using it	75	15	10
26. Tasks can be performed in a straightforward manner using this software	78,3	16.7	5
27. Using this software is frustrating	5	11,7	83,3
29. The speed of this software is fast enough	68,3	26,7	5
32. There have been times in using this software when I have felt quite tense	15	8,3	76,7
42. The software has very attractive presentation	68,3	23,3	8,3
44. It is relatively easy to move from one part of a task to another	75	16,7	8,3
48. It is easy to see at a glance what the options are at each stage	61,7	28,3	10

References

[1] F. Mantovani, G. Castelnuovo, A. Gaggioli, & G. Riva (2003). Virtual reality training for health-care professionals, *CyberPsychology & Behavior*, *6*, 389-395.

[2] Dalgarno, B., Bennett, S. & Harper, B. (2010). The Importance of Active Exploration, Optical Flow, and Task Alignment for Spatial Learning in Desktop 3D Environments. *Human–Computer Interaction, 25(1)*, 25-66. doi:10.1080/07370020903586670

[3] Lee, K. M., & Jung, Y. (2005). Evolutionary nature of virtual experience. *Journal of Cultural and Evolutionary Psychology, 3*, 159-178.

[4] American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders.* Masson

[5] J. Kirakowski & M. Corbett (1993). SUMI: The software measurement inventory, *British Journal of Educational Technology, 24*, 210-212.

An Audience Response System with iGoogle Gadgets Using Mobile Devices

Manuel CAEIRO-RODRÍGUEZ, Juan GONZÁLEZ-TATO & Martín LLAMAS-NISTAL
Department of Telematics Engineering, University of Vigo, Spain
Manuel.Caeiro@det.uvigo.es

Abstract: Currently, Audience Response Systems (ARS) are integrated widespread on face-to-face classes to improve the attendance and engagement of the students. In this paper, we present the current trend of using Web 2.0 to provide ARSs and our efforts to create a complete ARS over the iGoogle platform. We created a dedicated gadget to be used by the students as the traditional remote control system, which can be operated from any device with a Web browser (e.g. laptops, tablets or smartphones). With our approach, teachers can use gadgets to compose and make in-class polls or questionnaires and check the responses on-the-fly.

Keywords: Audience Response Systems, iGoogle, gadget, Web 2.0

Introduction

Audience Response Systems (ARSs) allow teachers to inquire learners during the course of classroom activity electronically. The basic operation of these systems involves the distribution of remote controls to the learners. Then, teachers may ask questions and learners provide their answers using the remote devices. Usually questions are shown to the students using a beamer. Learners' responses are collected by a system that process the information and offers a summary of the results. This system is usually accessible from the teachers' computer. This way, the teacher can get the results immediately and can vary his/her discourse or the classroom activities accordingly.

We present in this document our work for supporting an ARS over iGoogle [1]. We have developed a set of gadgets that provide ARS functionalities to users. Teachers may create and activate polls from his/her personal iGoogle page and students may answer, from their personal iGoogle page, using the device they want as remote control (e.g. laptop, tablet or smartphone).

1. Pi2E Project

1.1 Phylosophy

The Pi2E Project [2] aim is to provide an e-learning system via iGoogle [1]. iGoogle allows users to compose their own personalized home-pages, using gadgets and info sources that they organize in the screen as they prefer. Users have also available the well-known Google Web services, such as Google Search and Google Talk. The Pi2E approach is that the user can make use of these applications and gadgets to create his/her own e-learning environment. In addition to the general-purpose gadgets we have developed some e-learning specific gadgets to facilitate certain tasks.

We chose iGoogle home-page system because it offers a free service to any user with a Google account. It also offers its own set of gadgets, but anyone can add new gadgets from any page or repository (e.g. ROLE widget store [3]). In addition, users can also take advantage of the several presentations flavours available in iGoogle. Gadgets can be offered through smartphones and digital TVs. When the personal home page is accessed from a smartphone it is shown in a different way than usual to adapt to the reduced space. In this case we provide a multi-device and ubiquitous service easily.

The Pi2E specific e-learning gadgets that we have been developing provide functionalities related with the management of educational activities. We consider that learning plans involve a set of activities of different types that have to be carried out. For example: lectures, assignments, questionnaires, debates, etc. Therefore, we have taken some gadgets already available in iGoogle and developed some new gadgets to provide functionalities supporting these activities and mainly the management of these activities. Each gadget provides just a specific functionality and the whole e-learning environment can be achieved through the composition of several gadgets. Some of these gadgets may require a Web service that provides the contents or some storage capability. Therefore, we have also developed some Web services that support the provision of gadget functionalities.

Up to date, we have developed four specific gadgets for Pi2E [2,4]: Edu-GAM, Edu-GAAT, Edu-GAL and Edu-GAR. These gadgets are provided for different kind of users within the educational process: author of learning activities and lesson plans, teacher and student. Next sections provide a brief introduction to these gadgets while the rest of the paper focuses on Edu-GAR.

2. Edu-GAR

Edu-GAR (Educational – Gadget Audience Response) is the latest gadget we have created in Pi2E. Its purpose is to emulate the operation of a remote device in an ARS. In other words, it just enables students to answer questions that are issued by a teacher during the class. Other Pi2E gadgets support this functionality. Edu-GAAT must be used by the teacher to create the questions. Edu-GAM must be used also by the teacher to activate the questions and to view the (statistical) results.

1.1. Edu-GAR in Action

Edu-GAR was designed to be used by students during the class. It was developed using the Edu-GAL gadget as a reference, but in a simpler way. The teacher can activate or deactivate questions. Questions are shown to the students when they are active and then students can answer questions.

Unlike conventional remote controls, Edu-GAR involves a proactive behaviour. While remote controls just send answers to the receiver when the student push the button, Edu-GAR must periodically poll the server to check if a question is available. Then, the gadget shows to the student the question with the possible answers and waits until the student provides his/her result. Then, the student answer is sent to the server.

Edu-GAR can display not just single questions but complete questionnaires with several questions. Indeed, Edu-GAR supports IMS QTI [5] question types.

3. Conclusions

The ARS solution implemented in Pi2E gadgets differs from traditional ARS systems in two main ways. First, students can answer the questions using several devices. Moreover, they are shown the questions and answers in their personal device and not just in the classroom screen. Second, student answers can be recorded for future analysis. This way, the teacher has a new instrument to assess his/her students.

In addition, Edu-GAR also involves some other advantages. Traditional ARS systems just allow users to answer one question each time. The question is a multi-choice type, and it requires the user to choose one option out from several available ones. Using Edu-GAR the teacher can include several questions in a row and several question types are available.

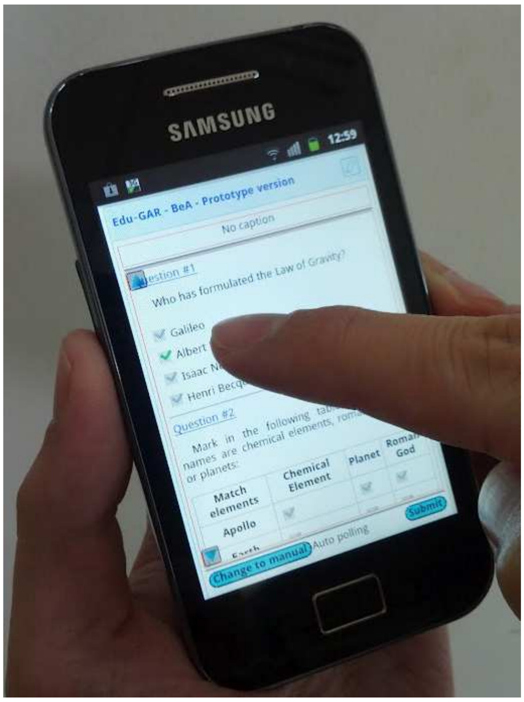


Figure 1. Student using Edu-GAR on a smartphone

Edu-GAR shares the advantages of Web-based ARS systems. The service cost is reduced, more devices can be used to answer the questions and students are less likely to forget their remote devices at home.

There are some aspects that need to be taking into account using this technology. As a Web-based system, it depends on the availability of a good Internet connection. In case the connection is damaged the users may not receive the questions or they answers may not be saved. In addition, the system has latency. Therefore, it is not suitable for fastest finger test types. It is also important to note that a student can answer the questions from anywhere, not just from the classroom. If this can be a critical issue, then some mechanism could be implemented, such as location filtering.

Acknowledgements

We thank Galician Consellería de Economía e Industria for its partial support to this work under grant "*PI2E: Páxina Inicial I-Google de E-learning*" (10SEC002E). We also thank Spanish Ministry of Science and Innovation for its partial support to this work under grant "*Methodologies, Architectures and Standards for adaptative and accessible e-learning (Adapt2Learn)"* (TIN2010-21735-C02-01).

References

[1] iGoogle. Retrieved from http://www.google.com/ig Visited: 2012/05/07

[2] Llamas-Nistal, M.; Caeiro-Rodríguez, M.; González-Tato, J. and Álvarez-Osuna, J. "Towards a Collection of Gadgets for an iGoogle e-learning platform". EDUCON 2012 Marrakesh, Marocco 17-20 April 2012

[3] ROLE widget store. Retrieved from http://www.role-widgetstore.eu Visited: 2012/05/07

[4] González-Tato, J.; Llamas-Nistal, M., Caeiro-Rodríguez, M. and Álvarez-Osuna, J.. "Implementando un PLE con gadgets de iGoogle". ICEM-SIIE 2011 "Old meets new – media in education". ISBN: 978-972-789-347-8

[5] IMS QTI Retrieved from http://imsglobal.org/question/ Visited: 2012/05/07

Evaluation of a Mobile Plant-Identification System to Support the Study of Vegetation Succession

Haruka INOUE[a*], Fusako KUSUNOKI[b], Yoshiaki TAKEDA[a], Etsuji YAMAGUCHI[a], Shigenori INAGAKI[a], Hiroshi MIZOGUCHI[c] & Masanori SUGIMOTO[d]

[a]*Kobe University, JAPAN*
[b]*Tama Art University, JAPAN*
[c]*Tokyo University of Science, JAPAN*
[d]*University of Tokyo, JAPAN*
* hareni.uku.ao@gmail.com

Abstract: The authors developed a smartphone-based *Pocket Plant Guide* to support the outdoor learning of vegetation succession, and evaluated the utility of this device by administering a questionnaire to 35 elementary school children. The results indicated that the *Pocket Plant Guide* not only functions as a supportive learning tool, but also offers a user-friendly interface.

Keywords: Mobile system, outdoor learning, vegetation succession

Introduction

The authors have developed the *Pocket Plant Guide*, a smartphone-based, mobile system for plant identification [1] [2]. The *Pocket Plant Guide* provides accessible information about six indicator plants for outdoor learning into vegetation succession. These plants have been selected on the basis of findings from research into vegetation succession conducted in the Rokkozan region of Hyogo Prefecture in Japan. A digital game using a similar group of indicator plants has previously been developed, in order to enhance vegetation succession studies (The Vegetation Game: An Vegetation Succession Game). However, as this game is situated within a virtual world, it has encountered problems in terms of real-world relevance. The *Pocket Plant Guide* addresses these problems, by providing information that focuses specifically on the six indicator plants in outdoor learning, and thus complements the information obtained from the game.

In this study, we evaluated the *Pocket Plant Guide*, concentrating on the ease of use of its interface and its effectiveness as a learning support, by administering a questionnaire to elementary school children.

1. The *Pocket Plant Guide*

The *Pocket Plant Guide* is designed to operate on an iPhone/iPod touch. It provides graphic and verbal information about the characteristics of leaves for the six indicator plants used in the vegetation succession game, so that the user can identify these plants in the field. Figure 1 shows the home screen of the *Pocket Plant Guide*. If you touch the icon of a particular tree, such as the *konara* (a species of oak indigenous to Japan), a black-and-white image of the corresponding leaf will be displayed. When you touch the "see characteristics" button in the bottom right corner, the screen will switch to an image with comments, as Figure 2 illustrates. The "select plants" button in the bottom right corner of this next screen returns the user to the home screen.

Figure 1: Home Screen

Figure 2: Image with Comment

2. Evaluation of the *Pocket Plant Guide*

2.1 Method

Subjects: 35 sixth-grade elementary school children (aged between 11 and 12)
Use of the *Pocket Plant Guide*: The students first played the vegetation succession game, and then spent half a day outdoors, searching for plants featured in the game. The *Pocket Plant Guide* was used for reference during the outdoor learning.
Tasks and procedures: We set two tasks. The first focused on plant identification. In order to investigate the extent to which the system influences the subject's knowledge of plants, we showed the subjects images of the leaves of the six plants featured in the *Pocket Plant Guide*, before and after usage of the system, and elicited the name and characteristics of each plant. The second task focused on user evaluation. In order to evaluate the accessibility of the user interface and the benefit of knowledge provided by the system, we asked the subjects to complete a questionnaire after they used the system. The questionnaire required them to respond to 17 statements with one of the following options: "I strongly agree," "I somewhat agree," "I somewhat disagree," and "I strongly disagree." Both tasks were performed using questionnaires, which took around 15 minutes to complete.
Timing: The evaluation took place on October 22, 2011.

2.2 Results

2.2.1 Plant Identification Task

The responses to the task of identifying and providing characteristics for the plants featured in the *Pocket Plant Guide* before and after usage of the system were evaluated according to the Wilcoxon signed-rank test. The results showed that the number of times a plant was correctly identified and described increased significantly after implementation of the system (median of before 1.0, median of after 11.5; Z=-5.24; $p < .01$).

2.2.2 User Evaluation Task

Table 1 illustrates the trends discovered in the subjects' responses to each question. We examined the trends by first separating the four choices into two groups, consisting of positive and negative responses, and then calculating the deviation using the Fisher statistical significance test. The results showed that there were significantly more positive responses than negative responses to 16 out of the 17 statements ($p < .01$). There was no significant difference in responses to Statement 13.

Table 1: Evaluation of the *Pocket Plant Guide*

	TS	STS	DQTS	DTS
01: It was fun to use the *Pocket Plant Guide*.**	25	10	0	0
02: I was absorbed while using the *Pocket Plant Guide*.**	14	19	2	0
03: The *Pocket Plant Guide*’s screen was easy to see.**	32	3	0	0
04: It was good to have buttons at the bottom.**	26	8	1	0
05: The buttons were of the right size.**	24	8	3	0
06: It was easy to switch from the screen to select plants to the screen where the image of the plant’s leaf is shown.**	33	2	0	0
07: It was easy to switch from the screen that showed an image of the plant’s leaf to the screen that showed the image with comments.**	34	1	0	0
08: It was easy to return to the screen to select plants from the screen that showed an image with comments.**	25	0	0	0
09: The screen showing the image of the plant’s leaf was of the right size.**	23	8	4	0
10: The screen showing the image of the plant’s leaf with comments was of the right size.**	22	10	3	0
11: I looked at the screen showing the image of the plant’s leaf repeatedly.**	33	2	0	0
12: I looked at the image of the plant’s leaf with comments screen repeatedly.**	21	13	0	1
13: It was easy to find the leaf of the plant I was looking for.	2	13	15	5
14: I compared what was described on screen showing the image of the plant’s leaf with comments with the real leaf of the plant.**	19	13	3	0
15: I compared what I observed on the real plant with what was shown on the *Pocket Plant Guide*. **	21	9	5	0
16: The *Pocket Plant Guide* was useful in outdoor learning. **	28	5	2	0
17: I want to investigate other forests using the *Pocket Plant Guide*. **	26	8	1	0

N=35. ***p* <.01 TS: Think so, STS: Somewhat think so, DQTS: Don’t quite think so, DTS: Don’t think so

3. Conclusion

The results demonstrate that the *Pocket Plant Guide* is an effective aid for students in identifying indicator plants for vegetation succession studies. The *Pocket Plant Guide* has been evaluated positively because it is fun to use and easy to operate, and is an effective support in outdoor learning. However, the responses to the statement “It was easy to find the leaf of the plant I was looking for” did not show a positive tendency. In this regard, it is necessary to improve both the information content and the interface provided by the *Pocket Plant Guide*. Additionally, it is important to consider the importance of teacher intervention in learning activities that make use of such a system.

Acknowledgements

This research was supported in part by the Grants-in-Aid for Scientific Research (B) (No. 23300303).

References

[1] Kusunoki, F., Inagaki, S., Yamaguchi, E., & Takeda, Y. (2011). Design of an environmental learning system using the vegetation game and smartphones. *Proceedings of Kinki Regional Conference of Society of Japan Science Teaching 2011*, 72. [in Japanese]

[2] Inoue, H., Kusunoki, F., Takeda, Y., Yamaguchi, E., & Inagaki, S. (2012). Evaluation of an environmental learning system using the vegetation game and smartphones. *JSiSE Research Report*, 26(6), 137-140. [in Japanese]

Using Activity Tree Structure to Edit Materials of the Multi-Touch eBook

Yuan-Hsun Liao[a*], Hsiao-Hui Li[b], Po-Jen Cheng[a], Ming-Hsiang Su[a], Cheng-Yu Tsai[a] & Pao-Ta Yu[a]

[a]*Department of Computer Science and Information Engineer, National Chung Cheng University, Taiwan, ROC*

[b]*Department of Information Management, National Chung Cheng University, Taiwan, ROC*

Abstract: A large number of learning systems have applied touch elements to fulfill their user experience. eBook, another well-known e-Learning material, has been developed quickly and plays an important role nowadays. Nevertheless, there are not any learning systems designed for editing an e-Book including multi-touch contents. This motivates us to find a novel approach to design hierarchical teaching materials of digital learning content and e-Book. Due to the multi-touch technology and activity tree structure applied in this paper, this proposed approach can significantly enhance the variety of teaching strategy and interactivity of students. The designed approach consists of two parts: editor and player. About the editor, instruction designers can load multimedia content on a haptic tree node, insert search pattern to the node, and then output it as a project. About the player, learners can load the project and then touch every node on the tree.

Keywords: eBook, multi-touch, material design, activity tree

Introduction

The touch technology has been applied on many devices such as smartphones and computer screens in the recent years. Touch screens emerged from academic or corporate research laboratories in the second half of the 1960s[1]. Actually, we have been using touch screens for a long time, such as Automatic Teller Machine (ATM) machines, retail settings, tourist attractions, and museums. There are some common technologies including resistive, surface capacitive, projected capacitive, infrared, surface acoustic wave (SAW), and optical imaging[2].

Based on the activity tree structure, we design an integration system to combine the concept of learning activity from e-Learning with the touch technology into a user experience environment. With the player, learners can interact with learning objects and achieve better learning effect by directly manipulating.

1. Material Design with Activity Tree

Only based on the conventional teaching materials including images and videos, most of teachers cannot do vivid teaching in the classroom. This leads to the situation in which their teaching progress falls behind the course progress. Although the problem has been solved in the interactive learning environment which learners can involve in situated learning, another problem is that the thoughtless learners or novices concentrate on the unimportant part of learning object or control casually the course progress. We extend the activity tree structure to translate PowerPoint's content as a vivid package of multi-touch material [3]. In Fig. 1, an example is shown to present this Computers material structure created by this paper method.

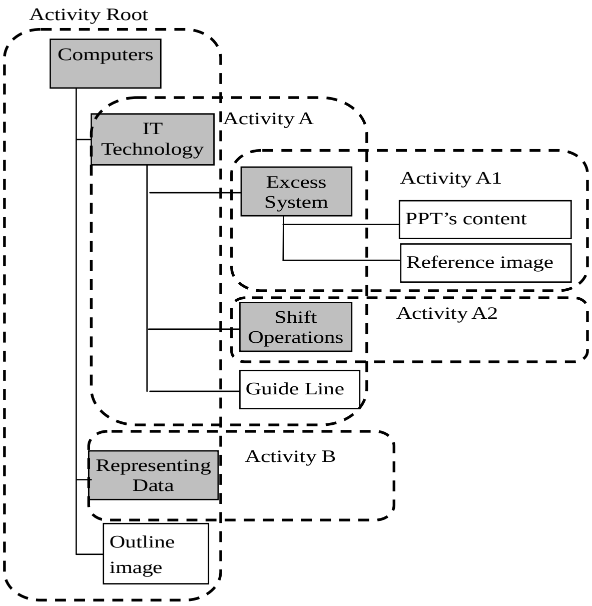


Fig. 1. A computer material structure created according to activity tree.

2. Editor Design

Multimedia learning has better learning effect than conventional method, but it is complicated to produce interactive learning objects. When facing numerous learning materials including videos, images and audios, our system can arrange these learning objects into a project or e-Book conveniently and then display it on a multi-touch screen for learners to manipulate and watch.

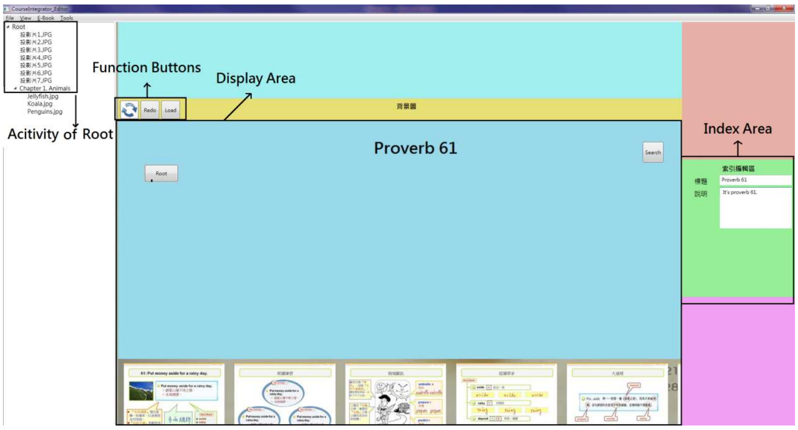


Fig. 2. User interface of root node display manager.

3. Player Design

The interface designs the multi-touch system to fit Natural User Interface (NUI) [4] with intuitive appearance and friendly interaction. Without using small display element and saving screen space for users to manipulate, the entire screen is a surface containing only two areas which are a manipulation area and a multimedia list area. The user interface of a sample project is shown in Figure 3 which is at branch node as well as a surface. To deserve to be mentioned, all the nodes on the surface are touchable.

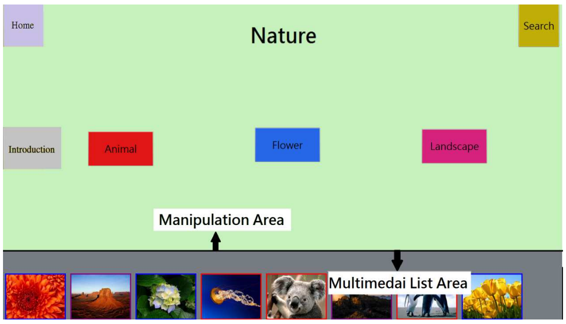


Fig. 3. User interface of a sample project.

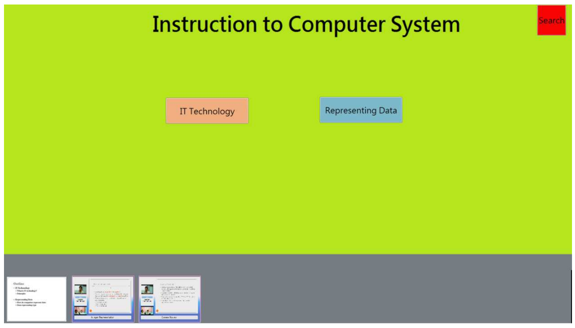


Fig. 4. Activity root of the computer material.

4. A New Design Concept for Teaching Material

We use the Editor to produce the package of multi-touch material according to book structure. In Fig. 4, this is activity root including cover, outline and guidance video of the course. The activity root provides leaf node button as activity A or B or more session of the course to link course activity for teacher teaching.

5. Conclusion and future work

By using screens supporting multi-touch, teachers can exhibit the touchable content either for student or for teachers themselves to manipulate the item, and students are easier to focus on their lessons and teachers are easier to interact with students because teachers have a lot of ways to display the teaching material. This paper has achieved the effect of learning such as interaction between teachers and students, easy navigation, and demonstration teaching method [5-7].

References

[1] Shneiderman, B. (1991). Touch screens now offer compelling uses. *IEEE Transaction on Software*, *8*, 93-94.
[2] Bhalla, M.R., & Bhalla A.V. (2002). Comparative Study of Various Touchscreen Technologies. *International Journal of Computer Applications*, *6*, 12-18.
[3] IMS Global Learning Consortium, I. IMS Simple Sequencing Best Practice and Implementation Guide. 2012, Available from: http://www.imsglobal.org/simplesequencing/ssv1p0/imsss_bestv1p0.html.
[4] Petersen, N. & Stricker D. (2009). Continuous natural user interface: Reducing the gap between real and digital world. *IEEE International Symposium on Mixed and Augmented Reality*, 23-26.
[5] Xingfu, D. (2001). Microcosmic theories of distance education. *China Distance Education*, *2*, 11-14.
[6] Croney, J.C., Jameson D.M., & Learmonth R.P. (2001). Fluorescence spectroscopy in biochemistry: teaching basic principles with visual demonstrations. *Biochemistry and Molecular Biology Education*, *29*, 60-65.
[7] Zimmerman, B.J., & Jaffe A. (1977). Teaching through demonstration: The effects of structuring, imitation, and age. *Journal of Educational Psychology*, *69*, 773-778.

Development of Supporting System for Nature Observation and Investigation Activities Around Users Using Smartphones

Yukihiko SATOU, Hisaharu TANAKA, Kenzi WATANABE & Yasuhisa OKAZAKI
Graduate School of Science and Engineering, Saga University, Japan
*yukihiko@ai.is.saga-u.ac.jp

Abstract: In this research, we develop a supporting system for nature observation and investigation activities around users using smartphones. By using smartphones equipped with GPS and a camera for nature observation, users can record the target pictures and position information to observe. Users can share many data by uploading these records with the others. Moreover, it is useful to data map creation of observation information by performing continuous use. We are developing the system which users enable to learn about natural environment by observation and investigation activities of dragonflies, being familiar with the nature around users using smartphones equipped with GPS. Since dragonflies are one of the index insects of waterside environment, users enable to learn about natural environment through observation and investigation activities of dragonflies. We aim at visualization of changes of natural environment by performing natural observational-research activities using this system.

Keywords: Smartphone, iPhone, Nature Observation, Position Information, GPS

Introduction

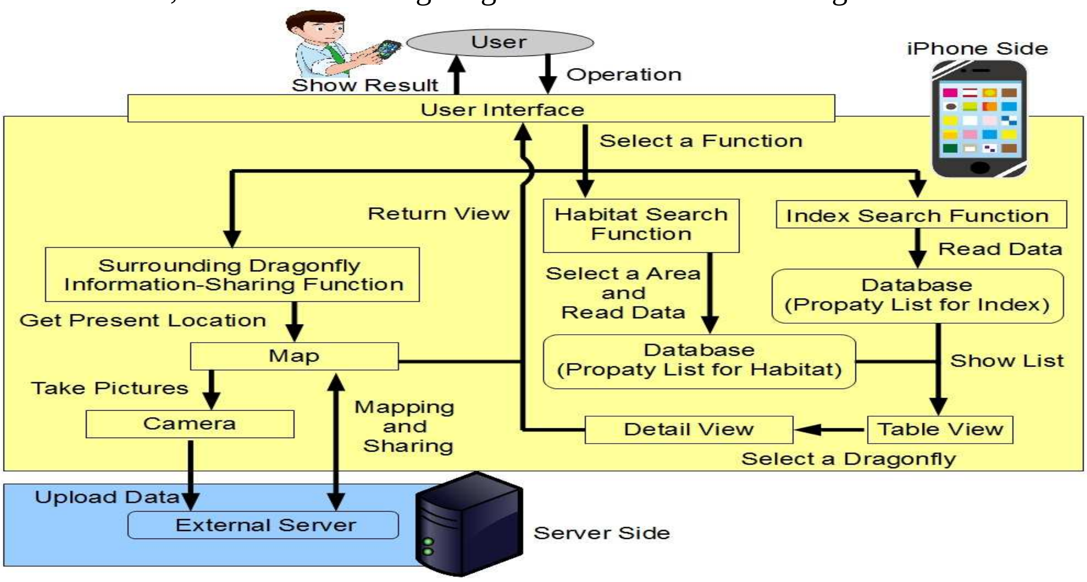
In recent years, smartphones and cellular phones are used for outdoor activities by many researches. For example, There are researches with GPS of smartphones, the camera and information sharing. The feature of the researches using the GPS function and the camera is to add information to a map, and other feature is offering learning support which is different for each location[1][2]. In addition, there are researches of information sharing of the outdoor activities by communication with mobile phones and servers[3][4]. We develop a new useful system for environmental learningby combining these advantages.
In this research, we utilize smartphones equipped with the camera and the GPS function taking advantage of the function of such a mobile terminal. Additionally, we develop the natural-environmental-learning supporting system through experiential activity in the outdoors about the natural environment of surrounding areas.

Dragonflies are one of indexes of waterside environment in fieldwork. For this reason, we develop the learning supporting system which enables to learn natural environment around users by observation and investigation activities of dragonflies using smartphones equipped with GPS. We have already developed and open an information system called a “Dragonfly Kingdom Saga”. This system is a system developed for PC. Therefore, it is not easy to use this system in outdoor activities.

So, we adopted iPhone as a terminal, and developed a new information system of dragonflies. As field learning new functions, we are developing a record function and a share function for the information on user's surrounding dragonflies. Users can investigate the dragonfly's feature and ecology information easily on that spot. Also, users can upload the information on a position and time to the server, when users take pictures of dragonfly by a camera. Users can refer to the uploaded information as an information map from a terminal. When users perform outdoor activities, users enable to understand the past life list for dragonflies at a glance. Moreover, when users continue such activity, it is possible to put the information in a database. Therefore, this system is able to create an ecology information map. We aim at visualization of changes of natural environment using this system.

1. System architecture

Figure 1 shows architecture of our system and behavior. This system uses fundamental UI and GPS function of iPhone, and the Exif information on a photograph. Exif information specifies the file format for recording a graphics file. Therefore, we can describe camera information in a graphics file, when we take a picture. Now, We have developed three search functions. They are a “dragonfly's index search function”, a “dragonfly's habitat search function”, and a “surrounding dragonflies information-sharing function”.



1: Flowchart of system

2. System functions

2.1 Index search function

The index search function can display the names of dragonflies on a table, and can search detailed information of dragonflies with the order of the Japanese syllabary from the name of the dragonfly.

By using this function, users enable to investigate the photograph of dragonflies, the classification, the form, the ecology, and the origin of name. In observation and investigation activities for dragonflies, when the name of the dragonfly is known, we are able to search quickly by slide a table, inputting a name into a search bar, or choosing the initial of the name of the dragonfly.

2.2 Habitat search function

In habitat search function, users can specify a habitat and retrieve the information on the dragonfly which inhabits the place. When a habitat is pinpointed, we look the table view of only the dragonfly which lives in the habitat. This function also has Index search function. In observation and investigation activities for dragonflies, by this function, users enable to choose a dragonfly from a habitat and enable to do Japanese syllabary search.

2.3 Surrounding dragonflies information-sharing function

The surrounding dragonflies information-sharing function is a function in which dragonfly information is sharable, by taking pictures of dragonflies and uploading to the server. When users take pictures, users can add the filming date, position information and so on to photograph by Exif. Since this function maps on the map based on the position information, users enable to investigate what kind of dragonflies lives around users. Users can upload photographs and dragonflies information, and users can share with retrieving the information on dragonflies from the map with other users. For this reason, In observation and investigation activities for dragonflies, users enable to know information of dragonflies.

3. Conclusions

In this research, we have developed the supporting system for nature observation and investigation activities around users using smartphones. In this system, users can investigate the feature and ecology information on dragonflies easily on that spot. Additionally, users can take pictures of dragonflies using a camera and upload to a server with the information on a position and time. By using the uploaded information, users can create the ecology map of dragonflies. When users perform outdoor activities, they enable to understand the past life list at a glance and offer useful information in the case of observation. In addition, when users continue such activity, it is possible to put the information in a database. Therefore, this system is able to create an ecology information map. We are going to conduct an evaluation experiment after developing the function which supports identification of dragonflies from the feature of the body of dragonflies.

References

[1] Miyata, H., Ishigami, M., Sannomiya, M., & Matsushita, E., (2010). Development and Evaluation of the Flower Identification Database for Mobile with a Geo-tagged Picture Map, Proceedings of the 18th International Conference on Computers in Education, Putrajaya, Malaysia, 344-351.

[2] Ogino, T., Suzuki, M., Kusunoki, F., & Hatono, I., (2008). Cellular Phone with GPS Based Learning Support System for Field Studies, Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008, Vienna, Austria, 1812-1817.

[3] Takenaka, M., Inagaki, S., Kuroda, H., Deguchi, A., & Ohkubo, M., (2006). Fieldwork support system using mobile phones: Evaluations of information sharing in the second grade's Life Environment Study, Proceedings of ED-MEDIA 2006, Orlando, Florida, 1325-1331.

[4] Milrad, M., Multisilta, J., (2009). Sharing Experiences with Social Mobile Media, Proceedings of the 11th international Conference on Human-Computer interaction with Mobile Devices and Services, Bonn, Germany, 1-3.

Interface Design of Three Modules for Child-Computer Interaction

Lai-Chung LEE[a*], Whei-Jane WEI[b]
[a] *National Taipei University of Technology, Taiwan*
[b] *Taipei Municipal University of Education, Taiwan*
*f10666@ntut.edu.tw

Abstract: The paper presents three educational module designs entitled tactile, haptic and multi-mode interaction different in open-ended solutions to the problems. For creative problem solving (CPS), module II is fully open-ended designed. The underlying theory is based on scaffolding inquiry and discovery learning. Module III is most helpful for CPS.

Keywords: Interactive technology, interface design, young children.

Introduction

Apple iPad and Kinect has been introduced in preschools since 2010 [1, 2, 4]. Preschool teachers spend much time in preparing tangible materials in need of storage. Interactive technology is more sustainable and space-saving. Through motion sensing, Kinect enables children to use their gestures rather than a mouse or controller to interact with the content on the screen. Kinect makes children's learning experiences become extraordinary and immersive.

1. Interface Design for Three Modules

1.1 Interface Design for Embodied Interaction

Hsu (2011) mentioned about the implications of Kinect that a user can make sounds and movements on the screen synchronized by his/her gestures and its movement detected by Kinect [3]. Figure 1 shows that children played in front of the screen to move the screen's object into the right position. The story related to the proverb of Wolf and Three Pigs. Children's mission was to build up the wall in order to prevent the invasion of the wolf. There were five materials on the screen such as brick, paper cartoon, wood, plastic bottle, and stack. Each material comprised different texture and weight. Therefore, children should make a decision on the choice of material and move it to the right position. The computer programming allowed children to stick to a chosen material about one second, then he/she smoothly moves and stack it onto the baseline on the screen. The children tended to use the different materials to construct the wall, by piling up one on top of another. While stacking up the materials, they realized that different texture and weight make it rigid and solid. The time limit for playing one game is 100 seconds per round.

1.2 Interface Design for Haptic Interact

For most of young children, drawing on iPad is an exciting experience. Module II was designed to empower children drawing on touchscreen. The story narrated about the Adventure of Henry in Forest. The first scene demonstrated that Henry encountered the heavy rain but he had nothing to prevent from being got wet. He needed to think about how to solve the problem. Seated Children drew a picture by which Henry can prevent himself from being wet. Children might draw an umbrella, a big leaf even a tree. The second scene illustrated that Henry was going to cross over one quickly flowing river without a bridge available. Children started brainstorming for creative problem-solving. Figure 2 showed a cooperative learning for CPS in which two children discussed what the cause of the problem was and what solutions were possible. They drew the solution projected via PC transmission. With the right solution, such as a bridge, Henry successfully crossed over the river to right side.

Figure 1: Module I

Figure 2: Module II

Figure 3: Module III

1.3 Interface Design for Multi-Mode Interaction

In Module III, we developed a combination of interactive technology, namely Kinect and iPad shown in Figure3. Children were doing cooperative learning for CPS. On the left-hand side, two children figured out what the problem was with a falling hot-air balloon and discussed possible solutions for CPS. Through wireless transmission, the solutions were popping up on the screen whilst one of children will wave his hand to move any right solution onto the right location for CPS. As a result, the hot-air balloon was rising up toward the sky. For Module III, children used their hand detected by the Kinect to grasp the yellow square received from the iPad and move it to the right location on the balloon by appropriate gesture. Even though, the red square cannot match the broken problem, this game could continue the progress till the problem was solved.

2. Method

We conducted an experimental study with 86 children (42 boys and 44 girls) mixed-age 4-6 at three (one private and two primary supplementary) preschools. First of all, we set up two video cameras in advance. For technical trial, we asked two children to play with the interactive games. And meanwhile, we adjusted all equipment including hardware and software. During the study of module I and module II, children participated in a 40-minutes interactive learning activity. In module III, children joined cooperative learning with iPad and Kinect. After the above session, we asked children to fill up a worksheet.

3. Results

3.1 Effect of Interactive Devices

For module effectiveness, evaluation of worksheets were developed with problems to be solved corresponding to digital contents of each module. After data collection and analysis, the results were summarized in Table 1 and Table 2.

Table 1: ANOVA on module effectiveness evaluation

Source of variation	*SS*	*df*	*MS*	*F*
Between Within	8010.023	2		8.737***
Within-Subjects	65353.167	85		
Error (module)	77928.397	170	458.402	
Total	151291.587	257		

$F_{(.95)(2,170)}$ =3.00, ***p<.001

Table 2: Paring comparison summary

	Module I	Module II	Module III
Module I	—	.024*	.000***
Module II	—	—	.025*
Module III	—	—	—

p<.05,***p<.001

Table 1 shows that the learning effectiveness of three modules reach significant differences at F(2,170) =8.737, p=.000<.001. The means of the learning effectiveness within three modules are Module I=31.76, Module II=39.616 and Module III=45.35. It indicates that haptic interaction can better promote children's creative problem solving (CPS) than the embodied interaction does. Multi-mode interaction shows the best effectiveness in enhancing children's CPS. Table 2 illustrates that an extremely significant difference has been reached between Module I and Module III. It means that multi-mode interaction plays much better influence than embodied interaction. Young children possess better problem-solving skills by haptic interaction [5]. Five- and six-years-old children perform better in CPS than those who are 4-5 years old.

4. Conclusion and Future Work

Interactive technology has been popularly applied in grade 1 to grade 10. However, it is rarely implemented in early childhood education. According to participatory design, we advised preschool teachers to discuss with kids about questions to be solved. According to children's ideas, we planned, designed and developed digital contents. For integration of hardware and software, we conducted technical trial and usability testing with a couple of children. We accomplish the design of haptic, embodied interaction, and multi-mode interaction modules for young children. Based on CPS approach, any possible solutions to the problems designed in three modules are open-ended without fixed answers. Our results provide some evidences that children become creative thinkers in problem-solving. The modules are significant in enhancing CPS. Posteriori comparisons show that pairing comparisons are significantly different. Especially, multi-mode interaction results the best effectiveness on CPS.

References

[1] Follmer, S., Ballagas, R., Raffle, H., Spasojevic, M., & Ishii, H. (2012). People in books: Using a flashcam to become part of an interactive book for connected reading. In S.E. Poltrock, C. Simone, J. Grudin, G. Mark and J. Riedl (Eds), *Proceedings of CSCW'12* (pp. 685-694). Seattle, USA: ACM Press.

[2] Hourcade, J. P., Bullock-Rest N. E., & Hansen, E. H. (2012). Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders. *Personal and Ubiquitous Computing*, *16*(2), 157–168.

[3] Hsu, H. M. (2011). The Potential of Kinect in Education. *International Journal of Information and Education Technology*, *1*(5), 365-370.

[4] Lee, L. C., & Wei, W. J. (2011). Impacts on attention with Computer-Supported Learning Devices. In W. Shen et al. (Eds), *Proceedings of CSCD'11* (pp. 722-726). Lausanne, Switzerland.

[5] Marco, J., Cerezo, E., Baldassarri, S., Mazzone, E., & Read, J. C. (2009). Bringing tabletop technologies to kindergarten children. In A. Blackwell et al. (Eds), *Proceedings of BCS HCI'09* (pp. 103-111). Cambridge, UK: ACM Press.

Rediscovering scientific laws in high school physics labs with mobile devices

Wing-Kwong Wong [a], Tsung-Kai Chao [b], Yunn-Wen Lien [c], Chao-Jung Wu [d]
[a]*Department of Electronic Engineering, National Yunlin University of Science & Technology, Douliou, Yunlin, Taiwan*
[b]*Graduate School of Engineering Science and Technology, National Yunlin University of Science & Technology, Douliou, Yunlin, Taiwan*
[c]*Department of Psychology, National Taiwan University, Taipei, Taiwan*
[d]*Department of Educational Psychology and Counseling, National Taiwan Normal University, Taipei, Taiwan*
{ [a]wongwk, [b]g9810815}@yuntech.edu.tw
[c]ywlien@ntu.edu.tw
[d]cjwu@ntnu.edu.tw

Abstract: Traditional physics labs in high school ask students to follow step-by-step procedure to collect data in order to confirm given physical laws. This study proposes to let students take up the role of scientist in physics labs. Modern physics labs can make use of powerful mobile technologies of smartphones and Lego Mindstorms NXT. These devices can be used to act as moving objects as well as data acquisition devices. Moreover, students are asked to find the functional relation between physical variables, e.g., pendulum length and pendulum period, by detecting patterns in experimental data. An empirical study was done in a real high school setting and some interesting preliminary results were found.

Keywords: Scientific discovery, physics labs, mobile devices, pattern finding, hypothesis generation

Introduction

Traditional physics lab is a step-by-step procedure leading to predictable results with no surprises. Some traditional physics labs lack precision. For example, traditional free-fall experiment marks the position of a falling object with a dot on a strip of paper. The precision is quite low. Some lab needs to make assumption. For example, in a pendulum experiment, a student takes the total time of 10 cycles and derived the average period by dividing the total time by 10, thus assuming the periods are equal. But they do not know how constant the period is. Moreover, they do not see the curve of harmonic motion with damping. In short, in traditional labs, students are asked to confirm given scientific laws with experimental data.

Our study takes an approach similar to the ThinkerTools Inquiry Curriculum (White and Federiksen, 2000), which follows the scientific method in the following steps: (1) question---students are asked to think the functional relationship between the period P of a pendulum and the length L of the pendulum; (2) hypothesize---students produces a new term by combining old terms such as P/L; (3) investigate---students carry out physics experiments using different lengths of pendulum and find out the average period for each length; (4) analyze---students study the data of P and L and try to find out their relationship; (5) model---students express P as a function of L like a scientific law; (6)

evaluate---students can do another experiment using a novel length and check the predicted period using the hypothetical model against the measured period. The novelty of our study is the use of a mobile device (smartphone and NXT) as an object in the experiment and a data acquisition tool (e.g., Peters, 2010; Wogt & Kuhn, 2012).

This article proposes to turn physics labs around in a way such that students can take up the role of a scientist (Figure 1). This can be achieved by using modern technology in acquiring high-precision data and finding patterns in these data to rediscover scientific laws (bottom-left of Fig. 1). Obtaining high-precision data is possible with low-cost equipment such as smartphones and Lego Mindstorms NXT (bottom-right of Fig. 1). After numerical data are acquired in experiments, students can be asked to find out an algebraic relationship between variables (representing physics features) with a computer-assisted strategy (top-right of Fig. 1). The environment is designed based on previous studies of InduLab (central component of Fig. 1, Wu et al., 2006) This is the stage of pattern finding and hypothesis generation. For the students doing these labs, their skills of scientific discovery can be enhanced (all corner components of Fig. 1).

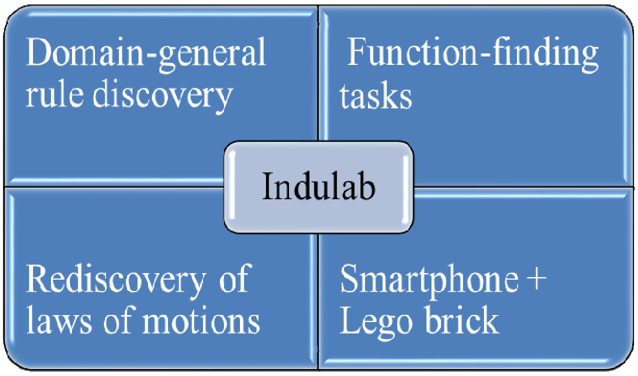


Figure 1: Components of a modern physics laboratory for teaching

Research Method

A class of 36 seventh grade students in a high school in Central Taiwan was chosen to do the experiment of physical pendulum for four weeks, with two fifty-minute sessions per week. They were divided into six groups randomly. In the first week, the researcher lectured on the objective and detailed steps of how to run this experiment. In the second week, three of the six groups did the experiment with Android phones while the other three groups used NXTs to acquire angle data with different lengths of pendulum. For each length of the pendulum, they measured the period of the pendulum for at most ten periods and averaged them. Then they recorded the lengths of the pendulum and the corresponding periods in a table. In the third week, the first three groups did the same experiment with NXT while the other groups used Android phones. In the fourth week, they did pattern finding and hypothesis generation for the Android experiment and for the NXT experiment.

Table 1 shows the sequence of variables proposed by one student. The student might have detected the pattern of decreasing T/L for increasing T. If so, T*T/L was a reasonable variable to try next. For the pattern finding and hypothesis generation part, 5 of the 35 students using smartphone data had once generated the target variable of T^2/L or L/T^2; 5 of the 34 students using NXT data had once generated the target variable.

Table 1. Sequence of variables proposed by one student

L(m)	T(s)	T/L	(T^2)/L	(100T^2)/L	(100T^2-4L)/L
0.27	1.03	3.81	3.92	392.92	388.92
0.31	1.10	3.54	3.90	390.32	386.32
0.37	1.20	3.24	3.89	389.18	385.18
0.45	1.30	2.88	3.75	375.55	371.50

Conclusion

Traditional physics labs in high school ask students to follow step-by-step procedure to collect quantitative data that fits a given physical law. Sometimes the precision of some measurement instrument is not of high quality and bold assumptions have to be made. We propose to transform the physics labs so that students can acquire data with high precision and find patterns in the data in order to rediscover physical laws. Data acquisition with high precision is achieved with modern mobile devices of Android phones and Lego Mindstorms NXT. Pattern finding and law rediscovery are achieved by a computer-assisted strategy of pursuit to produce new algebraic variable from existing ones. A few experiments were designed by the researchers and run by high school students in physics labs. One experiment was to find the relationship between the length of a physical pendulum (of NXT or smartphone) and the period of the pendulum.

Results indicated that the acquired data were of high quality and the period of a physical pendulum remained almost constant when the pendulum swings. For the part of pattern finding and law rediscovery, about 14% of students proposed the targeted law among many others. Most students needed more training in doing this type of pattern finding, which is an important skill in scientific discovery. Comments from some students said that they were excited to use modern mobile technology in physics labs.

Acknowledgement

This study is supported by National Science Council, Taiwan with contract 100-2511-S-224 -002- and 101-2511-S-224-001-.

References

[1] Haverty, L. A., Koedinger, K. R., Klahr, D. and Alibali, M. W. (2000). Solving inductive reasoning problems in mathematics: Not-so-trivial pursuit. Cognitive Science, Vol.24(2), 249-298.

[2] White, B. Y. and Fereriksen, J. R. (2000). Technological tools and instructional approaches for making scientific inquiry accessible to all. In M., J. Jacobson & R. B. Kozma (Eds), Innovations in science and mathematics education: Advanced designs for technologies of learning (pp. 321-359). Mahwah, NJ: Lawrence Erlbaum Associates.

[3] Peters, R. D. (2010a). Smart-phone sensor of pendulum motion, online at http://arxiv.org/abs/1012.1800.pdf.

[4] Wogt, P. and Kuhn, J. (2012). Analyzing free fall with a smartphone acceleration sensor. The Physics Teacher, Vol. 50, Issue 3, 182-183.

[5] Wu, Chao-Jung, Wong, Wing-Kwong, Cheng, Ying-Hao, Lien, Yunn-Wen. (2006). Generating and evaluating geometry conjectures with self-directed experiments. In Proceedings of the 30th annual conference of the International Group for the Psychology of Mathematics Education, 401-408, Prague, Czech.

A Two-Dimensional Approach to Creativity Assessment for Digital Game Story Design

Yun-Hsuan HUANG[a*] & Tsung-Yen CHUANG[b]
[a*] *National University of Tainan, Taiwan*
[b] *National University of Tainan, Taiwan*
D10055003@stumail.nutn.edu.tw

Abstract: This assessment design examines the deficiencies and flaws of existing creativity measures for digital game story design. Instead of using holistic scoring simply overlooking an overall story, a two-dimensional analytic approach to creativity assessment which combines distinct game story constructs with six types of creativity indicators is proposed. Besides, creativity descriptors commonly used by educators in scoring rubrics are also been characterized to correspond to the categories of creativity indicators. This creativity assessment not only helps evaluators like teachers measure digital game story creativity, understands which aspects of the digital game story are short of creativity, but also prompts digital game story designers like students and product developers to tailor creative and entertaining game stories.

Keywords: creativity assessment, game story design, game story creativity, digital game stories

Introduction

Creativity has been explored and pursued by human beings because many masterpieces and innovations are yielded through creativity. In the digital game world, one of the crucial factors attracting and immersing digital natives into game play is the game story, as a game involving a story enriches game substance and creates a vivid sense of reality in the fancy game world. How to make a game story fascinating to draw players into a vicarious game world requires special elements. Undoubtedly, creativity is a crucial element for a game story. Thus, creativity assessment for game story design is an imperative.

1. Background, Problems and Purposes

Despite the fact that game stories, story creativity, or creativity assessment has been extensively discussed [1]-[3], scarce information specifically addresses the intersection of the three-creativity assessment of digital game stories. As game stories, unlike other types of stories such as movies and novels, possess story-game duality [6]; thus, the creativity assessment for digital game stories demands to customizably measure whether creativity in digital game stories inspires players to immerse into and interact with the fantastic game world, thereby enjoying story- and game-like entertainment. However, creativity indicators of existing measuring instruments do not adequately target digital game story creativity. Aside from the digital game industry, pedagogically, digital game design courses also need to provide useful information. To fill up this deficiency, this assessment design aims at developing a two-dimensional measure specifically for digital game story creativity. In addition to provide useful creativity assessment for digital game stories and through it more precisely diagnose which aspects of game story design needs enhancement, this assessment design also highlights a dynamic way of inspiring creativity through game story design.

2. Literature Review

2.1 Creativity Assessment

Due to multitudes of perspectives for creativity, creativity assessment varies accordingly. According to Hocevar and Bachelor's [7] classification from more than 100 examples of creativity assessment, they inducted eight categories of creativity assessment tools. These categories are listed as follows:

Table 1
Types of Creativity Assessment Instruments

Category	Example
Tests of Divergent Thinking	Torrance Tests of Creative Thinking [4] Creativity Assessment Packet, CAP [5]
Attitude and Interest Inventories	Basadur Preference Scale [1]
Personality Inventories	Iowa Inventiveness Inventory [8]
Biographical Inventories	The Creative Achievement [9]
Rating by Teachers, Peers and Supervisors	Domino Creativity Scale, ACL [3]
Judgment of Products	Creative Product Assessment Matrix [2]
Eminence	Genius, creativity, and leadership: Historiometric enquiries [10]
Self-reported Creative Activities and Achievements	The construct of creativity: Structural model for self-reported creativity ratings [11]

Among the eight categories, most commonly adopted techniques are two: one is the Divergent Thinking approach; the other is the Consensual Assessment [12] approach which covers the method of Rating by Teachers, Peers and Supervisors and the method of Judgment of Products [7].

2.2 Creativity Indicators

Whichever creativity assessment it is, existing creativity assessments share similar creativity indicators but also label differentiated ones. For instance, the creativity indicators employed in Torrance Tests of Creativity Thinking [4] comprise fluency, flexibility, originality, average, elaboration, creativity index, abstractness of titles, and resistance to closure. The William Scale [5] for creativity thinking proposes curiosity, imagination, complexity, and risk-taking. The Pythagoras measure [13] of verbal and nonverbal creativity applies creative imitation, verbal interpretation, original production, and verbal aptitude.

Aside from the notions defined by researchers, a tremendous variety of vocabulary has been practically applied by educators to indicate creativity in scoring rubrics. Such words are like original, novel, unique, innovative, unusual, inspiring, complex, sophisticated, fantastic, unpredictable, dramatic, interesting, surprising, intertwined, unexpected, varied, or using colorful words, imagination, vivid images, suspense, foreshadowing, climax, conflict, irony, imagery, metaphor, symbolism, simile, producing curiosity, etc. These terms will be categorized later in Figure 2.

2.3 Narrative and Story Writing Assessment

In the area of narrative and story writing assessment, myriads of assessment rubrics have been dedicated by researchers and educators. The commonly used constructs for narrative writing, story writing, or creative writing include content, knowledge and understanding, plot, organization, communication, application, word and language choice, voice, mechanics, style, originality, focus, integration, etc [14-16]. Yet, these constructs are not all effective for game stories.

3. Development of Creativity Assessment for Digital Game Stories

3.1 Problems of Existing Creativity Assessment Instruments

No matter which kind of assessment it is, none can adequately and specifically target the assessed domain-game story creativity, as game story creativity simultaneously involves the dimensions of creativity in game and story contexts.

For general creativity assessment rubrics, they lack measurement in the attributes of game contexts and story writing. For creative writing assessment rubrics, they neglect the characteristics of game contexts and narrative or story genre. For narrative or story writing assessment rubrics, they do not closely address creativity as a core, nor do they thoroughly consider some significant game essences such as virtual characters, fantastic scenes, and exciting game play. To tackle these problems, this assessment design thus devices a two-dimensional analytic rubric model to target this specific domain.

3.2 Proposed Theoretical Framework of Game Story Constructs

For the purpose of this assessment design, five game story constructs are selected as manifested in the following framework.

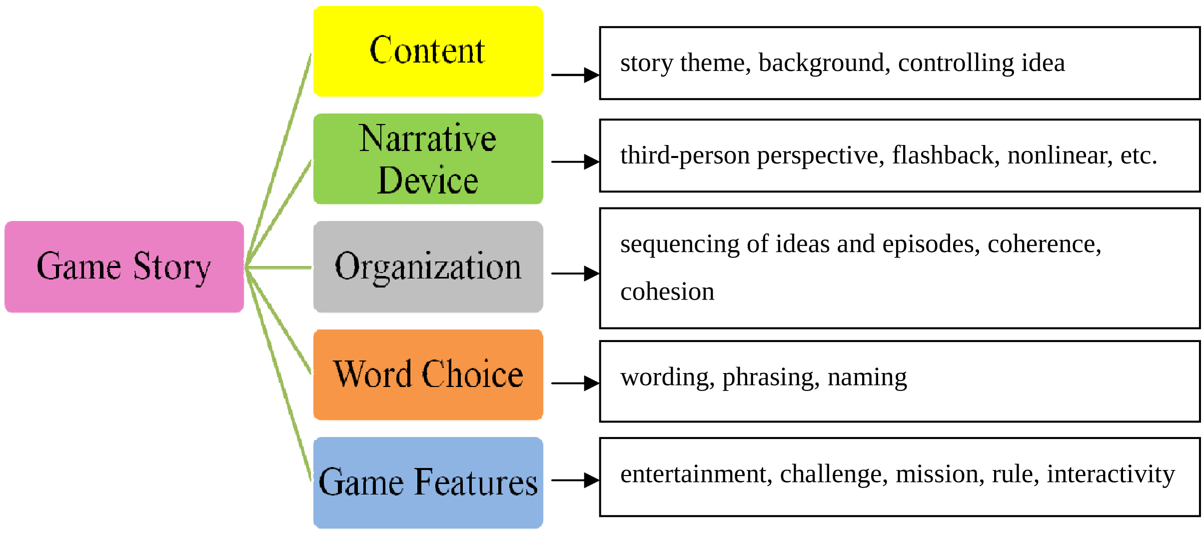


Figure 1. The theoretical framework of game stories and the statement of constructs

3.3 Proposed Creativity Indicators for Game Stories

This assessment design examines creativity in terms of the "product' dimension from Rhodes' 4P [17], and the assessment rubric generated is through the Judgment of Product technique. Six types of creativity indicators are selected: originality, complexity, imagination, flexibility, wonder and application. The more types of creativity indicators involve, the more abundance creativity presents. Besides, the aforementioned terms such as surprising, novel, imaginative applied by educators in scoring rubrics are characterized as creativity descriptors to interpret the six types of creativity indicators, which is illustrated below.

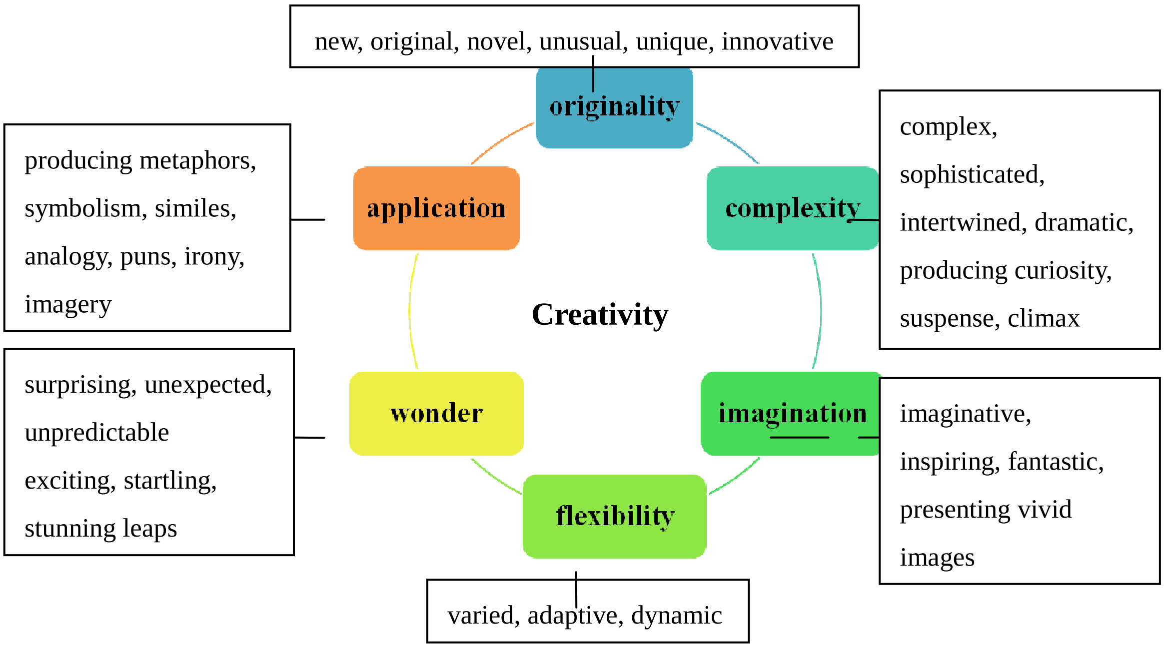


Figure 2. The framework of creativity indicators for digital game stories

3.4 The Analytic Scoring Rubric of Creativity Assessment

Based on the above frameworks, the scoring rubric is developed in the following.

Table 3
Analytic Scoring Rubric for Creativity of Digital Game Story Design

Construct	Statement of Scoring	
Content	6	Proficiently applies 5-6 types of creativity indicators
	5	Skillfully applies 3-4 types of creativity indicators
	4	Adequately uses 1-2 types of creativity indicators
	3	Unskillfully practices any type of creativity indicators
	2	Uses a common story theme, background, and controlling idea
	1	Displays an unclear story theme, background, and controlling idea
Narrative Device	6	Proficient conveyance with 5-6 types of creativity indicators
	5	Tactical conveyance with 3-4 types of creativity indicators
	4	Adequate delivery with 1-2 types of creativity indicators
	3	Unskillful delivery with any type of creativity indicators
	2	Smooth use of a common narrative technique
	1	No narrative technique

Organization	6	Proficient sequencing with 5-6 types of creativity indicators
	5	Tactical sequencing with 3-4 types of creativity indicators
	4	Adequate sequencing with 1-2 types of creativity indicators
	3	Unskillful sequencing with any type of creativity indicators
	2	Ordinary sequencing of ideas and episodes
	1	Poor or confusing sequencing of ideas and episodes
Word Choice	6	Proficiently depicts with 5-6 types of creativity indicators
	5	Brilliantly applies 3-4 types of creativity indicators
	4	Appropriately portrays with 1-2 types of creativity indicators
	3	Unskillfully practices with any type of creativity indicators
	2	Shows limited ability in wording, phrasing, naming
	1	Reveals many inappropriate wording, phrasing, and naming
Game Features	6	Excellent game features with 5-6 types of creativity indicators
	5	Skillful game features with 3-4 types of creativity indicators
	4	Impressive game features with 1-2 types of creativity indicators
	3	Ordinary game features
	2	Poor game features
	1	No game feature

4. Characteristics of the Analytic Rubric

Distinct from other assessment rubrics, this two-dimensional analytic scoring rubric comprising game story constructs and creativity indicators is characterized with four traits: (1) It consists of a breakdown of game story constructs, identifying the features of each game story construct for precise assessment; (2) The creativity indicators specifically target digital game stories, as they can detect the creativity that digital game stories need for stimulating players' excitement and enjoyment in the interplay of experiencing the story and playing the game, rather than simply for story readers; (3) Based on semantic implications, the creativity descriptors generally adopted by educators are categorized to correspond to creativity indicators termed differently by researcher; (4) Such a categorization integrates the perspectives and terminologies from both researchers and educators for creativity assessment.

5. Conclusion

The creativity assessment of digital game stories has not received deserved attention. Even though conducted, it is often treated as assessing general story writing without scrutinizing the nature and constructs of digital game stories as well as the adaptability between creativity indicators and digital game stories.

This assessment design devises a two-dimensional analytic approach to creativity assessment for digital game stories, identifying the significance of individual game story constructs and integrating the association of creativity indicators denoted by researchers and educators. This analytic rubric can not only assist evaluators like teachers to precisely assess digital game story creativity as well as effectively observe which aspects of game story writing are short of creativity, but also prompt game story designers like students and product developers to tailor creative and entertaining game stories. For digital game design courses, the proposed game story constructs, creativity indicators, and the analytic rubric can be used as effective teaching resources. More importantly, this assessment design accentuates the dynamic way of sparking creativity through digital game story design. With this assessment rubric developed, future study will further examine its validity and reliability.

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 98-2511-S-024-004-MY3, NSC 99-2511-S-024-003-MY3, and NSC 101-2631-S-001-001.

References

[1] Basadur, M., & Hausdorf, P.A. (1996). Measuring divergent thinking attitudes related to creative problem solving and innovation management. Creative Research Journal, 9(1), 21-32.
[2] Besemer, S.P., & O'Quin, K. (1987). Creativity product analysis: Testing a model by develping a judging instrument. In S.G. Isaksen (Ed.), Frontiers of creativity research: Beyond the basics (pp.367-389). Buffalo, NY.: Bearly.
[3] Domino, G. (1994). Assessment of creativity using the ACL: A comparison of Four scales. Creativity Research Journal, 7(1), 21-23.
[4] Torrance, E. P. (1966). The Torrance tests of creative thinking-norms-technical manual research edition-Verbal Tests, Forms A and B-Figural Tests, Forms A and B. Princeton, NJ: Personnel Press.
[5] Williams, F. (1980). Creativity assessment packet: Manual. East Aurora: N.Y.: DOK Publishers.
[6] Schell, J. (2008). The art of game design. Burlington, MA: Elsevier Inc.
[7] Hocevar, D., & Bachelor, P. (1989). A taxonomy and critique of measurement used in the study of creativity. In J.A. Glover, R.R. Ronning, & C.R. Reynolds Eds.),Handbook of creativity (pp.53-75). New York: Plenum Press.
[8] Colangelo, N., Kerr, B., Hallowell, N., & Gaeth, J. (1992). The Iowa inventiveness inventory: Toward a major of mechanical inventiveness. Creativity Research Journal, 5, 157-164.
[9] Lubart, T.I. (1994). Thinking and problem solving. In R.J. Sternberg (Ed.), Creativity (pp.289-332). New York: Academic Press.
[10] Simonton, D.K. (1984). Genius, creativity, and leadership: Historiometric enquiries. Cambridge, MA: Harvard University Press.
[11] Kaufman, J.C., Cole, J. & Baer, J. (2009). The construct of creativity: Structural model for self-reported creativity ratings. Journal of Creative Behavior, 43(2), 119-132.
[12] Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. Journal of Personality and Social Psychology, 43, 997-1013.
[13] Grove, G.A. (1999). PYTHAGORAS B / C: An experimental measure of verbal and nonverbal creativity. Retrieved from http://www.mysteriumsociety.org/psimetric%20solutions/pyth_bc.html
[14] Arter, J., & McTighe, J. (2001). Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance. Thousand Oaks, CA: Corwin Press Inc.
[15] Engel, T., & Streich, R. (2011). Yes, There Is Room for Soup in the Curriculum: Achieving Accountability in a Collaboratively Planned Writing Program. The Reading Teacher, 58 (7), 660-679. doi: 10.1598/RT.59.7.5
[16] Spence, L. A. (2010). Discerning writing assessment: Insights into an analytical rubric. Language Arts, 87 (5), 337-347.
[17] Rhodes, M. (1961). An analysis of creativity. Phi Delta Kappan, 42, 305-311.

A Comparison of Learning Effectiveness Among Serious Games with Varying Degrees of Playability

Ching-Sheng Yang [a*], Shu-Hua Lin [b], Kuo-Hua Wang [a]
[a] *Graduate Institute of Science Education, National Changhua University of Education, Taiwan*
[b] *Department of Biology, National Changhua University of Education, Taiwan*
*pureman720730@gmail.com

Abstract: Serious games are composed of three components, namely, learning contents, learning strategies, and game elements. To emphasize the significance of game elements on learning effectiveness in serious games, this study proposed a design pattern for courses of serious games, based on serious game courses which addressed the endocrine unit for students in the seventh grade. Since the courses were first developed and put to use three years ago, the courses have been implemented and modified many times. The courses adopted in this study were the fifth version. Employing a quasi-experimental research design, an experiment of two-week courses was organized in three groups, namely the traditional instruction group, the serious game group, and the low playability game group, which used courses with less playabilities modified on the basis of serious games. A total of 234 students participated in this research. The research instruments included achievement test. The results showed that serious game courses with sufficient playabilities outperformed the low playability group and traditional group in both the achievement test, indicating that playability was an important factor influencing the effectiveness of serious game courses. Additionally, with respect to the categorical variables, gender differences had no significant influence over serious game courses. Learning background had no significant influence on the outcome of achievement test of the serious game group.

Keywords: Serious game, Endocrine, Playability, digital game

Introduction

1. Research motivation and background

Digital gaming accounts for a substantial portion of the time spent on computers by students in Taiwan, and a similar phenomenon occurs in other countries (McFarlane, Sparrowhawk, & Heald, 2002). Naturally, computer games play a significant role in the everyday life of students. When applying digital gaming to learning practices, we must pay attention to the integration of game playability, in addition to learning contents. The games that maintain a balance between learning contents and playabilities are referred to as serious games. A serious game may contain a few key elements, such as meaningful learning content incorporating game elements, learning strategies guiding students' exploration of learning content, and game elements enhancing learning. Serious games are digital games characterized by the incorporation of both entertaining and instructional functions, while avoiding the boredom of lite games (Michael & Chen, 2006; Zyda, 2005).

Serious games are composed of three components, where the playability is defined by many scholars as the qualities enabling a game to be fun and entertaining to play (Alessi & Trollip, 1985; Malone, 1981; Prensky, 2001). Researcher summary these qualities and divid them into two reciprocally influencing categories, namely, “game attributes” and “design elements,” and then used them to describe the elements necessary in the development of digital game-based instruction. The two groups of course designs with different playabilities are distinctive from each other in their design elements.

Table 1
Essential design and game attributes of digital game

Design / Game factor	Competition & Challenge	Result & Feedback	Rule & Goal	Story
Problem solving	X	X	X	
Adaptation	X	X	X	
Successes	X	X	X	
Playfulness	X	X		
Interactive	X	X		
Funny	X	X		X
Fantasy				X

Learning contents need to be reconstructed by employing a combination of appropriate game elements and learning strategies, because simply mixing concepts in the textbooks into games may lead to page turner styled learning. Without appropriate learning strategies, serious games probably turn into nothing but entertainment. Because of time limits in course delivery, appropriate learning strategies may help shorten the time between flow and break down (Roth, 2006), preventing learners from wandering around aimlessly. This also prevents learning activities from being turned into purely entertaining games. One of the key characteristics of games is its problem solving. It is of practical significance to integrate problem solving strategies into digital gaming based courses. According to previous studies, a problem solving process may consist of six steps: emergence of problems, identification of problems, sifting through all kinds of information to seek key information for problem solving, seeking reasonable interpretation to develop an answer, judging or verifying the precision of the answer, and solving the problem by finishing a workable conclusion.

Previous studies have discussed the difference of learning performance using commercial entertainment games and lite games with variable playabilities. However, little research has been done to explore the effects of playabilities on learning performance through serious game courses. The reason for this is that previous studies have been unable to modify the particulars or programs of the serious game courses because they were not entirely designed by themselves. In this study, the serious game courses were jointly designed by the researchers and a graphic designer, thus we could modify and adjust any part of the course content as needed.

2. Research objectives and questions to be answered

In summary, this study develops serious game courses as well as low playability digital gaming courses modified on the basis of serious game courses, to understand the effects of playability on learning effectiveness. In addition, this study also explores the influences of background variables, which are commonly examined in digital gaming related research, such as gender differences and learning backgrounds, to understand whether these attributes influence the learning styles of different strategies. Previous studies have reported that different genders have different preferences toward game genres(Bonanno & Kommers, 2008 ; Wang & Wang, 2008) . This study is also concerned about whether such differences may influence the performance of different learning strategies. Hence, this study attempts to investigate the following questions:

- What is the difference in student achievement test using narrative, low playability games and serious game courses?
- How do gender differences and learning background influence achievement test?

3. Research Methods

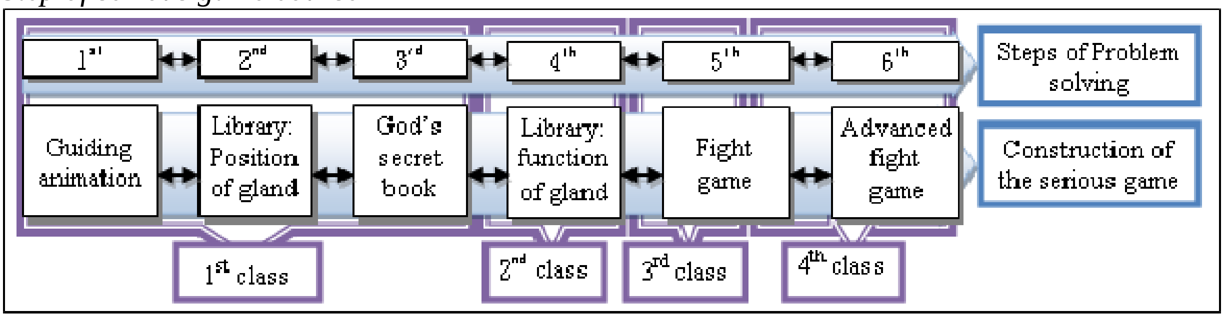
3.1 Serious game course

Adopting the design pattern of serious game courses presented in this study, the researchers developed a serious game course entitled, "Endocrine Warrior." The development tools were: Adobe FlashCS4, PHP5, & MySQL. Since its first release, the module has undergone four revisions within three years. Three professors specializing in education science and 32 junior high school science teachers participated. The courses have been implemented seven times, in 16 classes of the seventh grade and 14 classes of the ninth grade. Since its implementation, the courses have undergone three major revisions to reflect the comments and suggestions made by the teachers and education experts. The serious game adopted in this study was the fifth revision.

The serious game course were composed of five components, including the "God's secret book & Library," "Lab: Position of gland," "Lab: function of gland," and "Fight games and recorder of gaming experience." The instructional knowledge contained in each component was reconstructed according to their respective and adapted game elements. The six components were linked through an animated startup story, and the course design followed the serious game model designed by the researcher. The implementation of course lasted for two weeks and was delivered in four sessions.

Fig 3

Step of serious game course



3.2 Research Instruments

The main research instruments used by this study included an achievement test on endocrine concepts. The achievement test on endocrine concepts was compiled by the researchers themselves, and evaluated by ten science teachers from senior junior high schools and three professors specializing in education science. The test papers were pre-tested with trials to verify reliability, degree of difficulty, and discrimination. Altogether, 570 ninth grade students participated in the pre-test. The valid sample was 562 students, the KR20 value was 0.84, the mean degree of difficulty was 0.66, and the mean discrimination was 0.50.

3.3 Research Design

This study employed a quasi-experimental control-group design with repeated measures.. The research subjects were 234 students of the seventh grade from seven regular classes, who were then divided into three groups, namely the narrative course (n=66), serious game course (n=101), and low playability game-based course (n=67). The narrative group (NG) was the control group, in which traditional narratives were used for course delivery. The serious game group (SGG) delivered a complete set of serious game course. The low playability group (LPG) substituted the third cycle ofserious game with seminars to implement the last two steps of the problem solving (see Fig 2). With respect to the game design (see Table 2), the games in this group were less competitive and challenging than those in the serious game group.

All three groups took a pre-test of achievement on endocrine concepts prior to course implementation. Then, two weeks of course instruction was implemented. A post-test of achievement was taken two days after the courses were finished. The 1st delay-test of achievement was administered two weeks after the courses finished. Finally, eight weeks after the delay-test, a 2nd delay-test of achievement on endocrine concepts were administered to test the long-term memory retaining effects.

Table 2
The research design

Group	2 days ago	Course of implement (2 weeks)	2 days later	2 weeks later	10 weeks later
NG	pre-test	traditional narratives	post-test	1st delay-test	2nd delay-test
SGG		serious game			
LPG		Low playability of serious game			

3.4 Data Processing and Analysis

The classification of students' prior learning backgrounds was based on the average score earned in the last two monthly examinations on Biology preceding the endocrine unit. The students were divided into three groups according to their learning background. The students whose average grades ranked in the first 27% were regarded as high achievement (HA) students; the ones whose average grades ranked in the last 27% were regarded as low achievement (LA) students ; the remaining students were regarded as medium achievement (MA) students .

As to the data analysis aspect, the achievement test was administered four times. Each test randomly sorted the results. Given the same test paper was used, the data analysis should have been carried out using Repeated measures analysis of variance (ANOVA), rather than analyzing individual assessment, to prevent statistical errors.

4. Research Findings and Discussion

This study focuses on the essentiality of playability to the effectiveness of serious game. The subjects of the study were divided into three groups, namely the narrative group (NG), the serious game group (SGG), and low playability group (LPG).. The effectiveness of learning was measured through the scores obtained from a series of assessment, including a post-test, 1st, and 2nd delay-test. Moreover, the influences of gender difference and prior learning background on learning effectiveness were also explored.

Repetitive measurement was corrected using the Huynh-Feldt correction, with the primary effects of the research control reaching a significant level: F (2, 234) = 4.32， p = 0.01(Table 4). The post-hoc comparison was made using the Bonferroni correction. By and large, the SGG significantly outperformed the LPG and NG. However, the effect of LPG was not significantly higher than NG; instead, the achievements of the two were very close.

Table 4
Three-way repeated measures ANOVA results

Source	*SS*	*df*	*MS*	*F*	
Between group					
Achievement test (Repeated measure, A)	86237.82	2.93	29481.15	210.77	**
Control (B)	4196.59	2.00	2098.30	4.32	*
Gender (C)	61.83	1.00	61.83	0.13	
Background (D)	101664.70	2.00	50832.35	104.67	**
*A * B*	3194.18	5.85	545.98	3.90	**
*A * C*	27.59	2.93	9.43	0.07	
*A * D*	17286.03	5.85	2954.69	21.12	**
*A * B * C*	337.35	5.85	57.66	0.41	
*A * B * D*	2050.50	11.70	175.25	1.25	
*A * C * D*	624.62	5.85	106.77	0.76	
*A * B * C * D*	1122.70	11.70	95.95	0.69	
Within group					
Subject	104901.57	216.00	485.66		
residual	51209.71	430.00	119.09		
Total	372915.21	703.65			

$^{**}p<0.01$; $^{*}p<0.05$;

As interaction was present between the achievement and research control, analysis of simple main effect was performed (Table 5). The researchers were concerned about which didactics had the best performance. In the pre-test, no significant difference was found among the three groups, indicating that the students' prior knowledge before engaging in learning was not statistically different. Significant differences were identified in the post-test, the 1st delay-test, and the 2nd delay-test, which were put to a post-hoc comparison subsequently. The post-hoc comparison was conducted using the Sheffe technique, due to the unequal number of samples. By and large, the serious game group significantly outperformed the LPG in both the post-test and the 1st delay-test, and significantly outperformed the NG in the 2nd delay-test, suggesting the serious group had excellent long-term memory retention. Moreover, although the LPG without combat games and the NG did not reach a significant level, this suggests that without combat games, providing the same learning effects as traditional instruction is possible. If combat games can be provided, the learning performance can be further improved through game-based teaching.

Table 5
Analysis of simple main effect: different course and achievement test

Simple main effect	*SS*	*df*	*MS*	*F*		*Post-hoc*
Different achievement test						
NG	22550.28	2.20	10263.77	46.88	**	—
SGG	60010.58	2.56	23419.63	144.36	**	—
LPG	25153.26	2.58	9759.26	42.95	**	—
Different course						—
Pre-test	418.83	2	209.41	1.47		—
Post-test	3738.66	2	1869.33	5.26	*	2>3*; 2>1 ; 1>3
1st delay-test	6197.14	2	3098.57	7.32	**	2>3*; 2>1 ; 1>3
2nd delay-test	6119.89	2	3059.94	5.58	**	2>1*; 2>3*; 3>1

** p<0.01; * p<0.05; 1:NG; 2:SGG; 3:LPG

In the achievement test of repeated measurement, no interaction was present among the three factors: control, gender, and achievement grouping. Additionally, no interaction was present between the control and gender or achievement grouping, suggesting both achievement grouping and gender exercise little influence on learning effectiveness. In other words, whether male or female, and regardless of learning background, learning through serious games, or low playability games, or traditional instructions rendered no significant difference in the achievement tests.

However, an in-depth analysis revealed that in the 1st delay-test, the male students in the serious game group significantly outperformed the male students in the low playability game group ($F(2, 120)=6.14$ ，$p < 0.00$), suggesting that unlike female students, male students may achieve better learning effectiveness through more immersive game experiences.

Furthermore, the adoption of various didactics may create significant differences in the 1st delay-test for low achievement students ($F(2, 63)=3.99$，$p = 0.02$). Through post-hoc comparison, we saw that the students in the serious game group with lower achievement significantly outperformed the students in the LPG of the same achievement group. In addition, high achievement students in different groups had significant differences in the 2nd delay-test ($F(2, 66)=4.14$，$p = 0.02$). The post-hoc comparison revealed that through combat game-based learning, the high achievements students had better long-term memory retention than those in the NG. Therefore, higher playability was beneficial, to a certain extent, for both low achievement and high achievement students. However, because of the small sample size of background groupings, the analytical results by achievement groupings are pending further verification.

5. Conclusion and Suggestions for Future Research

The purpose of this study is to explore the role played by serious games in achieving learning effectiveness. Test results indicate that serious games with an adequate level of playability outperformed digital game-based courses with lower playabilities or traditional courses with no playabilities at all. Previous studies on digital gaming tended to emphasize hands-on skills (ex: Shen & O'Neil, 2006; Alkan & Cagiltay, 2007; Ko, 2002; Vogel, Greenwood-Ericksen, Cannon-Bowers, & Bowers, 2006; Robertson & Howells, 2008); however, serious game not only met the needs of teachers who were under the pressure of raising enrollment quotas (Kebritchi, Hirumi, Kappers, & Henry, 2009), but also transcended the effectiveness of traditional teaching and learning, eliminating the concerns of high critical tests surrounding onsite teaching.

No significant difference existed in the outcome of serious game between male and female students; however, male students slightly showed more demands for playability in the game-based courses. This is probably because of the game genres. This study primarily adopts role-playing games, simulation games, and strategy games. These genres are typically male oriented (Bonanno & Kommers, 2008 ; Wang & Wang, 2008). Therefore, it would stimulate male students to spend more time studying if more playabilities could be added to the games. This is why the serious game group outperformed the others in the delay-test. The gender difference, nevertheless, had little impact on the overall performance of the learning efforts.

Previous studies have discovered that learning background is inversely relational to digital game experience. Learners with mid-lower achievements are familiar with digital gaming, in line with the affordance of computers (Wijekumar, Meyer, Wagoner, & Ferguson, 2006). Therefore, for learners with mid-lower achievements, serious games with sufficient playabilities are more engaging, and are thus more facilitative for learning processes. By contrast, they have little impact on students with high achievement.

By and large, the courses developed on the basis of the design architecture for serious games in this study can exercise significant influences on learning achievements. Gender difference presents little influence on playability. As far as prior learning background is concerned, courses with more playabilities are more suitable for students with mid-lower achievements, while no significant difference was observed for students with high achievement.

Regarding the direction for future research, we make three suggestions based on the findings of this study as follows. First, serious game courses should be carried out progressively, with playability being added gradually, and the content being deepened continually. Because playability plays a critical role in the learning performance of serious game, playability must be incorporated into the content to prevent a page turner type of design. Given that the design architecture presented in this study is not difficult to accomplish, it can probably be used as a reference to further development of serious game courses. Second, more factors of influence can be taken into consideration in future studies to further explore the effects of playability, such as learning styles and proficiency of digital gaming experiences. Third, future researchers may experiment with more diverse learning modules to develop more serious games, such that the learning effects can be better observed.

Acknowledgements

The authors are grateful to the National Science Council for its grant for this study. The authors are also indebted to all of the experts, professors, teachers, and students who have taken part in the implementation of this serious game. We appreciate their kind comments and suggestions about which game can be further modified.

References

[1] Alessi, S. M. & Trollip, S. R. (1985).*Computer-based instruction: Methods and Development*. NJ: Prentice-Hall.

[2] Alkan, S., & Cagiltay, K. (2007). Studying computer game leaning experience through eye tracking. *British Journal of Educational Technology, 38*, 538-542.

[3] Bonanno, P., & Kommers, P. A. M. (2008). Exploring the influence of gender and gaming competence on attitudes towards using instructional games. *British Journal of Educational Technology, 39*, 97-109.

[4] Kebritchi, M., Hirumi, A., Kappers, W., & Henry, R. (2009). Analysis of the supporting websites for the use of instructional games in K-12 setting. *British Journal of Educational Technology, 40*,733-754.

[5] Ko, S. (2002). An empirical analysis of children's thinking and learning in a computer game context. *Educational Psychology,22*, 219-233.

[6] Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive science, 5*, 333-369.

[7] McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games. report of TEEM (Teachers Evaluating Educational Multimedia).* Retrieved November 8, 2007, from *http://www.teem.org.uk/publications/teem_gamesined_full.pdf*

[8] Michael, D., & Chen, S. (2006). *Serious games: Games that educate, train, and inform.* Boston, MA.: Thomson Course Technology.

[9] Prensky, M. (2001). *Digital Game-Based Learning*. New York: McGraw Hill.

[10] Robertson, J., & Howells, C. (2008). Computer game design: Opportunities for successful learning. *Computer & Education, 50*, 559-578.

[11] Roth, W. M. (2006). *Learning science: a singular plural perspective.* Rotterdam: Sense Publishers.

[12] Shen, C. –Y., & O' Neil, H. (2006). *The effectiveness of worked examples in a game-base learning environment.* American Educational Research Association, San Francisco, CA.

[13] Vogel, J. J., Greenwood-Ericksen, A., Cannon-Bowers, J., & Bowers, C. A.(2006). Using virtual reality with and without game attributes for academic achievement, *Journal of research on technology in education, 39*, 105-118.

[14] Wang, H. –Y., & Wang, Y. –S. (2008). Gender differences in the perception and acceptance of online games. *British Journal of Educational Technology, 39*, 787-806.

[15] Wijekumar, K. J., Meyer, B. J. F., Wagoner, D., & Fergusin, L. (2006). Technology affordances: the "real story" in research with K-12 and undergraduate learners. *British Journal of Educational Technology, 37*, 191-209.

[16] Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, *38 (9)*, 25-32.

Teachers' appropriation of game-based pedagogy: A comparative narrative analysis

Swati MEHROTRA[a], Yam San CHEE[a], Jing Chuan ONG[a]
[a]Learning Sciences Lab, NIE, Nanyang Technological University, Singapore
swati.mehrotra@nie.edu.sg

Abstract: This article examines the process of change through the narratives that two teachers tell to describe their journey over time while participating in the Statecraft X game-based learning program. Data from post dialogic session interviews is used to elicit the challenges and areas of improvement that teachers identify. The two teachers, whose case-study is being used in this paper, had contrasting experiences to share with respect to their professional goals and motivations. One of the teachers did not experience many perturbations, while the journey for the other teacher was loaded with personal and professional struggles. Insights from teachers' narratives indicate some patterns of change that made appropriation successful for one and much less so for the other. These include (a) moving from not adhering to lesson plan to real time 'thinking on the feet' (b) shifting from teacher centric to student centric classrooms, and (c) 'letting go of control' to facilitate more active student learning. These changes have implications for teacher professional development with respect to game-based learning and teachers' readiness for 21st century classrooms.

Keywords: Narratives, game-based learning, dialogic method, professional development

Introduction

Digital games have been used in classrooms over the past few decades with mixed reviews about its success. One of the challenges in advancing game-based learning is the preparation of teachers for handling innovative pedagogy [1]. The introduction of innovative pedagogies in schools often requires educators and students to make substantial shifts in their teaching and learning process. This is not easy and requires constant reflection and support. In this paper we focus on experiences of two teachers through their narratives as they enact the Statecraft X (SCX) game-based learning program.

1. Need for meaningful professional development

Digital games provide a powerful platform for 'authentic learning' where learners get an opportunity to experience complex situations first-hand as opposed to merely using these games for instructional and review purposes. However adopting and integrating serious games in the regular teaching and learning process has had a history of challenges, and the teacher's role in facilitating this uptake has been surfaced as an important requirement. A definitive shift in mindset is required for teachers to take-up game-based learning in the classrooms. However, this does not always come readily. Often support is needed in understanding how to ensure effectiveness of games use in class [2].

For game-based learning, effective professional development needs to be provided in terms of training, hands-on experience with game, and opportunities need to be built in for reflection and reconstruction of knowledge as well as for follow-up support. This is where professional development often lags and new pedagogy is misinterpreted. To overcome these barriers and to provide meaningful professional development support to teachers in the SCX game-based learning program, we planned a phase-based professional development program. We began by conducting a four-day workshop, where teachers were acquainted with the theoretical underpinnings of the game-based curriculum and experienced playing the game. During the enactment of the SCX curriculum, we supported their development through reflective, reflexive, guided appropriation model (RRGA) proposed by Chee [3]. In the model, reflection refers to the process of looking back and thinking upon one's actions, and reflexivity is the capacity to make one's own actions the target of critical interrogation with a view to improving existing practice. RRGA stresses the process of reflection that teachers need to do on their teaching practice along with being reflexive. This process aids in scaffolding and appropriation of an enhanced practice.

2. Research background and purpose

The current SCX project emanated from an earlier project that focused on development of the SCX game, its curriculum, and classroom interventions. The preceding project led to theorization of game-based learning in terms of three constructs - play, performance, and dialog [1] and demonstrated efficacy with respect to student learning. One of the felt needs of the previous project was to better prepare the teachers to enact the game-based curriculum in a manner consistent with the learning design. In light of this, the current project aimed to level up teachers' capacity to enact game-based learning in their classrooms. In order for teachers to be ready for 21st century classrooms, they need to make significant shifts in their practice. As a consequence, there have been numerous demands for preparation of teachers. Our approach to teacher professional development was guided by the following research questions: (a) What are the specific challenges that teachers face in implementing a game-based learning curriculum, and how might these challenges be addressed? (b) What are the trajectories and profiles of teachers' appropriation and ownership of game-based learning pedagogy in the classroom?

3. Research setting and method

3.1 Participants

For the purpose of this paper we have chosen two teacher participants, X and Y (names withheld to maintain confidentiality) from two schools where the SCX curriculum was implemented. Case study approach was followed to examine teacher's experiences "under the microscope." This approach also supported the study of factors that enabled them to change their practices, their personal challenges, and challenges they faced in the context of the school system. We chose these particular cases because they provide meaningful understanding of the teachers' journey and also form insightful cross-case comparison. X and Y were social-studies teachers in the age range of 25-30 years. Both believed that it was their responsibility to teach students the required content and that a well-managed classroom facilitated this process. The teachers were self-motivated and were willing and active participants in the project, and their schools supported implementation of the SCX program. Yet, from the beginning of the intervention, they differed in their teaching goals, their beliefs, motivations and their expectations from students. These differences led to different trajectories of the change process.

3.2 SCX game-based learning program

The SCX game-based learning program focuses on the principles of citizenship and governance that are a part of the social studies curriculum at secondary three level in the Singapore school system. The game is played on Apple iPhones, where players govern their towns, manage their resources, and thus perform the role of governors. In doing so, they enact citizenship as opposed to learning *about* citizenship. Central to the SCX curricular program is its dialogic pedagogy that is used to advance the understanding of citizenship and governance. Dialogic pedagogy draws on the theoretical idea of dialog in the Bakhtinian sense, where ideas are exchanged and lived and are full of personal values and judgments [4]. In dialogic pedagogy, understanding is co-constructed in the classroom and students learn concepts in personally meaningful ways. Teachers serve as facilitators of dialogue by helping students make connections to real-life situations.

The SCX curriculum comprises six sessions spread over three weeks. The first session is introductory, where students are acquainted with the game and a pre-intervention survey is administered. The next four sessions are dialogic, where each teacher facilitates a session comprising 20 students who are engaged in the game-play. The final (sixth) session comprises of student speeches, a summative assessment (essay) and a post-intervention survey. The summative assessment is also administered to a comparable control group, where students are taught Citizenship using the traditional method.

As a part of professional development, teachers were interviewed by the researchers once before the start of the intervention and then from sessions two to six, using a semi-structured interview guide. Each interview typically lasted between 30 to 60 minutes. It was audio recorded and later transcribed. The researchers observed all the classroom sessions.

3.3 Narrative Analysis

Narrative analysis refers to a variety of approaches for studying the "storied nature of human development" [5, p. x]. Clandinin and Connelly [6] draw on a Deweyan view of experience as characterized by situation (place), continuity (past, present and future) and interaction (personal and social) to view narratives. Narratives provide a way to keep experience and action unified. This allows for greater understanding of experiences [7]. In our study, we examined teachers' narratives to obtain a better understanding of their professional development during implementation of the SCX curricular program.

Narrative analysis has temporality as an important characteristic. It differs from discourse analysis which deals with the structure of spoken language and is organized in specific ways to make a particular reality appear 'real'. Narrative analysis entails systematic interpretation and representation of informant's stories. The process includes attending, telling, transcribing, analyzing and reading the experiences [8].This approach appeals to educators and has become an influential research methodology within teacher education [9].

4. Data analysis

This segment of the paper centers on the analysis of narratives of the two teachers who participated in the SCX program. In the sections that follow, we focus on teachers' journeys while they enacted the SCX program. We highlight the contrasts in their experiences and account for reasons that made their journey so different. These comparisons help us understand what the process of teacher change can entail.

4.1 Teacher X's journey enacting the SCX program

X decided to become a teacher early on in her life. She had about six years of teaching experience. Her decision was based on practical concerns of stability and her personality fit. She emphasized the importance of imparting values and worked towards making her students independent thinkers so that, 'even if I take myself out of the classroom, - the end goal would be [that] they are people who [will be] able to come to a certain conclusion on their own.' Teacher X was personally open to trying new pedagogies, but with caution. She was not a gamer and was initially against the idea of games. She saw them as, 'eat[ing] her curricular time' until she was convinced of the value of SCX by experiencing it herself.

X was nervous about facilitating the first SCX dialogic session and confessed to spending many hours on preparation. After her first session she felt that the lesson had gone 'better than her expectations' as her students had participated actively. In the post session interview, she elucidated the challenges with regards to physical arrangements of the room that she faced. After the second session, she seemed quite disheartened with the fact that she was not 'getting it'. Realizing this fact, in second post dialogic session interview she said:

> *But I still feel like I haven't gotten the hang of the whole technique, just that whole fluency, the whole process, it's still not there yet … but I'm a bit more conscious, a bit more aware … still need to figure out how you know to connect all those dots.*

As an experienced teacher, X believed in planning her lesson beforehand. For her, careful and meticulous lesson planning was essential to quality teaching. She found it 'annoying' when her session did not go well even after planning. This fact was surfaced by her early in the program when in the second post dialogic session interview, she said, 'I think as a teacher you don't feel very professional going in not knowing exactly what's going on because most of the time the slides are prepared [in advance].'

From the third session onwards, X started to improve, and she realized it herself. She got away from adhering to a pre-prepared lesson plan. She relied more on working from student's responses and holding a meaningful conversation with them. She consciously sought to bring about a change in her practice, which reduced her stress. In the final post dialogic session interview, she said:

> *I don't prepare. I really don't. … And I realize that it takes the burden away… preparation is ongoing … I am more interested in what is going on. … You don't really have to sit there and plan per se how the whole thing is going to unfold.*

In the same interview, she reflected on some of the practices and ideas that she previously had and how they had changed over time as she advanced with the SCX program. Her refined notion about questioning was that students' responses guided her to pitching her questions. She stated:

> *I think in the past I used to think of scaffolding as some form of structure, format, from step 1 2 3 4, but the interesting thing about the dialogic session is … how I want to scaffold the lesson according to the responses they give me … So actually the kids are the ones who are building up the entire lesson.*

Toward the end of the SCX program, X also realized that in the process of reflecting on her practice and being self-aware, she had adapted her teaching style to the needs of her students and achieved better engagement. In the final interview she commented:

> *I just see them [students] as individuals, they may have different needs or they learn slightly differently… So engagement may not necessarily come in form of academic kind of talk and I think sometimes they need breaks in between. So I am quite conscious of that, so with the 3D [referring to students in one class of the Sec 3 level] I do sidetrack a little bit sometimes — we talk about our shoes, hair, I think that helps.*

Our observations of her sessions showed that with deliberate change in her practice, the learning environment in her class had changed, with sessions becoming more intense and fluid. Students in her class performed well in the summative assessment task. X explained her journey through the SCX game-based learning program by a narrative that highlighted her personal growth and change:

> *You know how a pearl is formed—you have a grain of sand—it's like you know you have an oyster and then there is this grain of sand comes in and you are really irritated by it, really annoyed that there is this grain of sand but of course you kind of allowed that a grain of sand to come in. … some excretion and that's how the pearl is formed when it hardens. And then you become this pearl, that's something precious, and I think my journey is something like that.*

4.2 Teacher Y's journey enacting the SCX program

Teacher Y had completed a year of teaching at the time of the intervention. Her desire to become a teacher was to help struggling students by providing the 'right teaching method'. She placed great importance on education due her personal experiences and believed that education was the route to earning: 'I see myself, my parents are not educated so… I know how hard it is. … It really motivated me to work harder because in Singapore you have to study hard and when you study you get a qualification and that's how you earn money.' Y wanted to try game-based learning pedagogy in her class for aiding students to 'visualize something about governance and stuff.'

It was observed that in Teacher Y's school there was a perceptible lack of student involvement in the game. This could be attributed to the school choosing to use the web version of the game over the iPhone version. An additional reason might be that the number of hours students were permitted to play the game was rather low. As a consequence, in the first session, only one student had played the game in her group. Despite this, Teacher Y was content with her lesson and asserted that '… today's lesson did go quite okay for me, but it's just that I didn't get the feedback.' She identified the strength of her session as her ability to carry out the entire session with only one student's experience of the game.

Y also prepared her lesson in advance, and she would pose some questions to the students during dialogic sessions. She was confident that these questions would help her students to understand the game better. She conveyed disappointment in not being able to go through all the questions. In the interview after her first dialogic session she said, '… wasn't very happy… a bit sad, that my lessons did not go as planned. Today … for example I had some questions right, this is what I wanted them to tell me, but …' As the intervention progressed, Y's sessions did not improve, and the students remained largely uninvolved both in game-play as well as in the dialogic sessions. She justified their non-involvement by blaming it on excessive homework that the students had to deal with. Y was content that the SCX sessions gave her a platform where she could use a curriculum innovation (A) developed by her school. She was convinced that by combing the two programs, she would be able to meet the requirements of the school as well as that of the SCX project. From Y's interview after the third dialogic session, it was apparent that she found this mixture fruitful. She commented:

> *I'm feeling better than the previous lesson because … What I'm feeling is that I can make use of the Statecraft game at the same time I can infuse my A package in order to understand governance. So I am feeling very good.*

We observed that throughout the SCX program, Y did not reflect deeply on her pedagogical practices and dwelled largely on the game-mechanics. Throughout the program, Y held on to her belief about planning content for her lessons. During the interview after her third dialogic session, she was asked what she would tell her colleagues if she was asked about the SCX project. Her response highlighted the importance of preparing in advance.

What I basically told them is we have to prepare … because the questions are very important for dialogic. … So it's about questions, it's about preparation in a different way – not like worksheets, but it's the other way – it's more very high-level thinking.

As Y progressed to the fourth dialogic session, she found that dialogue facilitation required more preparation and that her students and she herself was bored with it. For Y, dialogic sessions were inflexible and restricted her to asking questions in each session. After the final dialogic session, she was asked if the SCX curriculum had delivered what she had wished. Her answer was:

I felt about the facilitation being more work for me because my students are not the type that talks much. … they do discuss about the game, but I have to always constantly make sure that they … don't go off line ... after two sessions we realized that we must do something different … I mean we cannot be repeating the same thing again and again, and it can get boring for us.

We observed that Y's sessions were very teacher centric and students spent most of their time answering questions posed to them. Another concern that Y shared in the final dialogic session was that despite her reminder to students, they 'did not take down notes' during the session, hence, they would 'forget'.

To summarize, Y's journey enacting the SCX program was relatively uneventful, characterized by dealing with challenges. As observers, we sensed that, even towards the end of the program, Y was not successful in engaging her students dialogically. The performance of her students in the summative assessment task was not very different from the control group students.

5. Discussion and implications

The two teachers had different motivations for entering the teaching profession. They had different expectations of their students, due in part to their personal background. During their participation in the SCX game-based learning program, X and Y had different experiences: X metamorphosed and Y wanted the SCX program shortened. In the case of X, an observable change in her classroom practices was evident. During interviews, she could consciously reflect on what it meant to be an effective teacher. For Y, her classroom practices did not change materially nor was she successful in reflecting on her practices. This section highlights the key reasons that contributed to making the journey of the two participants so different.

5.1 From adhering to lesson plan to real time 'thinking on the feet'

Classrooms are places that are in a constant state of flux. Teachers often make the mistake of meticulously planning their lessons in advance and then getting discouraged if their plan is not fully adhered to. The idea of teachers having a lesson plan ready is culturally rooted. During pre-service training, teachers are encouraged to plan their lessons in detail. Often, schools expect teachers to account for everything that s/he will do in the form of detailed lesson plans. However, this strategy works against the dialogic spirit that needs to be nurtured in 21st century classrooms where students' independent thinking is valued. Teachers need to move away from strictly following lesson plans to conducting sessions with greater openness of purpose. This requires teachers to work with student's ideas and to create an active learning environment in the classroom. In the SCX intervention program, initially both X and Y acknowledged following a lesson plan. However, X soon realized that this strict structure was holding her back, and she moved away from it. In contrast, Y held on strongly to the idea and prepared questions for the sessions. Due to Y's strong adherence to the notion of lesson plans and X's flexibility to 'go with the flow', it was easier for X to modify and adapt her practices.

5.2 Shifting from teacher centric to student centric classrooms

The teacher-centered classroom is associated chiefly with the transmission of knowledge. At the beginning of the SCX program, both X and Y saw themselves as central figures in the classroom. We saw evidence for this conception of teaching in the way both the teachers structured many classroom activities as teacher-directed tasks and walked students through them. Shifting responsibility to students was not easy for the teachers. X struggled to find a balance between how much she needed to talk and where students needed to take a more active role in their learning. As the SCX intervention progressed, X came to believe that her central responsibility as a teacher was to be a facilitator. She modified her scaffolding strategies and realized the flexibility that dialogic sessions offer. Our observations of her sessions provide evidence that she went through a deep process of change, whereby she came to understand what it really meant for a teacher to be a facilitator. On the other hand, Y continued her sessions using the format of asking prepared questions. This shift of mindset, from thinking that teachers are the source of knowledge to acknowledging that learning occurs when students communicate amongst themselves in the class, is crucial for 'authentic learning' to occur.

5.3 'Letting go of control' to facilitate more active student learning

Studies in various fields have reported teachers' difficulty in letting go of control in the classroom [10]. Efforts to give up control are often associated with fundamental and difficult shifts in conceptions of teacher and student roles. To shift to a facilitator's role requires time, awareness, and being able to reflect and question one's practice and assumptions in a given context. With the strategy of reflection and reflexivity adopted during post dialogic session interviews, X began over time to see value in getting students' to question, argue, and negotiate amongst themselves. She described it as 'letting go of control' in her classroom, where she would 'step down' from the authority position of a teacher and be a facilitator at the same level as the students. She expressed her eventual comfort with this strategy by saying, 'I am very comfortable letting the kids talk.' For game-based learning pedagogy to make its impact, it is necessary to allow students greater agency over their learning, with teachers acting as facilitators. It was difficult for Y to think of any other ways of facilitating a dialogic session except posing ready questions to the students. Consequently, she engaged in substantial teacher talk during most sessions, while students listened. The outcome was that active student learning did not place, and there were not many takeaways for them.

6. Conclusion

To summarize, X and Y participated in the SCX game-based learning program where they sought to learn how to facilitate student dialogic sessions as a complement to students playing the game. To support teachers in the implementation of game-based learning pedagogy, formal training and hands-on experience were provided. Using the RRGA model, opportunities for reflection and reflexivity were built in. This provided constant support and guidance to the participants during the program. Both the teachers received same training materials and had support from their schools, yet both teachers enacted the curriculum differently. These differences were largely due to differences between X and Y as teachers and as people and to their personal situations. Despite these differences, both teachers had manifested a common concern for students' performance in standard tests, and they felt that 'drilling' the students was the best way out.

Based on our observations of the SCX game-based program, we realize that professional development that engages participants in narrative inquiry can help them in reflecting concretely on their practices. In doing so, professional development becomes meaningful to the participants. In this study, the RRGA model for professional development worked differently for the two teachers. While it helped both teachers develop an understanding of what game-based learning entails, the intricate connections that teachers make with the pedagogy plays a significant part in how they enact the curriculum. It should also be noted that professional growth requires substantial time and effort and that teachers need to be given time to reflect on their practices. Any approach to professional development must thus be flexible and continuous. Game-based learning, which is finding its way into schools and classrooms of the 21st century, needs to be supported in ways in which its benefits can be maximized. The success of game-based learning depends largely on the ability of the practicing teachers to take full advantage of it. Teachers needed to be convinced that "alternatives to present practice exist and are worth trying" [11].

Based on our study, we find that for teachers to effectively appropriate game-based learning pedagogy, teachers must enact a shift (a) from adhering to lesson plans to real time 'thinking on the feet', (b) from teacher centric classrooms to student centric classrooms, and (c) from dominating the class with teacher talk to facilitating active learning. Supporting teachers' professional growth at the start of the program and during its implementation is critical to an enriching and sustainable use of game-based learning in classrooms.

Acknowledgements

The work reported in this paper is supported by research grant OER 2/11 CYS awarded to the second author. We wish to thank Tien Yuan and Liu Qiang for their constant help.

References

[1] Chee, Y.S. (2011). Learning as Becoming through Performance, Play, and Dialog: A Model of Game-based Learning with the Game Legends of Alkhimia. *Digital Culture and Education, 3*(2), 98–122.

[2] Sandford, R. & Williamson, B. (2006). *Games and learning: A handbook from Futurelab*. Futurelab.

[3] Chee, Y.S. & Mehrotra,S. (2012). Reflective, Reflexive Guided Appropriation: Facilitating Teacher Adoption of Game-based Learning in Classrooms. Paper submitted to the *European Conference on Game-Based Learning*, October 2012, Cork, Ireland.

[4] Sullivan, P. (2012). *Qualitative Data Analysis: Using a Dialogical Approach*, Sage, London.

[5] Daiute, C., & Lightfoot, C. (2004). Theory and craft in narrative inquiry. In C. Daiute & C. Lightfoot (Eds.), *Narrative analysis: Studying the development of individuals in society* (pp. vii-xviii). Thousand Oaks, CA: Sage.

[6] Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. San Francisco: Jossey-Bass. York: Collier Books.

[7] Dewey, J. (1985). *Ethics* (John Dewey: *The* Later Works, 1925-1953, Vol. 7). Carbondale: Southern Illinois Press. (Original work published 1932).

[8] Riessman, C. (1993). *Narrative Analysis.* Sage. London.

[9] Denzin, N.K. and Lincoln, Y.S. (2005). *The Sage Handbook of Qualitative Research*, (3rd edn.), Sage, Thousand Oaks.

[10] Wood, T., Cobb, P. & Yackel, E. (1991). Change in teaching mathematics: a case-study. *American Association Research Journal, 28 (3),* 587-616.

[11] Anderson, L. M. (1989). Implementing instructional programs to promote meaningful, self-regulated learning. In J. Brophy (Ed.), *Advances in research on teaching: Vol. 1. Teaching for meaningful understanding and self-regulated learning (pp. 313-343).* Greenwich, CT: JAI.

The Effect of Prior Knowledge and Flow on the Understanding of Poetry in Game-Based Learning

Hsin-Yi LIANG, Sheng-Yi WU and Sherry Y. CHEN*
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*sherry@cl.ncu.edu.tw

Abstract: Poetry is the most concise language, which is full of creativities. Reading poetry enhances imaginations. However, many teachers find it hard to teach poetry because of the lack of pedagogical role model and the difficulties of explanation. Moreover, the difficulties to interpret the meanings of poetry also decrease student's learning motivation. On the other hand, game-based learning has been widely used to promote student learning. Thus, to improve students' motivation and the level of their understandings of poetry, this study uses a game-based learning environment to guide students interpreting poetry. Previous research found that prior knowledge and flow have positive effects on student learning so this study also examines the effects of these two factors on the understandings of poetry in the game-based learning environment. The result shows that game-based learning improves students' motivation to read poetry, but the ignorance of descriptions and the difficulty of games should be concerned to improve learning effectiveness.

Keywords: Poetry, Game-based learning, Flow, Prior Knowledge, Game Experience

1. Introduction

Poetry is the most concise language (Müller-Zettelmann, 2005). Reading poetry not only enhances imaginations, but also provides a different view of everyday experience (Harbus, 2002).More importantly, poetry evokes the empathy and self-discovery (Lee, 2000). In other words, poetry is valued for moral education in humanities (Kuzu, 2010).

Although poetry is worth in many aspects, it is difficult to teach for many teachers (Cao, 2009). Hughes (2008) notes that the lack of pedagogical role models and difficulties to explain poetry are main issues in education. Besides, hard to interpret the meanings of poetry decreases students' learning motivation (Hughes, 2008; Cao, 2009).Therefore, to promote student learning, digital technology is employed in poetry education (Müller-Zettelmann, 2005; Black, 2008; Cao, 2009). Indeed, a new form of poetry writing has been developed. This new genre of poetry is called digital poetry. Many poets believe that digital poetry motivates student reading poetry for its playfulness and visual attractions (Freitag, 2008). However, because digital poetry is written by complex sign systems, it is even harder to understand when comparing to traditional poetry. As a result, it is argued that digital poetry is able to promote poetry (Schreibman, 2008).

Therefore, to motivate students learning poetry, there is a need to address the importance to improve the understanding level of poetry. From the theory of empathy, it is hard to understand poetry due to the gap of emotional experience between poets and students (Lee, 2000; Müller-Zettelmann, 2005). Furthermore, previous studies showed the positive effect of prior knowledge in learning (Dumas, 1999; Quaiser-Pohl, 2005; Moos, 2009). Therefore, to bridge the gap between the differences of emotional experience and prior knowledge, game is able to reproduce similar experience from narratives (Ryan, 2009). Moreover, game also avoids ambiguity by providing clear goals and feedback (McDougall, 2008). To this end, game-based learning has been widely used to promote positive learning effects (Cao, 2009; Papastergiou, 2009). In this vein, this study designs a game-based learning system to build students' interpretation skills of digital poetry. Indeed, the game-based learning system is expected to improve the level of students' understandings of poetry.

In brief, empathy may play a role to understand poetry. Therefore, this study also examines how the empathy affects students to interpret poetry. Moreover, there is a need to quantify the empathy level. Because flow refers to a state which a player is totally involved in the reproduction of similar experiences in games (Csikszentmihalyi, 1977), flow is referred to understand students' empathy level in this study. In summary, this study attempts to answer the followed research questions: (1) how prior knowledge, including game experience and literature experience, affects flow and (2) the levels of the understandings of poetry, indeed (3) how flow affects the levels of the understandings of poetry.

2. **Methodology**

2.1 Research design and experimental procedures

Figure 1 show the research model of this study. Prior knowledge was examined from two aspects, knowledge of system used (game experience) and domain (literature experience) (Dumas, 1999). The former includes frequencies and quantities of playing games while the latter contains frequency, reading preference, and writing preference.

Based on the empathy theory, we conducted an experiment to examine the effect on flow firstly, and then evaluated how it affected the levels of understandings of poetry. The experiment was taken place in a lecture for all of the participants. In the beginning, a 10-minute talk was given to introduce digital poetry, and then they were asked to play the game and clear all the levels. There are no limitations for playing time. In other words, players were allowed to decide their own progresses. After they cleared the levels, they were asked to fill out a questionnaire to give feedback. The total time was about 60-80 minutes.

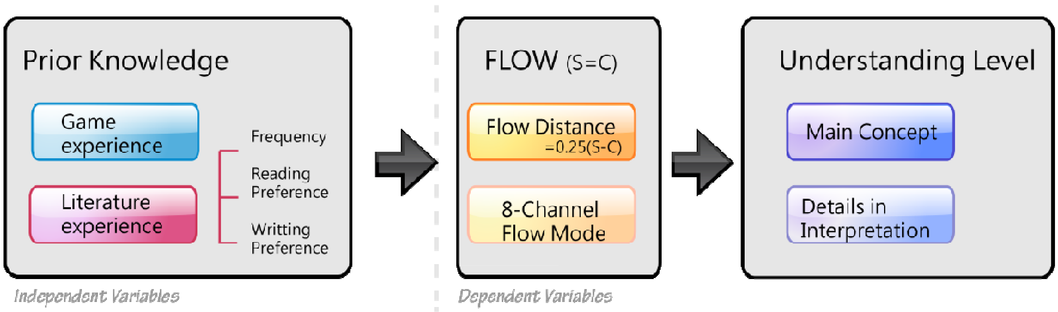


Figure1.The research model of this study

2.2 Participants

44 undergraduate students in Taiwan participated in this study. All participants had the basic computer and Internet skills necessary to play a game-based learning system but they do not have any experience and understandings of digital poetry.

2.3 Research Instruments

2.3.1 A web-based game: Skysea

A game-based learning system was designed to guide students interpreting digital poetry. The game-based learning system used narrative stories so that students' imaginations in a formal way of reading poetry are not limited.

The game included four levels, and each level was a digital poem written from single (the linguistic signs) to multiple sign systems (plus the visual and kinetic signs). Because students' actions played a key role to express the meanings of digital poetry, there were both short conversations and poems to guide students' actions in the beginning of levels (Figure2). Besides, to help students interpret the meanings of their actions, there was an animation, which provides further explanation after they cleared the level.

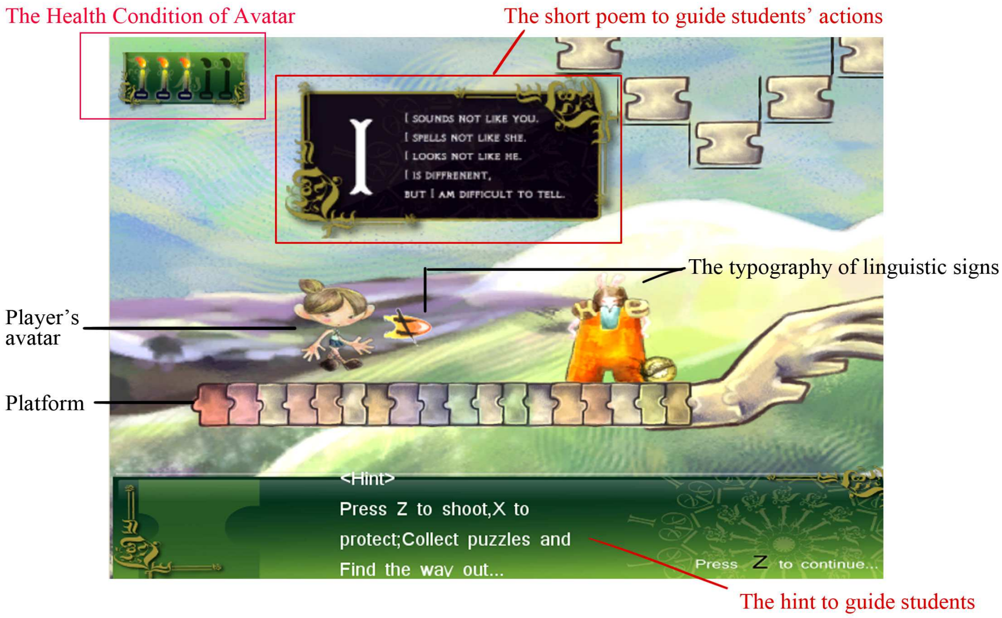


Figure2 .The opening hint of Level 2 of Skysea

2.3.2 Feedback Questionnaire

The questionnaires include three sections: (a) basic personal information and prior knowledge, including both game experience and literature experience; (b) a flow questionnaire designed by Novak (1997); and (c) 17 open questions to identify students' gaming experience and the levels of their understandings.

Regarding Section A, game experience and literature experience were assessed by self-report questions. Students were asked to rate the frequency of playing games, reading and writing. The frequency was rated in a 5-points scale. There were two questions assessing students' reading and writing frequency. The items were rated from 1 to 5, including "never", "seldom", "often", "usually" and "everyday". The frequency of game experiences was assessed based on the definitions by the International Game Developer Association (IGDA, 2012), of which classification is displayed in Table 1.

Table1. The classification of game experience

Number	Classification	Playing frequency	Game quality	Players' number (N)
1	Non-game players	Never	0	8
2	Casual players	Half hour per day	>=1	14
3	Mid-core players	Less than 1 hour per day	>=1	13
4	Core players	More than 1 hour per day	>=1	7
5	Hard-core players	Over 3 hours per day	>2 games at a time	2

The preference was assessed by 4 multiple-choice questions. Two questions asked students to select their preferred game platform and genre. The other two questions were related to the publication type and writing preference. For a statistic purpose, the literature genres were collapsed into"1" category and the "non-literature" genres were collapsed into "0" category. In the next section, the questions were adapted from Novak (1997)'s flow questionnaire. There were two parts in this questionnaire. The first part assessed students' skills and challenge from ten 9-points scale questions. All questions were rated from 1 for "strongly disagree" to 9 for" strongly agree". The last part assessed each channel of flow mode.11 questions were rated in 9-points scale from bipolar statements. The mean score of each channel was calculated for further data analysis.

The last section examined players' gaming experience and the levels of their understandings. Students were asked to answer five open questions, which are concerned with the meanings of poetry. Their answers were assessed by the details, logic rations and relationships with the main concept of the game. The total score was from 0 to 10.

2.4 Data analysis

The raw date of questionnaire was filtered prior to the statistic stage. The independent variables of this study were the participants' game experience, literature experience, reading preference, and writing preference. The dependent variables were flow and the levels of the understandings. The game experience was classified from 1 to 5 (Table1).The literature experience was summed by the score of frequencies and preferences. The flow was calculated from Pearce's flow equation (2005): Flow distance=0.25 (skill-challenge). And the scores of the levels of the understandings was summed by the questions mentioned in Section 2.3.2.

The data were analyzed by the Statistical Package for the Social Sciences (SPSS) for Windows (release 18.0). The Analyses Of Variance (ANOVA) was used to test the significance of each attributes. The Scheff's method was used for the posteriori comparisons because it was suitable for different numbers of the sample (Qiu, 2003). Additionally, ANOVA was used to test if there was a significance among the mean scores of 8-channel flow modes. Pearson's correlation was applied to examine the relationships between literature experience and flow and the levels of understandings.

3. **Results and Discussion**

3.1 The effect of prior knowledge on flow

3.1.1 Game Experience

The classifications of game experience were displayed in Section 2.3.2. According to Table1, causal players and mid-core players can be regarded as novice players while core players and hard-core players can be considered as experienced players. As showed in Table 2, there was a significant effect of game experiences ($F_{(4, 39)}$ =3.86, $p<0.05$). The post-hoc analysis showed that the flow distance of hard-core players was more than causal players ($p=0.03$) or mid-core players ($p<0.05$) but not between non-game players and core-players ($p>0.05$). According to the equation of flow distance mentioned in Section 2.5, novice players were closer into flow state than the non-game players and experienced players.

Table2. ANOVA Analysis of Game Experience

	SS	df	MS	F	Sig.
Between	1.931	4	.483	3.859	.010
Within	4.879	39	.125		
Total	6.811	43			

To further explain the result, another post-hoc analysis was used to examine the differences of skill, challenge, and four bipolar channels of flow mode. From the result, there were no significant difference of challenge, flow, and arousal channel. However, the hard-core players were better in control than non-game players, causal players, and mid-core players ($p<0.05$). Moreover, the hard-core players were less anxiolytic than the non-game players and novice players ($p<0.05$).There was also a significant difference of skills among the hard-core players and causal players and mid-core players ($p<0.05$), but not between non-game players ($p>0.05$). However, there was no significant difference of the above attributes between hard-core players and core-players.

The result showed game experience improved players' ability of control, indeed decreased player's anxiety during gamming. This result supports game experiences affects game enjoyment and performance (Dumas, 1993; Quaiser-Pohl, 2005; Richard, 2006). However, there was no significant difference between hard-core players and core players. A possible reason might be the difficulty of the design of design of this game. Because the hard-core players and core players have better performance in game, the difficulty might be too easy to compare the differences of skills.

Based on the equation of flow distance, flow is negatively related to the difference between skills and challenges. Therefore, because game experience is positively related to game performance, the experienced players may have larger flow distance because of insufficient challenges for them. Richard (2006) also found that the game experience was positive related to the achievement in motivation theory. In other words, the experienced players were more involved to challenge their skill limitations. Therefore, the hard-core players and core players were harder into flow than the novice players in this game.

In contrast, the insufficient skills of non-game players decreased their flow. The feedback from the questionnaire also supports that non-game players have more difficulties during their playing, such as hardly to clear a level, easily to die, or difficulty to control the avatar. Moreover, some players complained the over-challenges made them feel frustrated and nervous during the process of playing games

Our results showed that there were relationships between game experience and flow. Because experienced players tended to challenge their skill limitations in games, they might need more difficult challenges to be satisfied. However, for those players who lack such skills, there is a need to provide more help to improve their control abilities. Therefore, to improve players' flow in gaming, indeed to improve the learning effect, it was important to design a game with adaptive challenges for different players.

3.1.2 Literature Experience and Preferences

Prior domain knowledge also influenced flow (Dumas, 1999). Therefore, the effect of literature experience and flow was investigated. Pearson correlation showed no significant correlation between literature experience and flow (r= 0.20; p >0.05). The results of ANOVA also showed there were no significant effects of both reading preference ($F_{(1, 42)}$ =1.04, p>0.05) and writing preference ($F_{(1, 42)}$ =0.14, p>0.05).

As showed in Table 3, there were only two poetry readers in this study. As mentioned in Section 1, there was a gap of interpretation between poetry and other literature genres. Many poets also claimed that reading poetry required the association abilities and empathy, which was different from the other genres of literature (Harbus, 2002; Müller-Zettelmann, 2005; Freitag, 2008; Schreibman, 2008; Kuzu, 2010). This may be why there was no significance between literature-readers and non-literature readers.

Similarly with reading preference, only one student wrote poetry in her leisure time. Because of the gap of interpretation between poetry and other literature genres, there was no significance between literature-readers and non-literature writers.

Table3. The participant's number of literature preferences

Reading Preferences				Writing Preferences			
Non-literature(N=14)		Literature(N=30)		Non-literature(N=31)		Literature(N=13)	
Comics	1	Poetry	2	Dairy	31	Poetry	1
Newspapers	3	Classical Literature	4			Novel	1
						Commentary	11
Non-literature Novels	9	Literature Novels	24				
Prose	1						

3.2 The effect of flow on the levels of the understandings of understanding of poetry

As showed in Sections 3.1.1 and 3.1.2, game experiences affected flow, but the effect of literature experiences still needed to be identified. In next step, the ANOVA analysis was used to examine how flow affected the interpretation of poetry. The results showed that there was no significant differences of the levels of the understandings between flows ($F_{(28, 15)}$ =0.90, p>0.05).

The result shown in Section 1 was against to the poets' hypnosis. However, the result might be argued because of the low levels of understandings (M=2.25, SD=2.58). 77.3% of students' scores were less than the half of maximum scores (Max. =10, Min. =0). This finding suggests that most of students were able to understand the main concept. However, they might have difficulties to explain the meanings in details. According to the feedback of questionnaire, a possible reason Is that some students tended to ignore the descriptive words in games. As a result, their ignorance of further details of meaning affected the levels of their understandings. Thus, there is a need to prevent players to skip important clues.

3.3 The effect of prior knowledge on the levels of the understandings of poetry

As described in Section 3.2, the levels of the overall understandings might be too low to compare a difference between users with different levels of prior knowledge. Thus, the ANOVA showed no significance in this stage.

3.4 *Summary*

Figure 3 displays a framework based on the finding of this study. Our study showed there was a positive effect of game experience on skills. The causal players and mid-core players had better flow than the non-game players, core-players, and hard-core players. This result supported flow theory (Csikszentmihalyi, 1977), which indicated that flow was negatively related to the difference between skills and challenges (Pearce, 2005).

This study also found that there was a gap of interpretation abilities between poetry and the other genres of literature. Although there was no significance between literature experiences in this study, the lack of poetry experiences might be the reason. Therefore, the effect of poetry experience still needed to be identified.

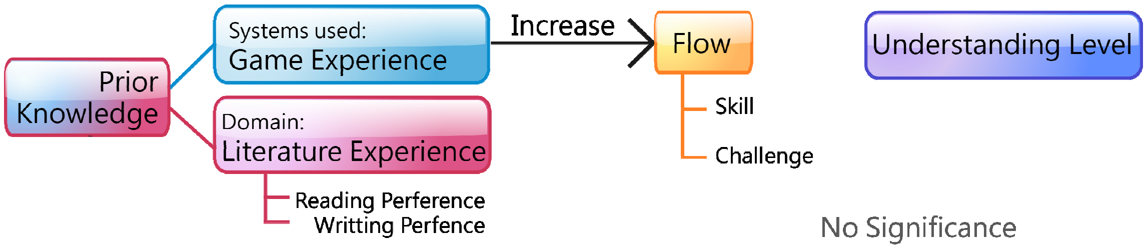


Figure 3.A framework based on the findings of this study.

4. **Conclusion**

In this study, a game was used to build students' interpretation skills of poetry. Moreover, we conducted an experiment, of which results showed that game experience was related to flow. Besides, flow is negatively related to the difference between skills and challenges. To be more specific, players with higher skills were satisfied with higher challenges. Due to the positive effect of game experience on game performance, the differences between game experiences should be considered to improve the flow. Moreover, the game was able to improve the learning effectiveness in game-based learning. Therefore, there may be a need to provide adaptive design, which provides suitable challenges for each student (Killi, 2005; Liu, 2009; Yun, 2010).

In contrast, there were no significant effects of literature experience and preference on flow. Interpreting poetry required more association ability, creativity, and empathy than other literatures (Harbus, 2002; Müller-Zettelmann, 2005; Freitag, 2008; Schreibman, 2008; Kuzu, 2010). Furthermore, this study also found that most students were non-poetry readers. Therefore, it is important to understand how non-poetry-readers interpret poetry so suitable game design can be provided for them to improve their learning effectiveness.

Finally, though there was no significance on the level of the understandings in this study, the effect of both prior knowledge and flow need to be identified because of an overall low understanding in this study. Moreover, there are some limitations of this study. The first limitation is the small sample. Indeed, there were unbalanced numbers of classifications of each attribute. Additionally, this study examined the effects of game experience on flow. However, because this game was an ARPG, players' different game experience tmay affect their performance, too.

To the end, this game demonstrated a way to guide students understanding poetry, and not to limit their imagination. It also demonstrated a way to promote poetry. Such evidence is not only helpful to develop game design for poetry education, but also is useful to understand how to develop personalized game-based learning systems.

Acknowledgements

The work presented in this paper is funded by National Science Council, Taiwan (NSC 101-2511-S-008 -010-MY3, NSC 99-2511-S-008-003-MY2, NSC 100-2631-S-008-004).

References

[1] Black, J. & Smith, K. (2008). Inspired by the poetic moving image. *Art Education, 61*(2), 9-25.
[2] Cao, Y. (2009). A web 2.0 personal learning environment for classical Chinese poetry. *Lecture Notes in Computer Science, 5686*, 98-107.
[3] Csikszentmihalyi, M. (1977). *Beyond boredom and anxiety* (2nd ed.). San Francisco: Jossey-Bass.
[4] Dumas, J. & Redish, J. (1999). *A practical guide to usability testing* (3rd ed.). UK: Intellect Ltd.
[5] Freitag, K. & Vester, K. (2008). *Another language: Poetic Experiments in Britain and North America*. London: Deutshe Nationalbiliothek.
[6] Hughes, J. (2008). The 'screen-size' art: Using digital media to perform poetry. *English in Education, 42*(2), 148-164.
[7] Harbus, A. (2002).*The life of the mind in old English poetry*, New York: Rodopi.
[8] IGDA.International game developer association. Retrieved, 2012, from http://www.igda.org/
[9] Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *Internet and Higher Education*, *8*, 13-24.
[10] Kuzu, F. (2010). Training through poetry within the sample of Nabi's "Hayriyye". *Procedia Social and Behavioral Sciences*, *9*, 1030-1035.
[11] Lee, S. (2000). *Western history of aesthetic*. Taipei: Suhsin Publisher.
[12] Liu, C. (2009). Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *International Journal of Human-Computer Interaction*, *25*(6), 506-529.
[13] McDougall, J. & O'Brien, W. (2008). *Studying videogames*. Poland: Auteur.
[14] Moerkerke, G. (1996). *Assessment for flexible learning: Performance assessment, prior knowledge state assessment and progress assessment as new tools*. London: Jessica Kingsley Publishers. doi:Oxford
[15] Moos,D. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition Learning, 4*, 197-216.
[16] Müller-Zettelmann, E. & Rubik, M. (2005). *Theory into poetry: New approaches to the lyric*. New York: Rodopi.
[17] Novak, T. (1997). *Measuring the flow experience among web users*. Interval Research Corporation,
[18] Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computer & Education, 52*(1), 1-12.
[19] Pearce, J. (2005). The ebb and flow of online learning. *Computers in Human Behavior, 21*, 745-771.
[20] Qiu,H.(2003).*Quantitative research and statistical analysis: SPSS (PASW) data*, Taipei: Wunan.
[21] Quaiser-Pohl, C. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences*, *40*, 609-619.
[22] Richard M. Ryan, C. Scott Rigby & Andrew Przybylski. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*(4), 344-360.
[23] Ryan, M. (2009). From narrative games to playable stories toward a poetics of interactive narrative. *StoryWorlds: A Journal of Narrative Studies*, *1*, 43-59.
[24] Schreibman, S. & Siemens, R. (2008). *A Companion to Digital Literary Studies*, Oxford: Blackwell.
[25] Yun, C. (2010). *PADS: Enhancing gaming experience using profile-based adaptive difficulty system*. Association for Computing Machinery.

Elementary Students' Concepts and Strategies in a Digital Stock Investment Game

Euphony F. Y. Yang [a*], Zhi-Hong Chen[b] & Tak-Wai Chan[a]
[a]*Graduate Institute of Network Learning Technology, National Central University, Taiwan*
[b]*Department of Information Communication, Yuan Ze University, Taiwan*
*euphony@cl.ncu.edu.tw

Abstract: To improve the design of game-based learning systems for elementary school students in the subject domain of financial management, this study investigates how students use concepts and strategy in a game-based stock investment environment. An interview approach was conducted on 23 fifth-grade elementary school students who used the system to collect their opinions on stock investment functions. The results showed that (1) students were easy to unilaterally interpret the stock descriptions and misunderstand some key investment vocabularies; (2) students' understanding were majorly influenced by significant others and media reports so that their key concepts were shaped. Based on these results, several suggestions for future development of financial management systems were also discussed.

Keywords: Game based learning, financial education, financial management, stock

1. Introduction

Recently, financial education has attracted increasing research attention because it closely links to our current society of knowledge economy in 21th century. Although financial knowledge is significant and has close relationship with students' daily lives, it does not always have positive impacts on students' attitude and behaviors [6]. This may due to two reasons. The first one is related with the complexity of the *content knowledge* itself, which might make students spend much time and effort in understanding the concept knowledge. The second one may be the complexity of *knowledge transfer*. Students can learn key concepts about financial management, but lack experiences of applying what they have learned to real situations.

To overcome the two complexities, some approaches are proposed to support financial education through information technologies [7], especially digital game technology. One example is the use of simulation games to facilitate students' financial literacy and skills [5]. Digital games not only can integrate information representation tools, but also enhance students' learning motivation [4]. When students are willing to invest their time and effort to learn, the complexity of *content knowledge* could be overcome, and students can gain better learning achievement. It should be noted that the learning achievement of students who learn with digital games are better than that of students who do without digital games [2]. Because digital games offer opportunities for students to learn, think, and make decision, which further enhance student learning [8]. Furthermore, digital games can immerse students in a learning context similar with their daily lives so that students' knowledge gained from the learning context can be transferred easier to practical settings. Thus, digital games could reduce the complexity of *knowledge transfer*.

Although digital games have great potential in students' financial learning, few studies emphasize on the investigation of how digital games influence students' learning process. In other words, most of previous studies emphasize on the outcome or learning achievements, rather than their learning process. Thus, there is a need to investigate how students learn during a game-based environment. To this end, this study investigates how students learn and practice what they have learned in a game-based environment. More specifically, the research question of this study is "*How do students learn and practice financial knowledge in a game-based learning environment?*" By doing so, we can obtain a deeper understanding about students' concepts and strategies during learning process, which could contribute to the development of game-based learning for a number of specific subject domains.

2. Method

To address this research question, this study conducts a case study so that participants' usages and feedback can be collected and analyzed in detail.

2.1 Instrument system

An instrument system is "My-Investment" system [1], whose purpose is to help elementary students learn financial literacy in a game-based environment (see Fig 1). More specifically, in My-Investment system, students' goal is to earn money for taking good care of their doggy, My-Pet. To this end, students can learn how to manage their money from a set of digital materials, and then invest their money to stock market. In particular, students' learning and practice in the stock investment is emphasized. Those functions include: basic financial knowledge, stock news board, and the table of stock prices (see Fig 2).

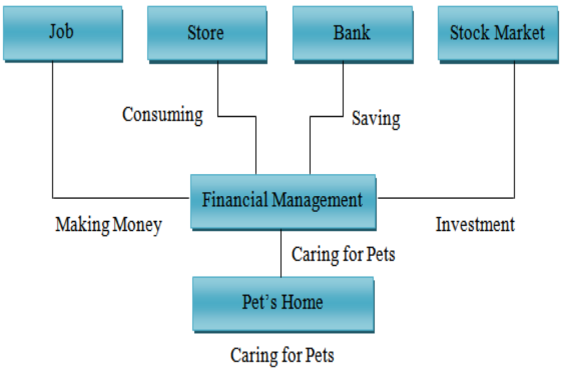


Figure 1. My-Investment system

Figure 2. Subsystem of stock investment

2.2 Participants

Participants were 23 fifth-grade primary students from three classes in north Taiwan. Every student used the system for forty minutes every week and lasted for eight weeks in school. In addition, students could also use this system after school if they were willing to.

2.3 Data collection

Harter and Harter [2] used multiple-choice to investigate whether students improved learning in a stock game. However, the format of multiple-choice might be limited in some aspects. For example, it was hard to find out what students have learned and the difficulties they encountered during learning process. Thus, this study uses an *interview approach* to collect students' data. Two trained researchers interviewed students one by one. The major question was to investigate what students' concepts were and what strategies they used in this stock game. To help students offer more detailed information about their learning process in the system, they were told to login the system to explain how they play and their reasons in this stock investment game during the interview session.

3. Result

3.1 Students' interpretation of stock information

Stock information was news that students can obtain from My-Investment. Thus, examining students' interpretation of stock information can examine what they have learned through explaining about the concepts of stock investment after reading the stock news board. There were three examples.

Example #1: Hard to understand the business terms.

R: How do you judge the auto stock may help you make money?

S2: Because its evaluation is good.

R: You feel that "sales will reach the break-even point" is good?

S2: It seems not the case.

S2 could not understand the business terms. When he saw the term "break-even point", he was likely to treat the auto stock as a positive description and worth be bought. Through the interviewer's questioning, the student started to doubt his judgment and responded the uncertain answer "It seems not the case". Without the rhetorical questioning process, the auto stock would be treated as a good stock and be the investing choice of S2.

Example #2: Positive phrases plus reverse narrative caused students' misunderstanding.

R: Why do you buy the financial stock?

S14: Because the information says that the financial company earned ten billion dollars last season which was lower than expectation. But all other stocks refer to decline, which is not good.

R: Do you know the meaning "It earned money which was lower than expectation"?

S14: It means exceeding…exceeding…

R: Is the original expectation higher than ten billion or lower?

S14: Lower. Is it lower than the expectation? (Murmur) It seems that…

R: So is the original expectation higher or lower?

S14: The original expectation should be higher.

R: So is this stock good or not?

S14: Not good at all.

Some specific terms with positive meaning in the sentence were easier treated as good situation by students and misleading them. The financial stock was thought as a deserved investment because the term "earnings" and the phrase of "performance is lower than expectation" show in the same sentence. Besides, the other three stocks are all referred to the key word "decline", which was not suitable for buying. In contract, the financial stock had no negative adjective apparently, but referred to the term "earned", which easier connected to a great situation. However, when it combined with the phrase "lower than expectation", it might be different, particularly when the interpreter misunderstood the actual meaning of the sentence. It seemed that S14 thought "lower than expectation" was a good description for the financial stock, so he responded "exceeding", which meant that the company earned money exceeding ten billion and thus deserved to buy. Not until the researcher asked him the meaning through every signal word or phrase, he was hard to discover his misunderstanding.

Example #3: Interpretation from the consumer position

R: "The auto stock published new products, and its sale performed well", "the technology of electric stock industry had not upgraded, thus their future did not have prospect", "3C stock company involved in an Embezzlement Case, so that the investors' confidence had adversely affected.", "the transportation cost magnitude of financial stock increased, and the profit performance would be reduced." Above four stocks, which one will you buy? How do you decide?

S4: I may buy the financial stock or the auto-stock. Although their turnovers are not higher than the others, but their sale performance and cost extent are good, so I won't consider the other two stocks.

R: What does good information mean?

S4: Their cost rose, so the price becomes lower.

S4 interpreted "the transport cost of financial stock increased" as cost-raising and price-reducing. In fact, as the cost increased, the price could only raise. The student was likely to misunderstand that the decreasing of company profitability was price down, and price down was a good thing. However, both cost increasing and price down are disadvantage for a company.

3.2 Major strategies within stock exchange

When the stock nature and operating mechanism were not clear, students were easy to use intuition to do stocks transaction. Their investment stressed to reduce risk and make money, so the timing of intuitively transaction for students was "buy the stock when it rises, and sell it when it drops". The method of selecting stock was alike to buy things: "stock price is the priority indicators of purchase". This might be because the growth of environmental constraints, and stock selection was "according to significant others ' preferences and media reports." We could clearly understand the intuitive thinking from the following three examples.

Example #1: Buying stock when the price rose and selling it as the price fall

S14: When the stock price decline, I sell it because I don't want to lose money. The stock price may continue downward.

R: But the price may go up quite substantially in contrast?

S14: I can buy it when the time comes.

Comparing with earning from the stock, most students cared about the deficit more. Students tended to exchange stock according to the price. Price was thought to be the crucial point in exchange stocks. The opportune moment of purchasing stock was price-rising, but otherwise it must be sold. Although stock price often changes (the price goes down this time, and may go up next time or next day), students did not want to take the risk of holding the stock. Worrying about the damage of the stock price would influence students' behaviors of stock exchanging.

Besides, it seemed that some students considered the stock ups and downs had domino effect. It meant that as stock price went down, it was down all the way. So the stock must sell it as soon as possible; if not, loss was definite. Similarly, as the stock price went up, it must continue boosting. In other words, the stock price was predictable. Thus, students would wait for the boost time to prevent disadvantage. S2 had the same idea. She referred that “I always make money, because I would direct selling it when losing”. S13 had another opinion with stock price and exchanges. He evaluated the stock price from its ups and downs, because “If stock goes up, it is very expensive, but if it goes down, the selling price should be cheaper.” A cheap stock would goes down, and I can’t afford the expensive one. Not buying any stock maybe the best choice”. Hence, for students, not buying the stock as it goes down and only buying when it goes up.

Example #2: Stock price was the prior purchase indicator

Some students treated stock as general merchandise and used cheap price as the first consideration intuitively. There were three investment types which bought the low price stock for different reasons. First, “*cheap priority and saving first*” type of students regarded the stock as general merchandise, so they considered the lower price was better. For example, S1 claimed that “cheapest stock rises and makes money easily”. In short, this type of students equated low price with saving and making money.

The second type was “*cheap priority and quantity win*”. This type of students loved cheap stock which was the same with the first type of students. However, students of “saving first” type emphasized on saving money, but students of “quantity win” type already had the notion of buying low price and selling high. The main goal of them was to buy many stocks and take advantage of rising prices to make profits. Limited to the amount of money, students only bought low price stocks to ensure buying enough, as S8 said “I would buy stocks according to the money I have. I buy cheaper stocks more, and earn the spread between selling and buying prices. If I don’t have much money and buy the expensive one, I can merely earn one thousand when the next day the share price rises a little.” The notions of two different students above were similar to small profits and quick turnover in business. In other words, buying more cheap stocks earned fewer price differences, but large amount would balance the insufficient margin of profit and obtain the maximum profit overall.

Final one was “*low price priority, but information determine*” type. This type of students also bought stock according to the price, but they tended to evaluate stock situation first. For instant, S4 determined stock by “finding the cheapest one, then read the information. If its information is not good, I may change the better one; if the better one is expensive, I may earn more money from other places. Then I come back to buy the stock later”. Although these students also considered the lower price first, they wouldn’t insist. They will also concern other aspects based on the positive information at the same time. So they were the investors who more wise and considered more aspects.

Example #3: Investment depending on the important person and media report

The stock students decided what to buy were easy affected by surrounding social environment, especial the people they contacted in daily life. The factors which influenced the stock price were various. However, students tended to trust the persons who had credibility, such as TV media, parents and teacher. When students lacked of related background, the preferences and experience of their important person were their priority even the only target. For example, when researcher asked S17 about why stock could make money but most people didn’t buy stock. He responded immediately that “they surely watch TV”. It was clear that he considered TV as the important channel to spread the information about stock deserved to buy or not, and the media usually reported stock investment defeated at the time. So he misunderstood that people had common behaviors were due to receiving the same information.

Some students used the news report in reality life as the base in their stock investment game. S10 thought the financial stock wasn't good, while the 3C stock deserved to be purchased. Because "the financial stock is disadvantage for Taiwan, since it' is not allowed by the government. In contrast, the 3C stock is good, because Taiwan sells 3C and is a 3C big country. So the 3C company makes money very soon". Student was very sure to get enough information from the reality, in spite of the game had its' own random stock information as the investment indicator and not correspond to the reality.

Personal experience and people's preference around students would also influence their predilection and misjudged. S1 referred that his father bought the 3C stock. He thought his father would not buy a money-losing stock, so buying 3C stock should be more likely to make money. Thus, he applied this criterion to the financial investment game. On the other hand, for students, in addition to credible parents, teacher was the most influential person. S14 said that "I usually buy financial stock but not others; because my teacher ever showed us the financial stock was a better choice". The stock which teacher recommend may only applicable to certain situation at that time. However, students never considered other options and merely chose financial stock teacher recommend as the only one choice when he wanted to buy stock.

4. Discussion and suggestion

4.1 Scaffolding of multiple representations and embedding financial knowledge

Students were less exposure to the stocks in real life. They were not sensitive to the business terms of the stock information in this game. Besides, they might have different text comprehensive abilities. Thus, the system was necessary to help students learn the stock and financial background knowledge, and foster their sensitivity of complex narrative. Following were three recommendations.

(1) Providing explanations or examples to help students understand the professional background and financial terms. In the present study, the students tended to decide what stocks to buy from one-sided message. But students might not interpret the message correctly. For example, students interpreted "the sales of auto stocks will reach break-even point", "the financial stock earned ten billion dollars last season which is lower than expectation", "the transportation costs magnitude of financial stock increase and the profit will be reduced," as the positive descriptions and deserved to buy. Students were unable to find out the keywords from the textual for the overall market assessment, and were not familiar with the business language. It might due to their language comprehension abilities. If the description of stock situation can provide a small dictionary for students to click and look up business terms, they might improve the shortage of situation understanding.

(2) Constructing students' correct investment concepts through the complementary cooperating between abstract concepts and the concrete operation. In this study, students who lacked of knowledge of stock market analysis and information might be difficult to assess the pros and cons, so they used the stocks potential and available information to determine trading. Thus they had herd behavior: buying stock as the price went up, and selling it as price went down. Although students preferred not to buy stocks by fear of losing money, stop-loss immediately, or not to take risk to gamble the next comeback for avoiding some huge financial loss, the learning system in the financial management game should offer proper investment content and appropriate financial instruction.

(3) Using multiple representations to represent the complex narrative context, and prevent framing effects. Students are easy to mis-interpret different explanations. For example: "this company estimated turnover increased 6% last season in original, and they

earned 10 billion finally which increased five billion than before” and “this company earned 10 billion which lower than expected" have the same meaning, but the former one makes people mistakenly believe as the blue chip stock, while the latter one is a money-losing stock. This is a framing effect. That is, using different manners of representation for the same problem, so that people have different decision-making results [3].

Except for the financial stocks game, there are many things that influence people's judgments and decision through framing effect, such as news reports. Banks encourage investment by using high remuneration advertisement and insurance provides preferential protection. Financial game design not only provides children opportunities to experience what their age can’t encounter and learn financial knowledge, but also needs to consider how to support transfer from financial management to the decision-making from different angles. By doing so, as children face various events and information processing in their life, they could foster their multi-thinking habits and make the right decisions and behaviors.

4.2 Using situational anchor to connect the learning content and game environment

The lacking of information would seriously affect the judgment of students’ stock selection. From the third examples, the results show that students’ life experiences influence their financial behavior. Real-life stock market, the recommendation of the TV media, teachers’ experiences and preference are all the information sources in this game. But even in reality, those information sources still can’t apply to different situations and timing. Hence, it’s crucial to provide sufficient knowledge for students to learn and practice in the game. Beside, how to design particular stock investment financial instruction and scaffolding in a game based learning environment for developing students’ awareness, and helping them understand the situation and meaning of various stocks, or financial investment within and outside the website. In addition, it is also important to learn how to overall analyze the stocks, and when to buy and sell stocks, and how to prevent from following the ideas of significant others to make wrong decisions.

Good connection with learning content and game environment depends on the appropriate context anchor. Knowledge created by students from the digital content in the game, where decision are often combined with the interpretation of students’ life experiences, and thus turned into resistance or power in their financial management practice. Game environment is good at producing context, while the situation is easy to anchor the problems and create further learning opportunities. Thus, the environment should promote students’ in-depth study and reflection, including whether it is a positive guidance that approaches life situations or whether the investment landscape matches the information with their real life, and deliberately produces cognitive conflict.

Students would bring some kinds of pre-existing views and knowledge to the game situations. Without taking the advantage of students’ understanding and available information in reality, we cannot control the new concepts and information and subsequently blending with the views of the outside world. Games can provide students with appropriate life context, like a vessel anchor, to connect with learning content seamlessly for a specific topic. Embedded reference information in stock trading page, such as the movie or animation included in the investment consultant's recommendations and their prediction of stock ups and downs in daily life like news broadcast. We may lead the students to engage in the financial game and have better stocks understanding. Based on the foundation, students often exposed to provide more investment knowledge and skills as a stock investment presented manner so that the opportunities for students to internalize knowledge can be increased.

5. Conclusion

This study provided students' operating behaviors of the stock concepts and practice strategies in the financial game through a content analysis of interviewed with elementary students. Students interpreted the descriptive information of stock market might occur some problems, especially one-sided interpretation or misjudge the meaning of certain words. This study also further explored the imagination of financial concepts and stock trading operation, and how those operating concepts influenced their investing practice. For example, in order to reduce risk, they followed other people to purchase the rising stock and sell the falling one, or preferred low price stocks to prevent large losses of stock decline, and treat the stock price as the prior purchase indicators. In addition, students were influenced by the significant others in their daily life, such as purchase experience and preference of a particular stock of parents, teachers, and television media, and select stocks based on these experience in the financial game, resulting in ignoring the game setting on the stock and is inconsistent with the reality and thus suffered losses.

Based on these results, this study provided further suggestions for improvement of the future development of financial management in game content and systems, including the framing effect of the learning content, scaffolding of multiple representations, complementing abstract concepts with concrete operations, and using situational anchor to connect the learning content and game environment. This study is expected to have real help for the future development of digital investment games.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (NSC 101-2811-S-008-009, NSC 101-2631-S-008-003, NSC 100-2511-S-008-013-MY3, & NSC 99-2511-S-008-002-MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

[1] Chen, Z. H., Wu, H. W., Li, Y. C., & Chan, T. W. (2010). My-Investment: Simulation Games to Help Primary Students Learn Financial Management. In S. L. Wong et al. (Eds.), *Proceedings of the 18th International Conference on Computers in Education* (ICCE2010). Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education, 548-552.

[2] Harter, C & Harter, J. F.R. (2010). Is Financial Literacy Improved by Participating in a Stock Market Game? *Journal for Economic Education, 10*(1), 21-32.

[3] Kahneman, D., & Tversky, A. (1981). The Framing of decision and the psychology of choice. *Science,* 211, 453-458.

[4] Keller, J. M., & Suzuki, K. (2004). Student motivation and e-learning design: A multinationally Validated process. *Journal of Educational Media, 29*(3), 229-239.

[5] Curland, S. R., & Fawcett, S. L. (2001). Using simulation and gaming to develop financial skills in undergraduates. *International Journal of Contemporary Hospitality Management, 13*(3), 116-119.

[6] Mandell, L., & Klein, L. S. (2007). Motivation and financial literacy. *Financial Services Review, 16*, 105-116.

[7] Way, W. L. & Wong, N. (2010, August). *Harnessing the Power of Technology to Enhance Financial Literacy Education and Personal Financial Well-Being: A Review of the Literature, Proposed Model, and Action Agend*: Paper presented at the Center for Financial Literacy Research Conference. (2010). Madison, WI.

[8] Yamashita, Y., Takahashi, T. & Terano, T. (2008). The Development of the Financial Learning Tool through Business Game. [Electronic version]. *Lecture Notes in Computer Science, 5178*, 986-993. Retrieved January 17, 2011, from http://www.springerlink.com/content/j166580u51u04112/fulltext.pdf

The Effect of Gaming on Secondary Students' Thinking, Beliefs, Creativity and Skills: A Preliminary Study in Social Studies Lessons

Chua Tee TEO[a*] & Yam San CHEE[b]
[a]*Psychological Studies, National Institute of Education, Singapore*
[b]*Learning Sciences and Technologies, National Institute of Education, Singapore*
*chuatee.teo@nie.edu.sg

Abstract: Statecraft X, a fantasy multiplayer strategic game designed to assimilate principles of governance and related concepts on citizenship played on Apple i-phones, is developed based on performance pedagogy. Two secondary three classes of students were involved in this study. The study intends to examine any significant shifts in the students' thinking and learning attitudes, beliefs, creativity and skills after playing with the i-phone game Statecraft X in their own time outside Social Studies lessons on governance. The experimental study was conducted at a secondary school in Singapore promoting the use of technology in learning. Pre- and post tests were used to check for experimental effects. Preliminary results of the study served to inform the researchers and teachers on the effects of the use of a game like Statecraft X in helping students to develop inquiry attitudes and process skills. During the intervention, teachers employed non-authoritarian, dialogical and non-indoctrinating pedagogical methods to facilitate independent thinking in students.

Keywords: Gaming, citizenship, governance, creativity, thinking, beliefs, skills

Introduction

To prepare secondary students for a world of unknowns in the 21st century means keeping up with the students' interests in games and technology, and being able to dialogue with teenage students at their mental frequencies. The previous pedagogical method of didactic teaching is probably not very appealing to the more intelligent and IT savvy students today who have tendencies to examine evidence, challenge assumptions and ask thought provoking questions. Teaching adolescents citizenship concepts is specifically not easy as adolescents may not accept modes of government as prescribed by text books. Understanding and knowing the logic behind the principles of citizenship in relation to governance are made easier if the teenagers are able to discover the principles themselves. These are key reasons that prompted the researchers to design Statecraft X, a fantasy multiplayer strategic game to be played on Apple i-phones which assimilate principles of governance and citizenship.

Statecraft X is a client-server game which lasts three weeks. Players are grouped into four factions with fictitious names, each with an ideological affinity [4]. Each player assumes the role of a free agent or active citizen in the game, acting as the first person to manage resources, perform transactions with other citizens and build the town in which s/he resides in the virtual game world. The game may be played any time as long as the student player logs in to execute actions as a virtual citizen. Students are supposed to learn the principles of governance-citizenship as they play and dialogue with their teammates, members of other factions and their facilitating teacher.

Thomas and Brown, (as cited in [3]), said that games help students to think in a manner that they would “learn to be” rather than “learn about” in traditional classrooms. In games, players take on the roles of the fictitious characters and they literally learn through empathetic thinking. Players have to use their imaginations to answer questions on “what if” when they make decisions of steps to take as they play the game. Games naturally have propensities to promote the development of creative thinking in the students playing the game. Personal attributes which are closely associated with creativity are the development of insights or the eureka moment, curiosity, the ability to make associations between and among ideas to make sense, and an inquiring attitude of problem finding, idea finding and solution finding as witnessed in creative problem solving [1; 6; 7; 8]. With the design of the game Statecraft X, it is hoped that assimilation using technology would contribute to the goal of having students to develop to *become* good citizens and creative thinkers and would not merely know or “learn about” citizenship and creativity.

The principal investigator of the project designed the Statecraft X based on the beliefs in reflective inquiry and the development of process and dialogical skills [5], performance pedagogy and Chee’s [2] belief in process philosophy for learning. The tenet is that students are dignified beings and as teachers nurture them to become mature beings in the classroom, they have to be treated with respect and most importantly, they have to be guided or led to reason and investigate the reality, and ask questions in the process of learning. Indoctrination or talking down to students will not bring about mature and thinking adults after they leave the school setting. When teachers invite the young teenagers to discuss and dialogue on problems, the adults give the youths opportunities to act or become pseudo adults. Games serve as effective scenarios for practice of such skills as no harm could result in the context of the game. Should the students make blunders or bad decisions in the game, the teacher is in control and in the worst case could simply reset the game. Thus, games also provide a safe environment for students to learn to make moral decisions as no harm is done to any living beings, and yet the students could learn their lesson in various ways since many choices are provided in the game.

1. The Statecraft X Educational Intervention Program

1.1 The Intervention Program

The Statecraft X curriculum spanned three weeks in the school curricula program in Social Studies. As students were initiated into playing the game, they would meet the teachers twice a week each week. The students played the game concurrently as they attended their Social Studies lessons to discuss their learning. Each class session was about 50 minutes. Teachers for each class asked questions of the dialogic pedagogy to focus on issues which arise when playing Statecraft X. The teacher would draw on the students’ experiences with the game and draw out their thinking and feelings in the process of executing the game and discuss the problems encountered. The teacher would discuss the students’ reflections on their learning of what good governorship entails. The results of making different choices in the game with regard to racial harmony, dealing with immigrants, preparing human resource through training, building infrastructure of the town, meeting the defence needs of the nation, managing expenditures and investments, and others [3] would be thoroughly expounded. Students would be assessed after the three weeks of intervention through an individual essay of 300 words either for entry on a blog or to give a speech on the issues concerning how to govern Singapore when they have to deal with the citizens and resources efficiently and creatively. Students were required to substantiate with evidence in the essay the issues raised and the concerns expressed in the manner befitting of good governance.

Teachers would score these essays to check if the Statecraft X had attained its functions of helping students to *become* good government and good citizens, thus fulfilling requirements of the Ministry of Education curriculum.

1.2 Procedures

The entire project was conducted in four secondary schools with six classes of secondary three students or 15 year-olds. This paper reports a pilot study with only one particular secondary school having two classes of secondary three Social Studies students. These were huge classes with 43 students of both genders each. The educational intervention program Statecraft X was first conducted with one class. When it ended, the next class then received the same intervention. Each intervention was taught by two teachers, with each teacher taking half the class in the dialogic sessions. One teacher had a little over a year of prior teaching experience while the other had three years. All the teachers in the project were trained in a 3-day workshop by the principal investigator at the National Institute of Education at the end of the previous year, that is, before the current school year began. The principal investigator helped to fine tune the statements on the self-report inventory, called the "Self-Knowledge on Beliefs, Thinking and Inquiry, Creativity & Skills (Student version), 2012" after the author designed it. Basically, it is intended to identify the attributes of thinking/inquiry, beliefs or mindsets, creative thinking and personal skills of the students before and after the intervention was administered.

2. The Instrument

The newly devised "Self-Knowledge" inventory consists of self-report statements in the areas of inquiry learning attributes, beliefs, creative thinking and skills, each with 7 to 12 items on a 9-point Likert scale, and an open-ended free response section for the students to pen any unique personal learning or thinking skills. Examples of items are "I challenge assumptions", "I develop insights &/ ideas during discussions", "I like the process of learning", "I believe that play is important to develop learning", "I derive great pleasure making new mental connections and modifying old ones", "I am able to express my ideas well", "I am able to learn through meaningful dialogues and conversations", etc.
Results of factor analyses of the "Self-Knowledge" scale are shown in Table 1 below.

Table 1
Psychometric characteristics of the "Self-Knowledge on Beliefs, Thinking, Creativity & Skills (Student's version)"Scale (n = 150)

Subscales (factors/variables)	Items	Cronbach alpha
Inquiry / Thinking Attributes	1, 4, 8, 9, 10	.84
Deep Learning	2, 3, 11, 12	.83
Beliefs in Inquiry	3, 5, 6, 7, 8	.84
Beliefs in Process (of learning)	1, 2, 4, 9	.85
Creativity / Creative Thinking	1, 2, 3, 4, 5, 6, 7, 8	.90
Skills	1, 2, 3, 4, 5	.88

The "Skills" subscale consisted of two variables: the one reported in Table 1 is comprised of personal skills to express ideas clearly in speech and writing and to be able to hold dialogues with others. The other factor was on citizenship. As there were two few items to measure the citizenship factor accurately, this factor was not included. As the number of students was not large enough for purpose of confirmatory factor analysis, it was not conducted.

3. Results

3.1 Quantitative Findings

The results of pre- and post-tests for each experimental class of students undertaking the Statecraft intervention program was analysed using independent *t* tests as not all students handed in both pre- and post-tests to their teachers. The first experimental class undergoing the intervention program reported a significant increase in the thinking or inquiry attributes of the students after the intervention [$t(60)=2.41$, $p<.05$] with a pre-test mean of 6.83 and a post-test mean of 7.43. No significant differences were reported for the other variables. It is noted that the "Beliefs in Process" factor was near significance.

3.2 Qualitative Feedback from students

Comments made by the students after the intervention at the open-ended section of the scale are grouped according to their similarities in themes and some examples are given below.

Inquiry/Thinking
I noticed that I am able to know what's more important.
I am becoming better at analyzing situations and then decide the best option for everyone.

Beliefs in the Process of Learning/Inquiry
I am able to learn more during discussion with other students &/or teachers.
I noticed that I am able to understand how learning in a playful experience helps me to understand.

Creativity/Creative Thinking
I noticed that I am able to think in a different perspective (different perspectives).
I am now able to connect ideas with other current suggestions/issues.

Skills
I am now able to apply what I have learnt to my everyday life.
I am now able to hold conversations confidently with others, both one-to-one and as a group.

Citizenship
I am better at knowing how to become a better citizen.

4. Discussion

Gaming is one of the innovative pedagogies to educate young adolescents on governance and citizenship. It is unlike the old ways of telling the students what to do to be a good citizen or government. The related concepts are discovered through inquiry and discussion as students assume the roles of the governor and citizens in a game. Failures are acceptable and mistakes may be reverted, while the players learn to adjust their decisions consultatively in a team of equals. Perhaps this is also the ideal place for the practice of true democracy, in a fictitious, friendly and safe manner.

Preliminary results of the experimental intervention show that the game Statecraft X was able to enhance the thinking ability of the teenage students significantly for one intervention class but not the other. The thinking or inquiry attributes captured in the "Self-Knowledge" scale include the students' ability to generate insights, curiosity, questions, reflections, evidence to substantiate their personal views and the ability to make sense of their learning. The other factor which was statistically almost significant was the beliefs in the process of learning which encompasses a liking for and the enjoyment of the process of learning, a belief that the process is more important than the products of learning, and finally the belief in learning to recognise patterns of connections and interactions. The study has to be replicated with more classes of Social Studies before the effects of Statecraft X may be generalized. Perhaps, control groups from other schools may be deployed to add rigour to the study.

Gleanings of the qualitative comments given by the students after the intervention program indicate that Statecraft X has been successful with individual students. Though not for large numbers of students, the Statecraft X appeared to be able to bring about the desired outcomes intended by the designers of the game. These cover having greater inquiry and thinking skills or the ability to develop new ideas, to support them with evidence, etc.; increasing beliefs in learning in a fun way and through discussions, etc.; having more creative thinking tendencies or the generation of insights, thinking out of the box, etc.; becoming better at listening to others, expressing self, dialoguing and communicating more confidently, attaining better study skills and even becoming more patient with others; and finally one student expressed that s/he is on the way to become a better citizen.

Perhaps, more items on citizenship should be included in the "Self-Knowledge" scale to identify the students' views on citizenship and governance in future studies. Paired *t*-tests, if possibly conducted, may serve as a better check for the effects of the intervention program Statecraft X. The study is on-going at the moment, and perhaps the students might be asked to fill in the "Self-Knowledge" scale again in the next semester to see if the effects of the intervention are lasting. The intervention program could have been extended to four or five weeks as three weeks may be too short a time to witness any real growth in the thinking of the students as the attitude of inquiry and discovery takes time to develop. The possible effects of the intervention could also be more rigorously ascertained if more schools and more classes could be involved. Inquiry and discovery learning is a method suitable for maturing adolescent students and this paper only reports its preliminary results.

References

[1] Carson, S. (2011). The Unleashed Mind. *Scientific American Mind, 22*(2), 22-29.

[2] Chee, Y. S. (2010). Studying Learners and Assessing Learning: A Process-Relational Perspective on the Learning Sciences. *Educational Technology, 50*(5), 5-9.

[3] Chee, Y. S. (2011). Learning As Becoming Through Performance, Play, and Dialogue: A Model of Game-based Learning with the Game Legends of Alkhima. *Digital Culture & Education, 3*(2), 98-122.

[4] Chee, Y. S., Mehrotra, S., & Liu, Q. (in press). Effective Citizenship Education Through Mobile Game Based Learning: The Statecraft X Curriculum. Paper to be presented at the *6th European Conference on Games Based Learning*, Cork, Ireland, October 2012.

[5] Chee, Y. S., Gwee, S., & Tan, E. M. (2011). Learning to Become Citizens by Enacting Governorship in the Statecraft Curriculum: An Evaluation of Learning Outcomes. *International Journal of Gaming and Computer-Mediated Simulations, 3*(2), 1-27.

[6] Piirto, J. (2004). *Understanding Creativity. Scottsdal*e. Arizona: Great Potential Press.

[7] Treffinger, D. J., Selby, E. C., Isaksen, S. G. (2008). Understanding Individual Problem-Solving Style: A Key to Learning and Applying Creative Problem Solving. *Learning and Individual Differences, 18*(4), 390-401.

[8] Willings, D. (1980). *The Creatively Gifted: Recognizing and Developing the Creative Personality*. Cambridge: Woodhead-Faulkner.

Analysis on the feasibility of introducing digital board game into classroom teaching – from teachers' perspective

Chin-Feng CHEN[a*], Po-Hsiang HUANG[a], De-Yuan HUANG[b], Chia-Jung WU[a] and Gwo-Dong CHEN[a]

[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*

**starsky1@gmail.com*

Abstract: An educational board game module integrating with textbook materials was introduced to current teachers. This module served as a teaching aid that was equipped with as set of educational equipment-Digital Learning Playground. The aim of the study was to investigate the feasibility of adding situated interaction with a board game structure into ordinary school settings.

Keywords: Board game, Digital learning playground, game-based learning, teaching aid.

1. Introduction

Board game learning has many advantages [5, 7]. Board game has a board. The board in the game is the place where players share information, show status, and interact with each other. Modern board game can be divided into Roll-and-Move, Open Movement, Worker Placement, Simultaneous Action, Role Selection and Cooperative Play [12]. The type Roll-and-Move has already commonly designed into textbooks.

Board game is extensively been used for education. Some researches combined Board Game with technical ability. And some combine Board Game with knowledge. Studies have shown that the board game to join the teaching element has a positive impact on teaching [2, 3, 8, 15].

2. Motivation & Related Work

In the school learning, many teachers put games into the classroom, to allow students to practice the new knowledge in the classroom through games. However, teachers make preparations for class contents spend very much time in thinking game. Chen (2011) used an existing board game and added learning activities that based on the theme of the board game [5], board game placed in the teaching site, and can enhance students' learning motivation and learning outcomes [8].

To help teachers combine knowledge into the board game in a short time. We have previously done a pilot study [9]. The purpose of our pilot study is translating existing English teaching materials to be Roll-and-Move type board game. Our design combine school knowledge and board game learning, and equipped with a set of educational equipment- Digital Learning Playground (DLP) [3, 5, 9, 11].

3. Instructional design & Learning equipment

We are going to try to take a board game learning system with textbook materials in classroom settings. We apply Digital Learning Playground (DLP) (Fig.1) that consists of a big screen presenting textbook materials with videos and audios, and a flat table serving as a game board that students can join in it. A PC control center, two projectors, and a Kinect device is the rest of DLP [4, 9, 11]. Kinect sensor captures the student's body movements. In additional, we design various interactive games for students using Kinect sensor. We use TPRS (Teaching Proficiency through Reading and Storytelling) model into our "Roll-and-Move" game structure. TPRS is proven that a common teaching methodology exposes learners to a sufficient authentic language with variety of comprehensive questions such as low-leveled, open-end, personal relevant questions [10, 12,14] (Fig.2).We use Kinect to achieve interactive scenario game. The interactive system runs through physical action, students enhance the relevance with game element by reality physical action. We call this Total Physical Interaction (TPI).

Total Physical Interaction (TPI) is like a visualizing version of Total Physical Response (TPR) [1] which uses body actions and movable contexts to enhance the connection between cognition and knowledge. In this instructional design, we design some interaction games from the vocabulary of textbook, such as "Wash Face." This face-washing game is trying to enhance the relationship with the phrase "wash face." After playing this game, the students will really know the action performed in the game means "wash face."

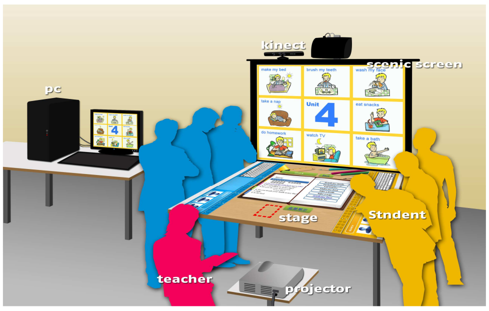


Figure 1. Digital Learning Playground Architecture.

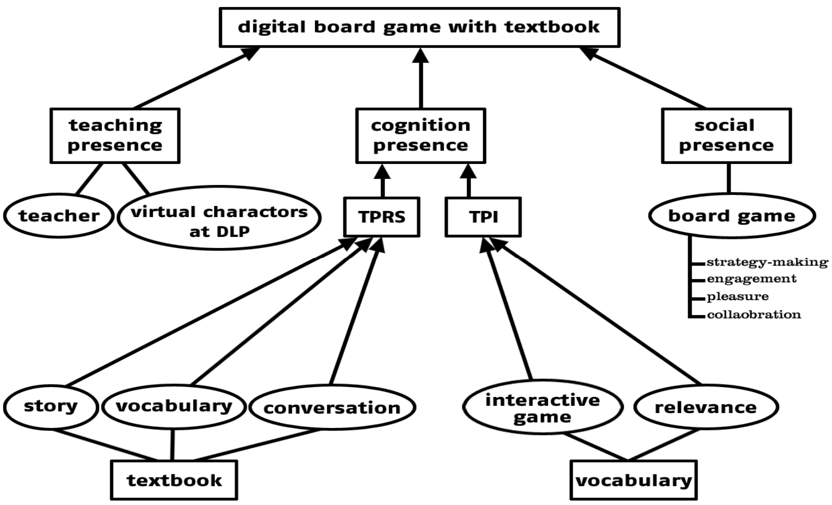


Figure 2. Design Flow with DLP Architecture

4. Experiment & Result

Game-based learning can enhance students' motivation, and students can learn in the game. After the pilot study, we confirmed that the design can effectively catch the students' attention, confidence, and satisfaction. We want to know that the design can help teachers' when they are teaching and preparing teaching materials. So we do the following hypothesis.

- Textbook materials into a Roll-and-Move board game can help teachers to enhance students' motivation in classroom.
- Textbook materials into a Roll-and-Move board game can help teachers to create more teaching situation.
- Textbook materials into a Roll-and-Move board game can provide students a chance to practice knowledge.
- Textbook materials into a Roll-and-Move board game can reduce teacher's burden of preparation.

After unstructured interviews with eight English teachers, they said that our system help textbooks vivid, and richer situational performance. Textbook materials into board game can catch students' attention to focus. Game always is a tool of teachers to promote an atmosphere because game could make students feel excitement and enhance students' concentration.

5. Discussion

In the interview, we found that the teachers all have their own teaching methods; however, most teachers would like to incorporate new technology in teaching if it could motivate students to learn.

Science and technology gradually changed people's behavior. Teachers seem to look forward to see the convenience brought by digital tools. Teachers could save more time to focus on students' progression. But most importantly, those teachers in Taiwan look forward to have a teaching aid that can sustain students' interest in language learning.

5.1 In social presence

Game assisted learning in the aspect of social presence can create human interaction, develop social skills and confidence [13, 16].

Ms. Chen expressed that confidence to students in EFL English learning is crucial. "If an educational game can make children be excited about learning and forget shyness, it would be a valid learning game design. Students with high English competency are full of confidence, while children with low or middle level feel inferior to speak out or even be put on the spot. You should avoid that letting students feel speaking English in front of everybody is sort of punishment."

5.2 In teaching presence

In this section, we would address the teachers' perspectives and expectation of using games as a teaching device. There were two types of teachers who had experience using games in teaching and who had had not.

One of the teachers, Miss Wu seldom used games in teaching because there were so many things to concern when she was delivering a lesson, such as classroom management, and time constrain. She seldom has time to jam a game in her teaching, if there is extra time, she may use games to do summary.

Ms. Lu is one of the teachers who often uses magic props, the game reference book "Give Me Five," for teaching [6]. The teacher said, "She spent so much time on game selects. The difficult part was how to combine teaching elements with simple games."A general agreement to the teaching aid with games is the aspects of simplicity. Teachers hope that games into classroom must be easy to understand.

5.3 In cognitive presence

Many teachers did not use Teaching Proficiency through Reading and Storytelling (TPRS) in teaching, however, many of them used Total Physical Response (TPR) method. In this section, we focused on how TPRS and TPI enhance students learning in teachers' perspective.

5.3.1 In our designed TPRS

Eight teachers haven't have heard about "Teaching Proficiency through Reading and Storytelling," but TPR. Miss Yeh used TPR in her class. She realized TPRS is similar to the activity of reading comprehension check. And our system incorporating their comprehension questions into a board game structure. This allows students to cope with high-level thinking problems with situations under a play-together learning atmosphere. Besides that, she also stated "Digital Learning Playground makes the textbook contents come alive with video and audio input.

5.3.2 In our designed TPI

Mr. Chen also commonly used TPR in class. He stated "TPR energizes a class and usually elicits high responses from students. Total Physical Interaction (TPI) is like a visualizing version of TPR, same as using physical action to enhance cognition, but with visualized and movable relevant contexts." The teacher showed his inclination to use TPI in language learning. He also mentioned the physical vocabulary games motivate a lot of students to try and put shyness behind. However, the operation of game play should be easier. Complicated game play caused frustration.

5.4 Challenges and Future Directions

The main concern which was addressed to most of the teachers is that the DLP equipment is suitable for small group of learning once a time. The classroom setting and the teaching flow with DLP need to be well-designed and avoid students to be idle. As to future work, some teachers also suggested our materials should be theme-oriented. As to our future implementation, the materials should be better classified in themes, such as food ordering, hobbies, and daily activities.

6. Conclusion

In this research, we design a module for Roll-and-Move board game, and provide a teaching aid which can accept by in-service English teacher. In our design, excessive number of students cannot play at the same time, and therefore this model is suitable for small-class. In order to strike a balance between play and learning, and considered that teachers is actual users. A digital board game equipped Digital Learning Playground (DLP) is constructed. Hence our teaching tool can enhance the convenience while teaching. In addition, after class the teaching tools can also assist teachers to prepare for the future courses. In the future, we will try to find a better board game-based pedagogical methodology to let students practice knowledge through board game structure.

Acknowledgements

This work is partially supported by the National Science Council, Taiwan, under the grant number NSC 101-2631-S-008-001 and NSC 101-2631-H-008-001.

References

[1] Asher, J. J. (1969). The Total Physical Response Approach to Second Language Learning. *The Modern Language Journal*, vol. 53, pp. 3-17.

[2] Bell, R. C. (1979). *Board and Table Games from Many Civilizations*. New York: Dover Publications.

[3] Caldwell, M. L. (1998). Parents board games, and mathematical learning. *Teaching Children Mathematics*, 4, 365-367.

[4] Chen, C. H., Chi, Y. L., Wu, C. J., Li, C. S., & Chen, G. D. (2010). Constructing a Digital Authentic Learning Playground by a mixed reality platform and a robot. *International Conference on Computers in Education*. 121-128.

[5] Chen, K. C., Wu, C.J., and Chen, G.D., (2011) A Digital Board Game Based Learning System for Authentic Learning. *Advanced Learning Technologies, IEEE International Conference*, pp. 25–29.

[6] Eisele, B. & Eisele, C. (2001). *Give me five*. Taipei: Caves Books.

[7] Harris, C. (2009) Meet the new school board : Board games are back-and they're exactly what your curriculum needs. *School Library Journal*, vol. 55, pp. 24–26.

[8] Hinebaugh, J. P. (2009). *A board game education.* American: Rowman & Littlefield Education.

[9] Huang, P. H., Wu, C. J., & Chen, G. D. (2012). An approach of building an engaging virtual authentic environment for classroom learning. *International Conference on Advanced Learning Technologies.*

[10] Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *JOURNAL OF INSTRUCTIONAL DEVELOPMENT*, vol. 10, pp. 2–10.

[11] Lee, W. J., Huang, C. W., Wu, C. J., Huang, S. T., & Chen, G. D. (2012). The Effects of Using Embodied Interactions to Improve Learning Performance. *International Conference on Advanced Learning Technologies.*

[12] Mayer, B., & Harris, C. (2010). *Libraries got game aligned learning through modern board games*. AMERICAN LIBRARY ASSOCIATION: CHICAGO.

[13] Piper, A.M., O'Brien, E., Morris, M.R., and Winograd, T. (2005) SIDES: A Collaborative Tabletop Computer Game for Social Skills Development. *Stanford Univ.: Technical Report.*

[14] Ray, B., & Seely, C. (2008). *Fluency through TPR Storytelling: Achieving Real Language Acquisition in School, Berkeley.* CA: Command Performance Language Institute.

[15] Selvidge, E. (2006). Journey to Egypt: A board game. *Montessori Life*, 18(4), 36-39.

[16] Zagal, J.P., Rick, J. & Hsi, I. (2006). *Collaborative games: Lessons learned from board games. Simulation & Gaming*, vol. 37 no.1, 24–40.

A Design Principle and Collaborative Learning Support Method in Game-based Learning Environment

Ryo TAKAOKA[a*], Masayuki SHIMOKAWA[a], Takuya NITTA[a], & Toshio OKAMOTO[b]
[a]*Faculty of Education, Yamaguchi University, Japan*
[b]*The Graduate School of Information Systems, University of Electro-Communications, Japan*
*ryo@yamaguchi-u.ac.jp

Abstract: The purpose of this study is to propose a design principle and framework of educational control (in particular, method for collaborative learning support) that induces and activates interaction between learners intentionally to create a learning opportunity that is based on the knowledge understanding model of each learner. In this paper, we explain the design philosophy and the framework of our game-based learning environment (GBLE) called "Who becomes the king in the country of mathematics?". In addition, we describe the method of collaborative learning support control that incorporates a "learner support agent" to support each learner and a "game control agent" to control the game into the learning environment.

Keywords: game-based learning environment, design principle, framework of educational control method, collaborative learning support method, pedagogical agent

Introduction

Many researches of education system which utilized a computer game effectively have been pushed forward [2, 3, 8, 9, 11]. Not only the study on motivational factors which a computer game has [7], but also the study on combining learning contents with game-based learning [5] effectively has been performed in this research field. Moreover, given recent progress in network (in particular, wireless), mobile and interface technologies, various types of game-based learning environment begin to be studied.

On the other hand, the research on CSCL in which the computer supports collaborative learning activity by plural learners is pushed forward [9, 12]. It is thought that the game-based learning environment incorporating the function of the computer game for collaborative learning and group learning becomes one of the effective learning environments. In these game-based learning environments, we need the following functions; for examples, a learner modeling function and a learning support function for an individual learner, a learning flow building function and a collaborative learning support for a group. We can propose a computer-supported collaborative learning environment which activates not only the interaction among learners but also the interaction among learning supporters and learners by designing and implementing these functions as pedagogical agents [1, 4, 10].

Therefore, the purpose of this study is to propose a framework of educational control that induces and activates interaction between learners intentionally to create a learning opportunity that is based on the knowledge understanding model of each learner. In this paper, we explain the design philosophy and the framework of our game-based learning environment (GBLE) called "Who becomes the king in the country of mathematics?". Moreover, we describe the method of collaborative learning support control that incorporates a "learner support agent" to support each learner and a "game control agent" to control the game into the learning environment.

1. Design principles in GBLE

The game developer and the expert about the game have been performing the various classifications and definitions about "the fun of a game" [6, 7]. Koster considers the game as a learning process which solves the given game task and skill by trial and error, and tries to solve the framework of "the fun of a game" from a viewpoint of cognitive formation of a game player [6]. He has stated that the four tasks that must be mastered exist in the learning process, and "the fun of a game" arises in the mastery process and success experience of these tasks [6]. Moreover, Malone proposed the following individual motivational factors; Appropriate challenge, Emotional fantasy, Cognitive and sensory curiosity [7].

Based on these classifications and definitions, we classified the fun which a learner feels in an educational game into the following four types: *Fun when a player achieves a goal, Fun from what a player is unable to predict, Elation when a player faces a challenging problem, Honor for the player*. These four types of "fun" in an educational game lead to the maintenance and improvement of a learner's motivation, and we believe that they give a game and the learning from it advanced power. Therefore, in the design of a GBLE, it is important that we incorporate these four viewpoints of fun into the scenes or phases of the game and the learning that comes from it.

In a GBLE, it is effective for the maintenance and the improvement of a learner's motivation to develop the support which utilizes the fun which a learner feels in an educational game. In this research, we have examined the design principle of a board game based learning environment for schoolchildren and junior high school students. Therefore, we set some design principles based on the four types of fun, which the learner feels in the educational game, in the development of the educational game design and learning support function design in the GBLE. These designs principles are shown in Table 1.

We have created an educational game design in the GBLE based on these all design principles explained above [12]. Moreover, we have designed and implemented the function and mechanism of the learning support based on the design principle and the support elements referred to literature (including Koster's suggestion mentioned in this section) as an educational control described in section 2. In addition, we accept that the case where concrete design principles differ according to the difference in the subject domain and learning style of a GBLE exists. Examination of the design principle in a different game form and leaning style is future work

2. Method of educational control in GBLE

The game developed in this study is a board game with a roulette in which there are four players (learners). The winner can become the next king of the mathematics kingdom. From the roulette, the learner receives a number to determine her/his forward movement. She/he then replies by trying to solve a calculation formula in the roulette with an unknown value. If the learner solves the problem correctly, she/he can advance only by the number of the answer. Next, the learner carries out an event, such as learning or the game, on the grid on

Table 1: Design principles of "Who becomes the king in the country of mathematics?"

- *Fun when a player achieves a goal*
  [Principle1-1]
  Existence of various types of grids
  [Principle1-2]
  Setting a reward by clearing the given event (For example, "prize" or "increasing the mark of a parameter")
  [Principle1-3]
  Selection and question of a learning subject according to the situation of each learner
  [Principle1-4]
  Setting the number of problems according to the situation of each learner
  [Principle1-5]
  Setting a time limit of problems according to the situation of each learner
  [Principle1-6]
  Motivation for fun when a player achieves a goal
  [Principle1-7]
  Setting a collaborative and competitive learning style
- *Fun from what a player is unable to predict*
  [Principle2-1]
  Setting a grid in which a special event occurs
  [Principle2-2]
  Selection of net learning style (Such as by a time trial, or having to check answers for other players)
  [Principle2-3]
  Selection and question of new learning subject
- *Fun from what a player is unable to predict*
  [Principle3-1]
  Setting the increase in the number of problems
  [Principle3-2]
  Setting a problem to which the degree of difficulty becomes high
  [Principle3-3]
  Shortening the time limit of problem solving
  [Principle3-4]
  Setting a cost when a learner fails in a learning subject
  [Principle 3-5]
  Motivation for challenge
  [Principle 3-6]
  Setting a collaborative learning style
- *Honor for the player*
  [Principle4-1]
  Setting players' ranking at the time of ending
  [Principle4-2]
  Motivation for aiming at a championship

which she/he stopped. The learner can increase the mark of a parameter (the learning power and the power of zest for living) of the character (avatar) that the learner operates by clearing the event. There are several types of grid in this GBLE: a "Learning grid," a "Zest for living grid," an "Item grid," a "Mini-game grid," and a "Special grid". On the "Special grid", every learner must stop forcibly. There is a "STOP grid" and a "TEST grid", which are special grids in the developed game environment. On the "STOP grid", the learner plays rock-paper-scissors with the computer. If she/he wins, a bonus point is given at random. On the "TEST grid", the learner must answer all the questions for each learning item correctly. If she/he makes a mistake, then she/he must return to a certain grid specified by the computer. This game is a type of educational game. Therefore, we need to control the game educationally in order to activate the interaction between learners and to be able to acquire the desired knowledge and skill for learners.

We have studied the interaction among groups, the direct support for each learner, and lesson support for a teacher as the extension of this GBLE to support the group. In this environment, we need the system configuration using agent which can support each learner or group, and interact with other learning supporter flexibly. So, we have been developing the system with agent model. In this GBLE, the learning and educational control in this GBLE is performed by two kinds of learning support agents (one "Learner Support Agent" per learner and one "Game Control Agent" in GBLE) (see Figure 1) [12]. We have designed and implemented the function of the learning support in these learning support agents based on the design principle of the fun which a learner feels in an educational game and learning support function described in section 1.

The support function of the learner based on [Principle 1-3, 1-6, 1-7, 2-2, 2-3, 3-2, 3-5, 3-6 and 4-1] is designed and implemented for the LSA("Learner Support Agent"). The LSA has three tasks: One is to determine the learning control type of the learner as a method for maintaining or improving the learning volition of the learner. The agent gives four questionnaires for a learner at random before starting a game. The learner replies with yes or no. Then, the agent tries to motivate the learner to increase the learner's learning volition appropriately. Moreover, the agent adds up the mark of four answered items and classifies it in three types (Controlling learner in the first half, Controlling learner in the latter half and No control) based on the total value.

The function of educational control in GCA("Game Control Agent") is implemented based on [Principle 1-2-1-5, 1-7, 2-2, 2-3, 3-1-3-4, 3-6 and 4-1]. The GCA performs an educational control of four players (learners) who play the game-based learning. This agent has the following global strategy: *The agent tries to check that the student has a value of more than US2 for all learning items until she /he reaches the TEST grid. The agent tries to make sure that the student stops all learning grids in which the value of the learning item is less than US1. The agent tries to ensure so that the learner can experience various learning forms in the game.*

Based on these strategies, the agent advances the game-based learning by using control rules such as questions, individual learning and group learning. Concretely, the agent receives information about the understanding state of learner and requests the next learning item from each learner support agent. The agent determines the learning item for the learner for her/his next turn and carries out the turn. When the learner needs learning control, the agent decides on a calculating formula and the answer by means of the roulette. The agent has three learning forms: personal learning in which the learner himself solves a learning problem, collaborative learning in which the learner competes or collaborates with other learners, and observation learning in which the learner learns from other learners' problem solutions. The agent chooses a learning form based on the state of the learner's understanding of the next turn and other learners' understanding.

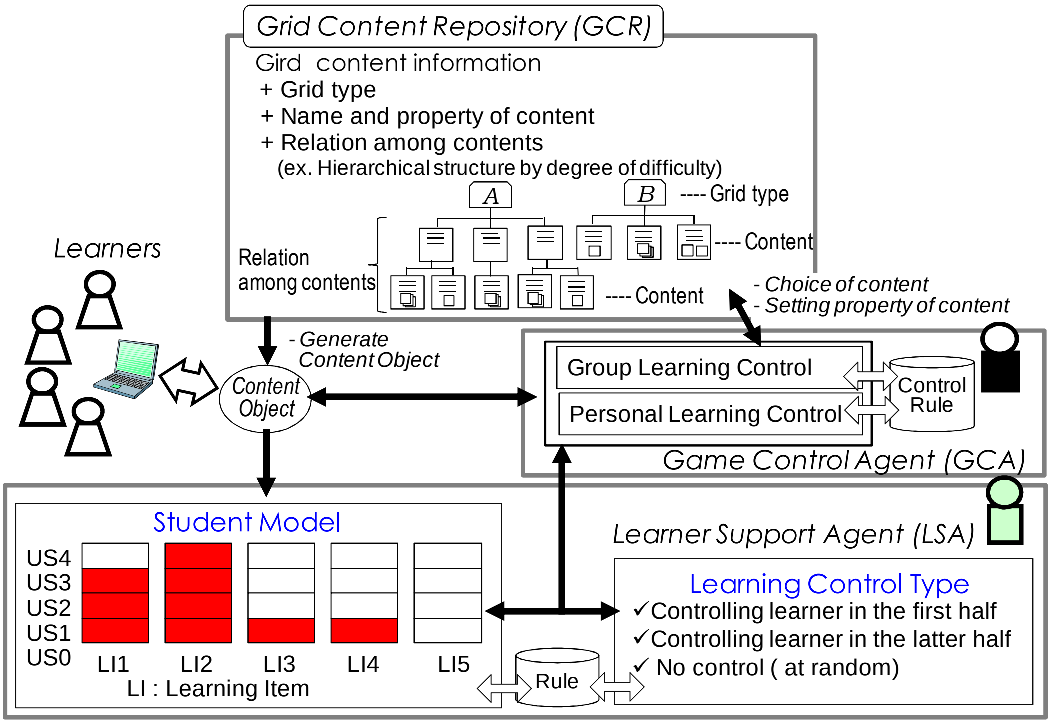


Figure 1: Task about learning and relation among GCR, LSA, GCA

Our system has the mechanism of the collaborative learning support. Furthermore, there are two interaction layers in the system. One is a "Learning Support Workplace (LSW)". The other is a "Learner Support Planning Workplace (LSPW)" to consider and decide the learning plan by interacting among agents. The flow of decide collaborative learning support decision making are as follows; Firstly, agents set the learning support goal, learning support method and, the member and the role of learning supporters. Moreover, they make a scenario of the collaborative learning support. Then, they move to LSW and start the scenario. If they can't achieve the learning support goal by the scenario, the challenge becomes the failure. Afterward, they return to the point (such as learning support goal setting process and learning support method setting process) of the factor.

3. Conclusion

In this paper, we explained the design philosophy and the framework of our game-based learning environment (GBLE) called "Who becomes the king in the country of mathematics?”. In addition, we described the method of collaborative learning support control that incorporates a LSA to support each learner and a GCA to control the game into the learning environment. As further work in the future, we need to elaborate and rebuild the method of collaborative learning support control.

Acknowledgements

This research has been supported in part by the Ministry of Education, Culture, Sports, Science and Technology in Japan under a Grant-in-Aid for Scientific Research (A) No. 24240104 (2012).

References

[1] Baylor, A.L. & Kim, A. (2004). Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and instructional Role, *Proceedings of ITS2004* (pp.592-603).

[2] Bourgonion, J., Valcke, M., Soetaert, R. & Schellens , T. (2009). Exploring the Acceptance of Video Games in the Classroom by Secondary School Students, *Proceedings of ICCE2009* (pp.651-658).

[3] Cheng, H.N.H., Wu, C.M.W., Liao, Y.C.C. & Chan, T.W. (2009). Equal Opportunity Tactic: Balancing Winning Probabilities in a Competitive Classroom Game, *Proceedings of ICCE2009* (pp.713-717).

[4] Girard, S. & Johnson, H. (2010). What Do Children Favor as Embodied Pedagogical Agents?, *Proceedings of ITS2010* (pp.307-316).

[5] Habgood, M. P. J. (2007). The Effective Integration of Digital Games and Learning Content, University of Nottingham, United Kingdom, Retrieved September 18, 2012.

[6] Koster, R, R. (2005). A Theory of Fun for Game Design, IN: Paraglyph Press.

[7] Malone, T. W. (1981).Toward a Theory of Intrinsically Motivating Instruction, *Cognitive Science*, *4*, 333-369.

[8] Muldner, K., Burleson, W., Sande, V.V. & Vanlehn, K. (2010). An Analysis of Gaming Behaviors in an Intelligent Tutoring System, *Proceedings of ITS2010* (pp.184-193).

[9] Ogan, A., Aleven, V., Kim, J. & Jones,C. (2010). Instructional Negotiation with Virtual Humans: The Effect of Social Goals on Gameplay and Learning, *Proceedings of ITS2010* (pp.174-183).

[10] Ogata, H., Yano, Y. & Matsuura, K. (2002).Human-Agent Interaction in Comuputer-Supported Collaborative Learning Environments, *Journal of the Japanese Society for Artificial Intelligence*, 17(6), 672-678.

[11] Seelhammer, C. & Niegemann, M.H. (2009). Playing Games to Learn – Does it actually work?, *Proceedings of ICCE2009* (pp.675-681)

[12] Takaoka, R., Shimokawa, M. & Okamoto, T. (2012). A Development of Game-based Learning Environment to Activate Interaction among Learners, *IEICE TRANS. INF.&SYST., E95-D(4)*, 911-920.

Designing a Typing Game with Chinese Words to Sustain Flow Experience of Children

Ellen C.C. LIU [a*], Calvin C.Y. LIAO[a], Tak-Wai Chan[a]
[a]*Graduate Institute of Network Learning Technology, National Central University, Taiwan*
*Ellen@cl.ncu.edu.tw

Abstract: In this study, we argue that the development of flow-based typing games should not only stimulate motivation but also sustain motivation of the design of typing Chinese words. In particularly, this system combined level- and star-mechanism to adjust the content of level (challenge) and the speed of typing (skill) of typing games; it possible enables children to sustain the flow experience. Findings revealed that the potentiality of flow-based typing game for sustaining the flow experience of children and facilitating the accessibility of numerous Chinese words. Besides, some implications about the experimental results were also discussed.

Keywords: typing game, flow theory, game-based learning, Chinese input

1. Introduction

In general, “Learning” is defined as the act or experience of the learners who acquired the knowledge or skill by instruction. To engage in this process of gaining knowledge or skill, the learners must be motivated [1, 2]. *Flow theory* is a way to implement and achieve motivation. According to the definition of flow theory [4], many people reach a perfect balance between challenge and skill, and then they will find a flow experience to keep the user’s experience within the user’s flow channel. This paper will focus on how to design a typing training exercise to sustain the practice and help students construct their typing skills. We will try to motivate the learners to improve their typing skills rather than decreasing the patience. In order to make an enjoyable learning environment, we design a system, named My-Pet-Typing. It is not only enhancing the typing speed of learners, but sustains their typing behavior.

2. How to learning the typing skill

Typing is a simple mechanical action. If you want to improve typing skill, you just practice again and again. Most of the way to practice the typing skill makes use of typing software which differentiated between drill-based and game-based software. Generally speaking, drill-based typing software does not have an element of entertainment like game-based software, but it usually has a virtual keyboard and the shape of a hand. It aims to help the beginner to know the standard fingering, and they will get the feedback score of speed, accuracy, and duration time (Figure 1(a)).

In fact, there are so many typing websites provide the learners to practice on World Wide Web, some of them include, in addition to many typing information, statistical chart, most typed, most mistyped, and so on (Figure 1(b)). Game-based software usually contains the little animation, colorful pictures, and multimedia, even offer instant feedback in visual and audio form. It is often most effective with particular learners who enjoy learning with game. On the other hand, some of the learners can type faster and faster when they indulge in game mode.

(a) Game interface (b) Typing information

Figure 1. Type Fu (From: http://type-fu.com/)

In Taiwan, the students learn to use the keyboards correctly at the first, in order to find the BoPoMoFo (Zhuyin) on the keyboards. After a period of time, they must learn how to type a “key sequence” of Chinese character, typing in phonetic symbols and tone. It can require the learners to type the vocabulary or an idiom, and then proceed to the next stage, and then they can complete sentences and whole article.

3. Design of Flow-based Typing Game

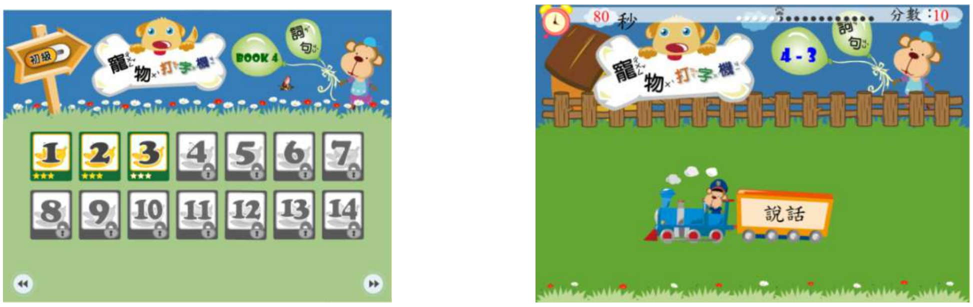
In previously study, we developed a game-based typing system, entitled My-Pet-Typing [6], was established to develop the typing ability of elementary school students, and provided a pleasant effective learning environment for typing, also can help the students enhanced the Chinese words of them.

3.1 Game design

Typing exercise is boring and dry, but it is proficiency can be built by repeating to practice. Most situations in using Chinese typing software, the learners are not interest in typing practice because of it usually repeated the same materials and played the unexciting game. Such game lacked a serial of process to cause the players have to spent long time to concentrate without interrupt. So we design a game with breaking through the barricade for the elementary school students to develop the Chinese typing skills, and record the progress of typing.

It provides the students who play at an appropriate level for their ability have to explore it by one way and accomplish to unlock it. In other words, students have to reach the lowest standard for passing the mission, then lock the next mission and get the basic reward in My-Pet-Typing. It just spends three or five minutes in each mission. For the next time, the students can play the progress continue due to all of typing record will be saved by computer server.

When the game starts, it assigns the question through the train by moving (Figure 2(b)), and the students have to type the answer in the specific area. Then, the results of mission will transfer to rewards through the “Star”, and it will be wrote down on the bottom of mission diagram (Figure 2(a)).



(a) Missions (b) Game interface

Figure 2. The interface of Flow-based typing game

3.2 Balance to Challenge and Skill

In order to keep the students are able to play the typing game continue, the game environment which support flow and enable learning must closely match the skill level of each student and also provide the clear goals and immediate individual feedback of tasks [1]. In flow-based typing game, the challenge refers to the learning materials and the number of practicing Chinese words. About the learning material, level of game is designed according to formal education in elementary school of Taiwan. The students have to type all vocabularies in the game level and assure of all of the answer are correct. And then, to calculate typing speed is according to the right score and the duration of typing activity, on the other words, typing speed (wpm) is that a student can input the amount of words in one minute.

To maintain the student's Flow experience, the exercise must balance the challenge of the game and the student's skill to address and overcome it [2]. According to Csikszentmihalyi's [5] flow theory, the learners reach a state, called *Flow*, which let a person has higher learning performance and pay attention to achieve the goals of task [1, 3, 7]. The Flow-based typing game provides the challenge which follows the textbook in formal education becomes difficult with the grade of students. If the students do not achieve the lowest standard of mission, it does not unlock and means it reduces the opportunity of outside the flow zone for students.

4. Preliminary Evaluation

4.1 Research Design

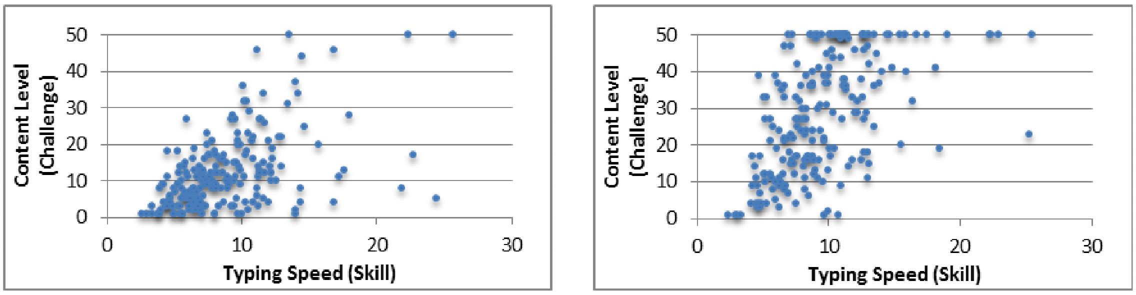
The participants are 205 elementary school students from 8 classes of grade 2 students who aged 8 to 9 years. In this study, all of students have owned a small tablet computer themselves and they are in a wireless environment classroom. The game was started to play at last semester. Four levels of typing game played, and total missions are 50 (Level 1 has 8 missions. Level 2, Level 3, and Level 4 have 14 missions for each.). Each mission had twenty to thirty Chinese words, including the content from grade 1 to grade 2. So far they had played for six months.

Table 1. Chinese Words of Level

	Time	Level	The Number of Mission	Range
Grade 1	1st Semester	1	8	1~8
	2nd Semester	2	14	9~22
Grade 2	1st Semester	3	14	23~36
	2nd Semester	4	14	37~50

4.2 Results

To calculate the total typing speed for all students during the duration of six months, the average is 9.57 wpm (Word per Minute). Figure 3 illustrates the student distribution between the typing speed and content level after one month a) other one is six month later b) At first one month, most of the students are at a low-challenge and a low-skill state. After six months, a group of students who are at low-challenge and low-skill move the upper challenge and raise the skill slowly. However, more and more students' level already pile up at the highest end. The phenomenon maybe impairs the ability of investigators to determine the central tendency of the data.



(a) One-month later (b) Six-months later

Figure 3. The student distribution between content level and typing speed

5. Discussion

5.1 Sustaining the Flow Experience of Children

The objective of this paper is to design a typing game to sustain flow experience of children, and so many students have done it (complete 50 levels). But as a matter of fact is students make less progress of typing speed during the game duration. Furthermore, while they have accomplished all of missions, they will no longer play the game when they feel it is not necessary.

5.2 Facilitating the Accessibility of Numerous Chinese Words

Moreover, there are many students who although get the high-challenge (Level 4), typing skill is lower than average speed (9.57 wpm) after the game starting for six months. On the other hands, it is difficulty to draw up the standard of typing skill with elementary school student because of the past typing exercise have no records to show the typing data with such age of students. When students do not familiar with the Chinese words and input method, the present model provides a simple learning environment to help them practicing the typing skills and know the vocabularies. After a period of time, while the students can master the typing skill and access numerous Chinese words through the typing game, then it will ask the students to strengthen their skilled and raise the typing speed. In order to pass the mission to get more rewards or compete with peers in the game, the students just type as fast as they can.

6. Next Step

In this study, we hope to the students who get the flow experience can help them to sustain the typing exercise. It is important to balance the challenge and skill of *flow theory*. The original design of the typing game in order to let the students easy to play, but it seems to cause the challenge of game is too easy to improve the typing speed difficultly and slowly. In other words, the challenge design does not take the typing speed into consideration that the relationship between the challenge and skill is too low. Next, we will add the time pressure for developing efficiency in game and the game will become more excited with students. Finally, Flow-based typing game will start a serial of researches and designs which can help the students construct their typing skills to sustain the practice until the typing no longer becomes the block of learning in the future.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (NSC 101-2811-S-008-009, NSC 101-2631-S-008-003, & NSC 99-2511-S-008-002-MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

[1] Bizzocchi, J., & Paras, B. (2005). Game, motivation, and effective learning: An integrated model for educational game design.

[2] Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, 50(4), 31 34.

[3] Cheng, H. N. H., Deng, Y. C., Chang, S. B., & Chan, T. W. (2007). EduBingo: design of multi-level challenges of a digital classroom game. *Digital Game and Intelligent Toy Enhanced Learning, 2007. DIGITEL' 07. The First IEEE International Workshop on* (pp. 11–18).

[4] Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*. San Francisco: Jossey-Bass.

[5] Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience. *New York: Harper & Row*.

[6] Liu, E. C. C., Liao, C. C. Y, & Chan, T. W. (2012). Developing a Flow-based Typing Game to Sustain Typing Behavior of Children. The Global Chinese Conference on Computer in Education (GCCCE2012).

[7] Wu, W., Cheng, H. N. H., Chiang, M. C., Deng, Y. C., Chou, C. Y., Tsai, C. C., & Chan, T. W. (2007). AnswerMatching: A competitive learning game with uneven chance tactic. Digital Game and *Intelligent Toy Enhanced Learning, 2007. DIGITEL' 07. The First IEEE International Workshop on* (pp. 89–98).

Design of Penalties against Useless Plays in Educational Games

Takanobu UMETSU[a*], Tomoka BABA[a], Tsukasa HIRASHIMA[b], Akira TAKEUCH[a]
[a]*Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan*
[b]*Department of Information Engineering, Hiroshima University, Japan*
*umetsu@ai.kyutech.ac.jp

Abstract: In this paper, we present a design of penalties against the action in contravention of the rule to discourage the tendency of unthinking and useless plays, observed while playing the automatically-generated educational games by our design method. Monitoring player's activity and penalizing illegal moves are potential ways of preventing unconsidered activities. Because of the illegal moves and adequate penalties are strongly linked in each educational game, it is necessary to design the penalties individually for each educational game, therefore, we investigated the automatic design of penalties for educational games, generated by the EPIC method. We also introduce a function to automatically implement the penalties, and report the experimental evaluations of educational games with penalties.

Keywords: Educational game, Authoring system, Design method, Automatic improvement

Introduction

We proposed a design for an educational game by exemplifying existing card games [1] and developed an automatic generation function for computer-based educational games [2]. In this study, we present a design for penalties to discourage the tendency of unthinking and useless plays, observed while playing the educational games designed by our method. We also introduce a function to automatically implement the penalty-based design, and report experimental evaluations of educational games with the penalty.

1. EPIC method

We present the embedding problem-solving exercises into a card game (EPIC) method, which transforms an existing card game into an educational one by substituting original cards for new cards that present problems to the players. To move forward in the game, players have to solve the problems provided on the substituted cards, instead of using the property of the original cards.

We explain a simplified version of the Memory game (also known as Concentration), transformed using the EPIC method into an educational game that uses basic arithmetic formulas. The Memory game is a card game in which cards have the same number when flipped are collected by the player, and those that do not are turned face down and left as is. The Memory game is transformed into an educational game in which cards with arithmetic formulas have the same answer when flipped are collected by the player.

If a player fails to solve the problems in playing an educational game made by EPIC method, the player will probably move cards in contravention of the rules of the game. However, because it is necessary to prevent the illegal move, the computer-based educational game identifies the player's mistake and allows the player a second attempt to choose from a multiple choice-selection.

2. Analysis of Plays of Generated Games

We conducted an experiment to demonstrate how many times a player was able to solve a problem when playing the educational game designed by the EPIC method. Ten subjects played each of the three computer-based educational games for ten minutes. Three educational games were designed from the Old Maid game, Memory game, and the War game. All three games were changed to incorporate exercises that were based on arithmetic formulas using three-figure numbers.

Table 1 shows the number of times the problem was solved, the number of illegal moves, and the number of legal moves during the ten minutes of the game. The ten subjects are indicated by letters from A to J. An illegal move meant that the subject moved cards violating the rules. For example, in the educational Memory game if cards with a different answer were picked up when flipped, the move was considered as illegal.

In this experiment, subjects C and I made more illegal moves than the other subjects. In addition, although they made many more moves, they solved fewer problems. When two subjects were asked about this, they said that they moved cards at random without actually solving the problems until the move was legal by chance. The subjects also said that they did not care that illegal moves were identified by the educational game system.

Table 1. Activities of plays during the one ten-minutes game

	Subjects	A	B	C	D	E	F	G	H	I	J
Old Maid	Solving	10	25	1	38	18	12	28	31	1	16
	Illegal moves	0	1	146	0	0	1	0	0	137	0
	Legal moves	22	27	8	26	21	28	26	21	5	19
Memory	Solving	18	21	3	24	11	14	22	31	0	21
	Illegal moves	0	0	28	0	1	0	0	0	29	0
	Legal moves	27	24	78	45	32	21	48	60	87	28
War	Solving	18	28	0	39	16	16	31	33	0	19
	Illegal moves	0	0	61	0	0	0	0	0	58	0
	Legal moves	18	28	112	39	21	16	31	35	101	19

3. Penalties Corresponding to Game Objective

We decided to penalize illegal moves to discourage unconsidered activity as well as the failure to solve problems. However, the level of penalties imposed differs depending on each game. Although an automatic generation function has been developed for computer-based educational games, it is necessary to design the penalties individually for each educational game. Thus, in this paper, we investigated the automatic design of penalties for educational games that are created by the EPIC method.

Based on an examination of 186 existing card games, we designed a system of penalties. The penalties are given by imposing a disadvantage. This disadvantage deviated from the game's objective. Game's objectives are categorized into four types. Thus, we designed penalties that were in line with the four game's objective, as follows:

- Objective: to score as high as possible/to not score as low as possible
  Penalty: subtract one point from the player's score/add one point to player's score

- Objective: to increase the number of cards /to decrease the number of cards
  Penalty: skip next step that increases the cards/skip next step that decreases the cards
- Objective: to make a high-ranking hand/to make a low-ranking hand
  Penalty: player's hand will slide down one rank in the next game
- Objective: to stay in the game as long as possible/to leave the game as early as possible
  Penalty: give a penalty in a sub-goal of the objective

4. Evaluation

We developed a function to automatically implement penalties in educational games as part of the authoring system [2] on the basis of EPIC method. The penalty function identifies the objective of a game from a combination of words or sentences that are within its rules and automatically implements the corresponding penalty rules.

We conducted an experimental evaluation to confirm whether the educational game improved by the function of implementing penalties to prevent unconsidered activity. For the evaluation, the function improvised the three educational games as mentioned in Chapter 2. We instructed the ten subjects, they were the same members who participated in the evaluation conducted in Chapter 2, to play each game for ten minutes. In addition, we explained to them the penalties were imposed when they made an illegal move.

Table 2 shows the results of subjects C and I. We excluded the results of the other subjects because the results were the same as those described in Chapter 2. The results suggest that the two subjects solved problems when playing the game with penalties. Their activities were almost the same as those of the other subjects playing the normal game.

Table 2. Number of illegal moves in the improved ten-minute game

		Old Maid		Memory		War	
		Normal	Penalty	Normal	Penalty	Normal	Penalty
C	Solving	1	13	3	18	0	26
	Illegal	146	0	28	1	61	0
	Legal	8	21	78	31	112	26
I	Solving	1	14	0	16	0	19
	Illegal	137	0	29	0	58	0
	Legal	5	21	87	24	101	21

5. Conclusions

In this paper, we present a design of penalties against the action in contravention of the rule to discourage the unconsidered plays, observed while playing the automatically-generated educational games by the EPIC method. We also introduce a function to automatically implement the penalty-based design. In experimental evaluations of educational games with the penalty, we confirmed that penalties prevent unconsidered activity.

References

[1] Umetsu, T., Hirashima, T., Takeuchi, A. (2006). Property Exchange Method for Automatic Generation of Computer-Based Learning Games. *Proceedings of ICCE2006* (pp. 283-490).

[2] Umetsu, T., Azuma, T., Hirashima, T., Takeuchi, A. (2011). Ontological Approach to Support Authoring for Game-Based Learning Environments. *Proceedings of ICCE2011* (pp. 26-33).

Full-body Interactive “Board”game for Learning Vegetation Succession Based on Identification of People and 3D Position Measurement

Masafumi Goseki†, Takayuki Adachi†, Hiroshi Takemura†, Hiroshi Mizoguchi†, Fusako Kusunoki††, Masanori Sugimoto†††, Etsuji Yamaguchi††††, Shigenori Inagaki†††† and Yoshiaki Takeda††††

†Department of Mechanical Engineering, Tokyo University of Science
2641 Yamazaki, Noda-shi, Chiba, 278-8510 Japan
††Department of Information Design, Tama Art University
2-1723, Yarimizu, Hatioji, Tokyo, Japan
†††Interaction Technology Laboratory, Deppartment of Frontier Informatics, University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
††††Graduate School of Human Development and Environment, Kobe University
3-11, Tsurukabuto, Nada, Kobe, Hyogo, Japan
Email: j7511629@ed.tus.ac.jp

Abstract: We are developing a digital game, "Human SUGOROKU," in which enables game players as pieces to learn vegetation succession. In order to realize this game, it is necessary to robustly measure 3D position of people and identify them. In this paper, we proposed a method for 3D position measurement and identification of people by integrating ultrasonic sensors and Kinect sensors. The evaluation results show that proposed method is more robust than methods using ultrasonic sensor and Kinect sensor separately.

Keywords: Practical study, vegetation succession, ultrasonic sensor, Kinect sensor

Introduction

In recent years, global environmental problems are becoming serious. Under these circumstances, the environmental education with experience has been more required. Digital game,"Digital SUGOROKU of vegetation succession", that targeted the conservation woodlands as environmental issues have been developed.

Figure 1 shows the overview of this game. The game visualizes changes in forest ecosystems in large temporal scales by simulation. Practical study of learning environment has been advanced by using this game [1] [2]. The game is backgammon on the computer screen. An issue of Connection with virtual world and the real world had been left for the learner. Therefor we are developing a new system of the game, "Human SUGOROKU”, which enables people to learn vegetation succession more realistic.

In that system, people walk on board which was placed on the floor. That person, made applicable to the plant, can experience the change of pseudo physically planting ecosystem. Figure 2 shows the overview of the system “Human SUGOROKU” . The overall size of board is approximately 10m square and the size of a frame is assumed to be 1m square. The number of participants is six. Up to three people may enter in one square.
In order to realize this system, it is necessary the two technologies.

First, even if there are some people within one frame which is one meter square, they can be measured position and identified. Second, it is possible to measure positions of people in large space of 10m square. A method to measure the position has been proposed by using ultrasonic transmitter tags with unique identifiers [3] [4]. But ultrasonic sensors may become unstable due to the measurement of the directivity of the ultrasonic wave. The range which ultrasonic sensor can stably measure transmitter's position is limited. This is a problem because there is a need to measure position stable in a wide range. If we try to realize this system using only ultrasonic sensor, huge amount receivers are required. On the other hand, Microsoft's Kinect sensor can measure the position of the people by the visual information. Therefore the sensor can't measure if occlusion occurs due to the overlapping people. This is a problem because there are 6 people in the game. So we intended to develop a system to measure 3D position of people and identify them if they are in a narrow range by integrating Kinect sensor and ultrasonic sensor to compensate for the weaknesses.

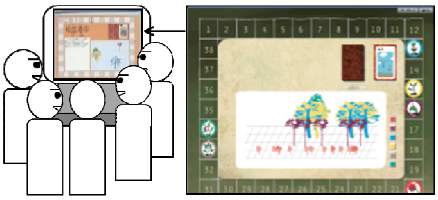


Fig.1 Digital SUGOROKU

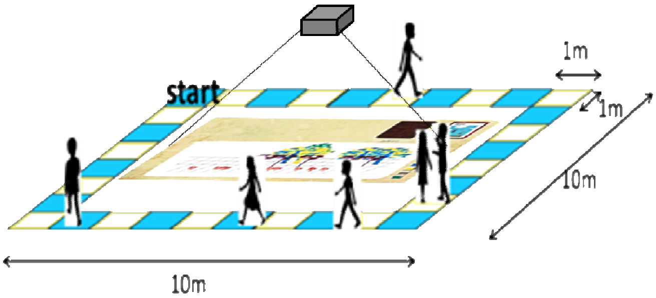


Fig.2 Human SUGOROKU

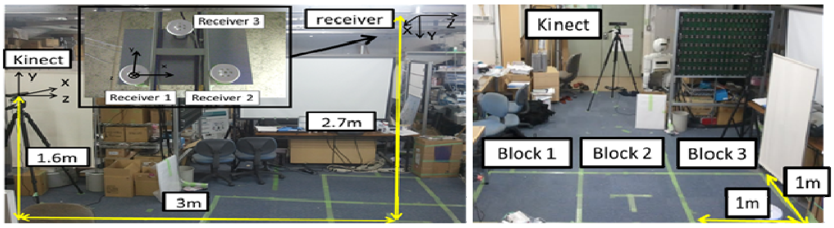
1. Experiment

By integrating the sensors, the weaknesses of each sensor are complemented. Then we investigated the stability of the measurement. In preliminary experiments, the position of a person was measured considering the weakness of each sensor. We compared the stability of the measurement by using each sensor separately and the measurement by integrating sensors. In main experiment, we investigated how much the system can measure position of two people.

Figure 3 shows the experimental environment. Receivers of the ultrasonic sensor were placed on the ceiling. The receivers were arranged such that an equilateral triangle of sides 300mm. Three frames placed on the floor under the receivers. The frames were allocated a number as figure 3 (b). The frames were 1 meter square. There is a center of the frame 2 on the z axis of the ultrasonic sensor. Border line is parallel to axis x and axis y. Coordinate system of the Kinect sensor was rotated 90 degrees counter-clockwise X axis of the ultrasonic sensor and moved in parallel.

Two people attached a transmitter on his head moved in the frames of three. They were sometimes hidden from the Kinect sensor due to overlapping. Figure 4 shows the experiment method. Ultrasonic sensors measured the position of the transmitter attached to the head. Kinect sensor measured the position of the head. Sampling rate is 25Hz and sampling number is 1500. Figure5, 6, and 7 shows the relationship between the sampling number and position of two people in x direction.

Because receivers cannot receive ultrasonic wave, there is the range which ultrasonic sensor does not measure position well. In addition, there is the range which Kinect sensor cannot measure position of the people due to overlapping. It is possible to robustly measure position of two people by integrating the sensors



(a) (b)

Fig.3 Experiment Environment

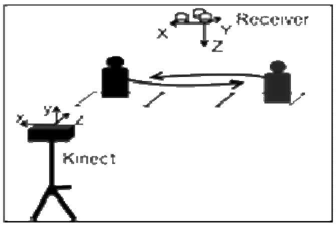


Fig.4 Experiment Method

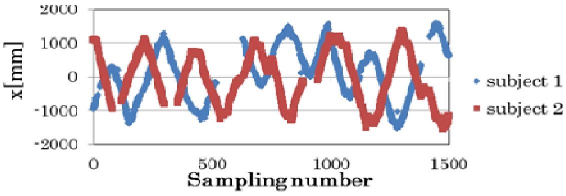


Fig.5 Measured by Ultrasonic Sensor

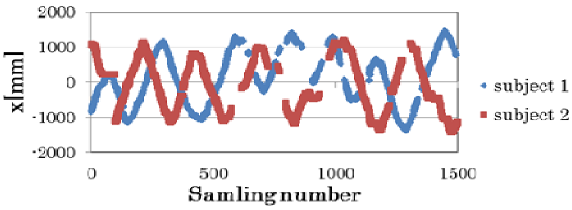


Fig.6 Measured by Kinect Sensor

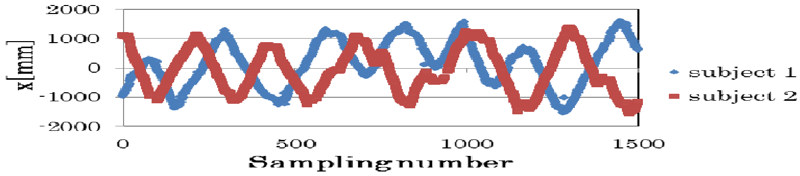


Fig.7 Measured by Integrating Ultrasonic Sensor and Kinect Sensor

2. Conclusion

This paper describes a method for 3D position measurement and identification of people by integrating ultrasonic sensors and Kinect sensors for developing "Human SUGOROKU". Ultrasonic sensor can measure 3D position and identify tagged people. Kinect sensor can measure 3D positions of people and tracking them with OpenNI. By combining the Kinect sensor and ultrasonic sensors, we made it possible to measure positions of people and identify them. The experiment results show that the proposed method is more robust than methods using only ultrasonic sensor or only Kinect sensor.

The proposed method robustly measure position of people in three squares. In the future, in order to realize "Human SUGOROKU", we will extend this system. We will develop a system which can measure the positions of people and identify them in a wide range of approximately 10 meters squares.

References

[1] T. Matsumura，Y. Takeda，"Relationship between species richness and spatial and temporal distance from seed source in semi-natural glassland, " Applied Vegetation Science, 13:336-345,2010.

[2] A. Deguchi, S. Inagaki, F. Kusunoki, E. Yamaguchi, Y. Takeda, M. Sugimoto, "Vegetation interaction game: Digital SUGOROKU of vegetation succession for children," Entertainment Computing-ICEC, Lecture Notes in Computer Science LNCS6243, 493-495, 2010.

[3] Y. Nishida, H. Aizawa, T. Hori, N.H. Hoffman, T. Kanade, M. Kakikura, "3D Ultrasonic Tagging System for Observing Human Activity, " IEEE International Conference on Intelligent Robots and Systems (IROS2003), pp. 785-791, 2003 .

[4] A. Nishitani, Y. Nishida, T. Hori, T. Kanade, H. Mizoguchi, "Global Calibration Based on Local Calibration for an Ultrasonic Location Sensor," in Proceedings of 1st International Conference on Sensing Technology (ICST2005), pp. 11-16, 2005

Training My-Dragon: Using Educational Agents to Facilitate Student Learning

Zhi-Hong Chen, Chia-Chi Yeh
Department of Information Communication, Yuan Ze University, Taiwan
zhchen@saturn.yzu.edu.tw

Abstract: In this paper we develop a My-Dragon system, in which students not only learn Chinese idioms, but also train their pet dragons through learning by demonstration. More specifically, students need to demonstrate correct procedure and answers for their dragons so that their knowledge can be improved. By doing so, students not only can be engaged in a game-based learning context, but they are offered more opportunities to master their subject domains. This is a work-in-progress study and the current status and future work are described in this paper, too.

Keywords: Game-based learning, educational agent, learning by demonstration

1. Introduction

Interface design, in general, has been a significant issue to facilitate human computer interaction, and further enhance students' learning achievement. For learning, *interactive agents* have been incorporated into the interface design [1]. Interactive agents are essentially embodied as virtual characters, which on the one hand attract students' attention, and on the other hand facilitate the system's communication with students [2]. For instance, *animated pedagogical agents* are used to enhance students' motivation and facilitate communication bandwidth through the agents' body language, such as nodding, gestures, and eye contact [3].

Although these studies offer valuable experience in designing intelligent agents, there are great challenges in terms of development cost and participatory motivation. One possible reason might lie in the fact that educational agents are often designed from a "smart" (i.e., intelligent and autonomous) perspective, and so are used in learning technologies as intelligent tutors offering vivid and adaptive interactions with students. In addition, another possible drawback to the "smart" agents is that students might not be so interested in interacting with these educational systems, especially when they find that these educational agents are not as intelligent as expected [4].

On the other hand, purposely "non-smart" agents have also been designed as a means of evoking active and responsible attitudes in students, such as peer tutees [5, 6] and teachable agents [7, 8]. Studies demonstrate that such agents have positive effects on student learning, meaning that it is worthwhile to investigate how they meet the two challenges. In such vein, this study aims to develop an educational agent, named the My-Dragon, so that the effects of such a "non-smart" agent on student learning can be investigated. In particular, the My-Dragon in this study is incorporated with the characteristic of pet-nurturing to develop a close relationship with students. In addition, the model of "learning by demonstration" is used to train the My-Dragon, during which students' mastery level of Chinese idiom could be improved because they are required to demonstrate Chinese idiom to their dragons.

2. My-Dragon system

1.1 Training mode

In this study, Chinese idioms are used as the subject domain because Chinese idioms form a significant part of Chinese literature [9]. A numerical format scoring system is used because such a numerical format is useful to classify students' learning status of Chinese idioms into three aspects: Remembering, understanding, and applying. More specifically, a "*learning by demonstration*" model is used in a game scenario, in which students improve their learning by training their dragons. To this end, two functions are offered: *Awareness* and *improving*. The former is designed to help students understand their learning status and their dragons' status, including the mastery level and progress for a specific Chinese idiom. By doing so, students can know which topics they have or have not mastered yet. The latter is to further encourage students to improve their learning by demonstrating correct learning process for their dragons so that students' "effort" value and their dragons' "knowledge" value can be enhanced. Figure 1 illustrates an example, where a student needs to pick out four words one by one in the correct sequence according to the given context, avoiding the selection of similar but erroneous words. During this process, the dragon will watch and learn, increasing their knowledge.

Figure 1. Training mode in the My-Dragon system

1.2 Nurturing mode

Regarding the nurturing mode, previous studies have indicated that developing relationships with pets can enhance students' participatory motivation, and learning achievement [10]. To this end, two functions are offered: *Feeding* and *nurturing*. As shown in Figure 2, the former is to enhance students' sense of being a "master", who is responsible for their dragons. Some types of food are used to feed them. In addition, the "hungry" value of the dragons is also presented to indicate whether they are hungry. By doing so, students can obtain feedback and further adjust their behaviors. Likewise, the latter is to enhance the students' sense of being a "care-giver", who not only satisfies the dragons' needs, but also influences their "loyal" value, implying that students can different kinds of services to enhance this value, and further offer participatory motivation and opportunities.

Figure2. Nurturing mode in the My-Dragon system

3. Current status and future work

The My-Dragon system described above is now developing. In particular, to make students perceive that training their My-Dragon is an interesting and meaningful task in terms of affective and cognitive aspects, two mechanisms are planned to further incorporated into the system: emotional feedback and growing-up mechanisms. The former might enhance the quality of human-computer interaction; the latter might maintain students' long-term participatory motivation and the goal-pursuing. In addition, an experiment is also planning to conduct in an elementary school, of which purpose is to examine the effects of the My-Dragon system on student learning. The results are expected to offer valuable comments on the further development of educational agents within a game-based learning environment.

References

[1] Preece, J., Rogers, Y., & Sharp, H. (2002). Interaction design: Beyond human-computer interaction. Hoboken, NJ: John Wiley & Sons.
[2] Wolf, B. P. (2009). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Burlington, MA: Morgan Kaufman.
[3] 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, 47-78.
[4] Shneiderman, B. (1998). Designing the user interface: Strategies for effective human-computer interaction. Reading, MA: Addison-Wesley.
[5] Davis, J. M., Leelawong, K., Belynne, K., Bodenheimer, R., Biswas, G., Vye, N., & Bransford, J. (2003). Intelligent user interface design for teachable agent systems. International Conference on Intelligent User Interfaces, Miami, Florida, 22-33.
[6] Leelawong, K., Davis, J., Vye, N., Biswas, G., Schwartz, D., Belynne, T., Katzlberger, T., & Bransford, J. (2002). The effects of feedback in supporting learning by teaching in a teachable agent environment. In Bell, P., Stevens, R., & Satwicz, T. (Eds.), Keeping Learning Complex: The Proceedings of the Fifth International Conference of the Learning Sciences (ICLS), 245-252. Mahwah, NJ: Erlbaum.
[7] Biswas, G., Leelawong, K., Schwartz, D., & Vye, N. (2005). Learning by teaching: A new agent paradigm for educational software. Applied Artificial Intelligence, 19, 363-392.
[8] Chase, C., Chin, D. B., Oppezzo, M., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. Journal of Science Education and Technology, 18, 334-352.
[9] Luk, R. W. P., & Ng, A. B. Y. (1998). Computer-assisted learning of Chinese idioms. Journal of Computer Assisted Learning, 14, 2-18.
[10] Chen, Z. H., Liao, C. C. Y., Chien, T. C., & Chan, T. W. (2011). Animal companions: Fostering children's effort-making by nurturing virtual pets. British Journal of Educational Technology, 42(1), 166-180.

Development of Air-squat Supporting System using Microsoft Kinect

Youji Ochi
Faculty of Science and Engineering, Kinki University, Japan
ochi@kindai.ac.jp

Abstract: Sports training is intended to promote health and improve skills of the sport. We think that the Study of the education system application of motion capture equipment such as Kinect become popular in the future. In this paper, we propose a new approach for air-squat, which is a type of training, behavior training to provide support using squat Kinect there. Finally, this paper reports on the development and operation verification of training support systems.

Keywords: Air-squat Supporting, Kinect, Skill Learning

Introduction

Sport training is intended to promote health and improve skills of the sport. Most people don't know the correct way of training. Their own way of training cannot be gotten the effect as expected and leads to an injury. There are some researches of sports supporting systems using virtual environments [1]. It is very important for educational system research to recognize behavior of learners or teacher [2][3]. In addition, research has also been movement analysis using motion capture equipment. We think that the education system using motion capture equipment such as Kinect becomes popular in the future [4]. However, Kinect has some constraints on the specification. Therefore, we propose a new approach for air-squat, which provide the training environment using Kinect. In this paper, we report the development and evaluation for the verification as training support systems.

1. Requirement for air-squat supporting

In this study, we define the following requirements for the air-squat supporting system.

(1) Support to check the position of the hip and knee

It is difficult for the user to check the position of her/his hip and knee. Because when s/he checks them by her/himself, s/he often slouches forward. It means the wrong movement. Therefore the environment where the user can check the position of the knee and hip is necessary.

(2) Support of checking the motion speed

Squat requires hold for at least one second after bending down. Therefore the user must look at a watch continuously to check the time. It is difficult for the user to check it during the squat exercise. Therefore, it is necessary environment where the user can check the time of squatting.

2. Our approach

2.1 Introduction of Kinect and the constraints on the specification

We think that it is appropriate for the introduction of motion capture equipment to support body movements. In this study, we focused on the Kinect. Kinect can output the coordinates of 20 joints as user's skeleton information. It is better to show a picture of the user from the side angle. However, Kinect outputs the human body model from front view only. Therefore Kinect cannot output the skeleton of the user correctly from the side view. Moreover, Kinect cannot output the correct skeleton information without full-body photograph. Therefore, the shooting location is important.

2.2 Attitude estimation and visualization of the using depth information

We focus on the depth information of skeleton from Kinect. The depth information is a Z coordinate of skeleton. The infrared sensor of Kinect calculates it. By using this depth information, it is possible to determine the Z coordinates direction of the knee. Then, I propose the projection way based on the depth information. The way is that the vertical axis is Y and the horizontal axis is Z. I think that information supports to estimate the angle from the side and check her/his skeletal position of the toe, hip and knee.

3. Prototype system

We have developed a system air-squat training system using the Kinect for SDK beta2 for the implementation of this system. The following describes in detail. The user interface is following four areas (Figure 1). "Body position view area" shows Z-Y axis skeleton information, the number of squat, and squat keeping time. "Mode show area" displays the current operating mode by the arrow shapes and text for user. "Video view area" displays a video image of Kinect for the user in order to understand the shooting range. "Skelton view area" displays an image of the user's skeleton form.

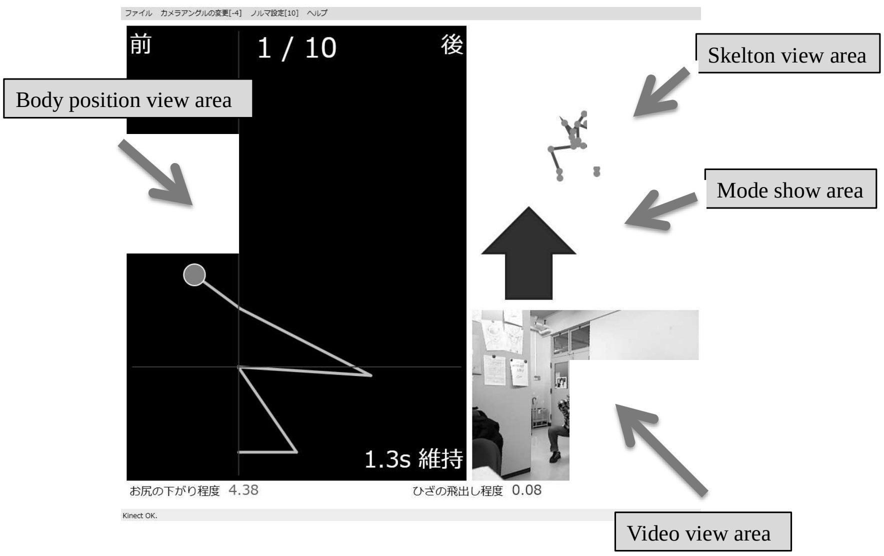


Figure 1 User Interface

4. Evaluation

We have prepared two trial environments (T1 and T2). T1 had the system navigation (Showing Main Window and Speech message) and T 2 had not it. We made 11 examinees (A-K) squat 10 times in the both environments. Then, we investigated the number of squat recognized by the system. We have described the definition of the air-squat to the examinees before this trial. Figure 2 shows the result. The “Failure” in the figure means the examinee did not bend down over one second in the keep mode. In T2, the system judged that most examinees did not bend down to the base position of their knees. Therefore, the system could not be recognized as their squat motion. The number of squat is less than 10 times about most examinees. In T1, the system was able to recognize all the squat of all users. The examinees were trying to bend down to the base position by watching the screen. The number of most users less than 10 times. I think the reason that the air-squat is difficult for the examinees.

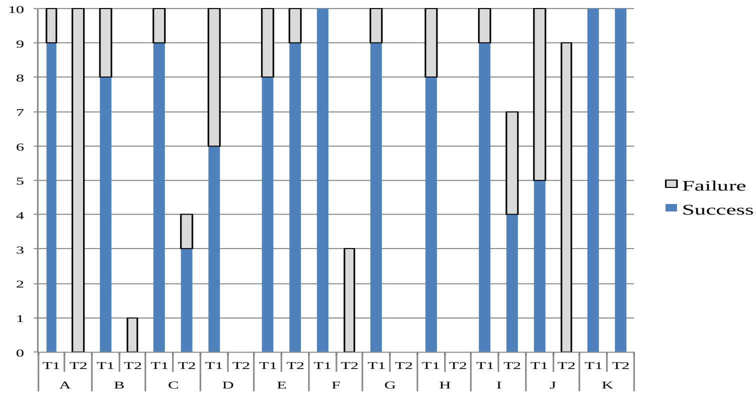


Figure 2 The Number of Squat

5. Conclusion

In this paper, we reported on the development and operation verification of training support systems. The users tend to drop the hip was seen more deeply by using this system. However, there were some different between the system’s posture and the user’s one in the examinees. My future work is the improving the system’s validity

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research(C) 23501122 from the Ministry of Education, Culture, Sports, Science & Technology in Japan

References

[1] Helen C. Miles, Serban R. Pop, Simon J. Watt, Gavin P. Lawrence, Nigel W. John (2012). A review of virtual environments for training in ball sports, Computers & Graphics 36, 714–726.

[2] Shimada Atsushi, Suganuma Akira, Taniguchi Rin-ichiro (2004). Automatic Camera Control System for a Distant Lecture Based on Estimation of Teacher’s Behavior, Proc. of the Seventh IASTED International Conference on Computers and Advanced Technology in Education, 106-111.

[3] Yong Rui, Anoop Gupta, Jonathan Grudin and Liwei He (2004) Automating lecture capture and broadcast: technology and videography, MULTIMEDIA SYSTEMS,Volume 10, Number 1, 3-15.

[4] Hui-mei Justina Hsu, The Potential of Kinect in Education (2011). International Journal of Information and Education Technology, Vol. 1, No. 5, 365-370.

Video Gaming Scale Effect on Spatial and Graphical Patterns Recognition on Eye Movement Behavior

Ben CHANG[a*], Meng-Lung LAI[b] & Yu-Cheng TSAI[a]
[a]*Department of E-Learning Design and Management, National Chiayi University, Taiwan*
[b]*Department of Early Childhood Education, National Chiayi University, Taiwan*
*bchang.tw@gmail.com

Abstract: The purpose of this study is using eye-tracking technology to explore the relationship between the video gaming scale, and spatial and graphical patterns recognition tasks, focusing on eye movement fixation number analysis. Nineteen sophomore students were recruited to participate in this experiment. The results showed highly negative correlation between the video gaming scale measure and the fixation numbers collected in the process of the advanced spatial recognition and graphical patterns recognition tasks.

Keywords: Fixation numbers, video gaming scale, eye-tracking

Introduction

Video games offer an attractive outlet in which students are willing to spend an ample amount of time engaging in virtual environments. This phenomenon has caught the attention of many researchers and has sparked a lot of game-related research. However, research exploring the video gaming scale effect on students' physical behavior is still rare, especially focusing on eye movement behavior. Owing to evidence pointing to younger generations spending an ample time on gaming, research into how gaming scale affects reading eye movement behavior will be worth further exploration. The goal of this study is to use a direct physical detection device, an eye-tracking machine, to detect the gaming scale effect on spatial and graphical patterns recognition tasks.

1. Testing Materials

Figure 1 shows the spatial and graphical patterns recognition testing materials which are comprised of four sub-tasks: (i) a simple spatial recognition (identified as S1); (ii) an advanced spatial recognition (identified as S2); (iii) finding the three non-correlating spots among the horizontal figures (identified as G1); (iv) finding the three non-correlating spots among the vertical figures (identified as G2). In S1, the simple spatial recognition testing, the cube should be rotated horizontally 90 degrees, and then be turned 90 degrees counterclockwise to match the answer, choice three. In S2, the advanced spatial recognition testing, the cube should be rotated counterclockwise horizontally 90 degrees, and then be rotated at z-axis 180 degrees to match the answer, choice four. By comparison, the second question (S2) requires more spatial manipulation to complete which deems it the more difficult task. In the G1 and G2 tasks, the experimenters were asked to identify differences between two nearly identical pictures. There were three differences between the two provided pictures.

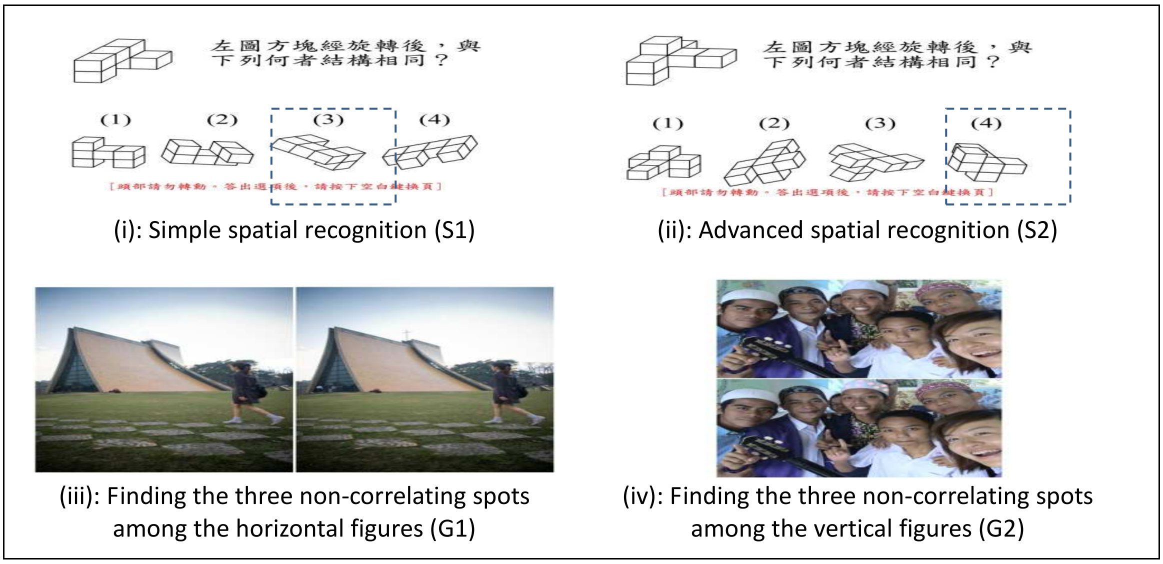


Figure 1. Spatial and graphical patterns recognition testing materials.

1.1 Gaming Scale Measurement

A video gaming scale questionnaire revised and translated from an online gaming scale assessment (Kim, Namkoong, Ku, & Kim, 2008) which has a reliability test (α=.90) was applied. The questionnaire is a five-point perception Likert item questionnaire (1 = almost never, 5 = quite often) with twenty questions about video gaming scale behavior. The questionnaire items are about the users' gaming habit self-reflection. Before the testing, all the students were asked to fill out the questionnaire. According to the questionnaire, all the students can be ranked with a number from 20 to 100, which reflects the scale of the gaming. The higher score represents the higher video gaming scale, and vice versa.

2. Experimental Procedure

Nineteen sophomores participated in this experiment. Among them, five were male and fourteen were female. Before the formal experiment, the students were assigned a warm-up exercise to familiarize students with the experiment testing materials. Meanwhile, they were asked the fill out the video gaming scale questionnaire. During the formal experimental phase, the environment was designed as an individual reading space where the testing materials were displayed on the eye-tracking machine. The students' eye movements were logged by the Gazetracker for further analysis.

2.1 Pearson correlation coefficient analysis on overall fixations and video gaming scale

When given a same task, a user who requires more fixations to complete the task correctly has lower encoding performance in comparison to someone who requires fewer fixations. Table 1 lists the statistics of the overall fixation numbers of the involved students on the eleven different types of reading materials. In order to understand the relation between gaming scale and the overall fixation numbers, Person correlation coefficient analysis approach was applied to this study. Only the correct answers which are meaningful when analyzed are counted.

Table 1. Statistics of the fixation and regression numbers.

ID	Gaming scale	Spatial recognition fixation #		Graphical patterns recognition #	
		S1	S2	G1	G2
1	57	48	69	**43**	121
2	64	70	**36**	114	130
3	54	**25**	82	**47**	**42**
4	27	25	34	**75**	**165**
5	25	19	27	**84**	**177**
6	27	**40**	32	**75**	81
7	25	**36**	**58**	24	14
8	60	50	40	**50**	**79**
9	51	**29**	**26**	**41**	**34**
10	35	**21**	21	**46**	37
11	29	**30**	19	24	**90**
12	43	**40**	26	**46**	**78**
13	31	**22**	**68**	**77**	209
14	48	**15**	**41**	**20**	233
15	43	55	**54**	**85**	264
16	27	**47**	42	27	**146**
17	45	43	**45**	185	128
18	33	18	**53**	**47**	**77**
19	46	53	83	**32**	**30**

* The bold and underlined numbers represent the tasks completed without error by each student.

Table 2 lists the Pearson correlation coefficient analysis with video gaming scale and correct answers of the overall reading fixation numbers. The Pearson correlation coefficient analysis results indicated that significantly negative correlation coefficients for spatial recognition (S2) and graphical patterns (G1, G2) were found. When given a task, a user who requires fewer fixations to complete the task correctly is deemed to have higher encoding performance. According to this phenomenon, the statistics result showed that game playing can help students increasing their graphical recognition ability in both in spatial recognition and graphical patterns recognition.

Table 2. Gaming scale and overall fixations of Pearson correlation coefficient analysis.

	Spatial Recognition		Graphical Patterns Recognition	
	S1	S2	G1	G2
Gaming Scale	-0.48	-.803*	-.629*	-.776**

3. Conclusions

This study was devoted to exploring the gaming scale effect on students' visual search behavior. The gaming scale was assessed by a video gaming scale questionnaire, and the overall fixation number index was applied to analyze the eye movement. The study result indicated that the subjects who spent a large amount of time playing game had the less overall fixation numbers for the spatial recognition and graphical matching patterns tests. The results supported by most findings that the video gaming can improve students' spatial and graphic pattern recognition ability.

References

Kim, E.J., Namkoong, K., Ku, T., & Kim, S.J. (2008). The relationship between online game addiction and aggression, self-control and narcissistic personality traits. *European Psychiatry*, *23*, 212-218.

Rayner, K. (1998). Eye movement in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422.

Exploring Teachers as game-based learning courseware and scenario co-designers for enhancing student English learning

Shelley S. C. Young & F.W. Lin
Institute of Information Systems and Applications, National Tsing Hua University, Taiwan
scy@mx.nthu.edu.tw

Abstract: The purpose of this study was to explore and describe how to transform the fixed text-based English learning materials into a flexible inviting and interesting online game-based courseware through the collaboration of a research team composed of three parties: school English teachers, experts from university and a software company that might not be possible otherwise. This paper describes the rationale of applying Universal Design for Learning (UDL) into the curriculum design, system frame work, and game scenario. In addition, a formative evaluation was conducted to the target students. The results of the evaluation indicate that most of the learners were positive to the game-based courseware design in terms of learning content and interface design. Further research is presented at the end of this study.

Keywords: Universal Design for Learning (UDL), Game-based learning, Competitive learning, English vocabulary learning

Introduction

Nowadays online games have become an indispensable role in many children's lives and leisure activities. Games provide engaging activities and bring fun and joyfulness, and moreover, games might provide a useful and attractive new method of learning. Ministry of Education (MOE) in Taiwan has been promoting integration of information technology (IT) in teaching and learning for years. However, most teachers remain to feel it difficult and challenging because most of the online game systems do not offer learning activities that are designed with integrated learning content, nor do they offer teachers the opportunity to incorporate appropriate learning materials into the systems (Yang, Chen, & Jeng, 2010). This paper reports on a systematic longitudinal study through intensive interactions and observations. The authors investigated and described how the three parties, school teachers, Learning Technology Lab in university and a software company, applied the curriculum design model Universal Design for Learning (UDL) (Rose & Meyer, 2002), to the common courseware stipulated by the Ministry of Education (MOE) and transformed the paper-based fixed English curriculum into a flexible game-based online learning scenario enriched with multiple learning modalities in the hope that the jointly designed courseware could attract students and engage them through multiple learning modalities to support class-based and after school learning or self-paced learning during their free time. The following research questions guide this study:

(1) How could the research team transform the English curriculum into an inviting and interesting online courseware that would bring more learning opportunities to adapt to diverse student needs in the afterschool settings?
(2) How should the research team design the text-based curriculum into a flexible online

learning materials and environment? What curriculum design model could the team use and what metaphor and design principles could be concluded?

(3) What would the curriculum turn out like in terms of an interesting learning environment? How would the target students perceive the online courseware and learning environment in terms of learning content and interface design?

1. Methodology

The methodology adopted in this study could be divided in two parts: curriculum development and online courseware evaluation. For the curriculum development, it is primarily a qualitative approach. This study on curriculum development is interpretive and descriptive in nature and uses the case-study method. Case studies are used to examine a specific unit such as an event, a program, an organization, and a time period in depth and detail, in context, and holistically. Triangulation was used to improve the probability that findings and interpretations would be reliable.

2. Results of curriculum development

2.1 Design of game-based online English learning environment

The research team followed the three principles of the Universal Design for Learning (UDL) (Howard, 2004; Rose & Meyer, 2002) including: 1. To provide multiple means of representation; 2. To provide multiple means of action and expression; and 3. To provide multiple means of engagement. Each party on this research team could have its role in engagement in design of the curriculum. Through the brainstorming design based on the guidelines of UDL, the research team transformed a courseware into the online English game-based learning materials. The design principles and scenario are concluded as follow:

(1) To ensure the learners' commitment to the game, role-playing and the plot of the play are important components of the game design.

(2) To create a more enjoyable and comfortable environment, the plots of the play, the colors and the scenes of the game are designed to fit the learners' age level.

(3) To provide a "multi-battle" model as a tactic to promote learners' motivation.

(4) To provide different levels and barrier-battles for learners to compete with each other to increase their English vocabulary power through the competitive process.

(5) To be adaptive to learners' different abilities so that earners can set their own learning goals and tactics as their abilities grow.

(6) To be a useful vocabulary-building strategy-tactics multimedia model, combining text, audio and video components.

2.3 Game design as to recruit learner interest and sustain efforts

The metaphor of a "Magic island" game model was designed into this learning content and scenarios. The online learning environment was developed and named "English Magic Rainbow Island" that offers 12 levels of challenging courseware for self-study and competitive activities. Two game modes are designed: a barrier mode, where learners can play different roles to set out on their English adventure journey; and a multi-battle mode, where learners compete with either themselves as genies or others in the "Magic Island". Through the learning by playing process, they could learn more vocabulary that empowers them to pass 12 different-level challenges represented by 12 different genies in the barrier mode. These magic genies also demonstrate 12 different vocabulary learning levels. The multi-battle mode is called “Arena”; arena is a place for learners to learn the vocabulary in the game. The vocabulary competition is controlled by the computer which generates questions based on the learner’s vocabulary level; the learner should answer the questions correctly and faster than their competitors in order to win the game.

2.4 Learning materials and pedagogical design

The goal of the game-based English courseware was to provide instruction and self drill-and-practice with vocabulary and sentences. English teachers in the elementary school differentiated the vocabulary and sentences into 12 levels based on the frequency of their appearance in the textbooks, and the students’ learning situations in the real world. The learning materials in the game were divided into two major parts, a “word-building zone”, and a “sentence-practicing zone”. The two major zones provide multimedia learning opportunities for learners, including: words, pronunciation, graphics, and example sentences. Learners will be able to understand the meanings of the words and understand more about their usages in authentic living contexts. To evaluate the learners’ performance in listening, pronunciation, reading and spelling, the related exercises and questions have been designed and integrated into the learning contents.

3. Conclusion and implications

The results indicate that this “English Magic Rainbow Island” jointly designed and implemented under the research team composed of three parties overall has received users’ positive feedback. A software development company could be as a partner to assume the system development work. The role of the education and technology expert in the university could be an important interface between the school teachers and the software company to provide advices based on relevant learning theories and research experiences. The implementation of the highly-demanding online courseware would one the one hand offer more flexible learning opportunities to satisfy diverse student learning needs in the after school settings and, one the other hand, would help ease teacher’s teaching load. The teachers involved in this joint research project expressed their thoughts that their technological knowledge and pedagogical content knowledge Technological Pedagogical and Content Knowledge have been trained and enhanced through the long-terms engagement and they could appropriate apply what they incorporated into the game-based learning courseware to associate with their classroom-based teaching as to maximize teaching effectiveness and efficiency. The empirical data on how the “English Magic Rainbow Island” is actually integrated into English learning contexts and in the after school learning settings and how students learn and benefit in the game-based environment will be collected and analyzed and reported in terms of learning outcomes, students’ motivation, and online learning strategies in a separated paper.

References

[1] Howard, K. L. (2004). Learning and Leading with Technology, 2004 - 4.17.143.133. accessed 2010.12.20. http://4.17.143.133/udl/downloads/LLT.pdf .

[2] Rose, D. H., & Meyer, A. (2002). Teaching Every Student in the Digital Age: Universal Design for Learning Alexandria, VA: ASCD.

[3] Yang, J. C., Chen, C. H., & Jeng, M. C. (2010). Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning, Computers & Education. 55, 1346-1356.

Learners' Perceptions of a Commercial Adventure Video Game for Learning English as a Second/Foreign Language

Christine Ting-Yu YANG & Howard Hao-Jan CHEN
English Department, National Taiwan Normal University, Taiwan
*christiney37@gmail.com

Abstract: This paper presents a study of a group of 60 Taiwanese EFL learners' perceptions of a commercial adventure video game for second/foreign language learning. The study results reveal that learners held a positive attitude toward the use of the adventure game for English learning and considered the game particularly beneficial to their listening ability, reading ability, and vocabulary knowledge. Learners also reported their enjoyment of using the game for language learning and stated enhanced motivation in English learning. Nevertheless, learners reported several language barriers encountered while gaming (e.g., fast dialogue, fleeting subtitles) and identified some weaknesses regarding the game designs. The research results of this paper identified both strengths and weaknesses of adventure games for EFL/ESL based on learners' perceptions, and should encourage more studies on the investigation of using adventure games in language learning.

Keywords: Adventure games, video games, learner perception, language learning, English as a second/foreign language

1. Introduction

1.1 Commercial Games in Education

The widespread popularity of digital video games over the last decade has aroused many researchers' interests to examine their educational value. These interests urged many researchers to design and investigate games for educational purposes (a.k.a. games-based learning, GBL), but more are on exploring the usefulness of commercial-off-the-shelf (COTS) games in learning. In their review, Connolly, Boyle, MacArthur, Hainey, and Boyle [3] collected 129 papers on computer games and found that more than half of the papers were research on games for entertainment.

The reasons for using commercial games more than self-designed educational games are not hard to understand. Since designing and integrating educational games requires both investments in training and funding [13], adopting existing commercial games will be a more available choice for most educators. Advocates of commercial video game-based learning also contribute to the acceptance of using COTS games in learning by proposing arguments highlighting these games' educational value [7][8][9][15][16]. For example, James Paul Gee regards video games as a form of "experiential learning" [8] and values highly on its educational potential. Foreman [6] and Prensky [16] also highly commend video games as motivating, learner-centered, and sensory-rich learning tools.

1.2 Commercial Games and Second/Foreign Language Learning

Researchers' interests in video game-based learning have contributed to the widespread occurrences of studies examining its effects on various subjects. One of the widely investigated areas is the learning of second/foreign languages. Studies have been conducted to investigate how commercial video games such as simulation games [12][17], massively multi-player online role-playing games [2][14][18][19][20][22], sports games [4], and music games [5] could facilitate second/foreign language acquisition, and many of the studies had reported encouraging results in language gains and positive learning attitudes.

In addition to the aforementioned game types, there are other game genres that might benefit second/foreign language learners as well. Characterized as highly narrative, story-based games that stimulate discovery learning, adventure games are believed to be particularly language-learning-beneficial by many researchers and DGBL-related professionals. Kirriemur and McFarlane [10] commented that adventure games allow learners to experience contextual learning by having them role-play while playing. Baltra [1] argued that video/computer adventure games are useful tools for developing communicative fluency. Tomasz Szynalski [21] also believed that adventure games can help increase learners' motivation and understanding of spoken English and thus proposed using adventure games in language learning.

Considering its narrative, motivating nature and the advocators' supporting arguments, it seems worthwhile to investigate how adventure games can benefit language learners' learning. However, few studies have targeted the use of adventure games on language learning, and even fewer have investigated how learners feel about this kind of games for English as a second/foreign language (ESL/EFL) learning. Thus, the current study aimed to investigate ESL/EFL learners' perceptions of playing a commercial adventure game for English learning and to answer the following research questions:

1. What language skills and knowledge can be improved via commercial adventure games based on ESL/EFL learners' perceptions?
2. What are the strengths and weaknesses of commercial adventure games based on ESL/EFL learners' perceptions?

2. Methodology

2.1 Participants

Sixty undergraduate freshmen at a national university in Taiwan participated in this study. These participants were all non-English majors from different subject disciplines, including Education (n=34), Liberal Arts (n=9), Science (n=8), Business Management (n=4), Fine Arts (n=3), Music (n=1), and Technology (n=1). Both male (n=10) and female (n=50) students participated in this study, and their average age was around 19.

2.2 Instrument

A commercial adventure game called *Back to the Future (BTTF): The Game* was purchased and provided to the participants. The video game is largely based on the story of the famous science fiction movie Back to the Future. In the game, the player controls the leading character, Marty, to explore the 3D environments by interacting with the non-player characters (NPCs) and surrounding objects (Figure 1). Some objects in the game can be picked up and stored in the inventory, and then can be used later to interact with other characters or objects to proceed in the game.

As a commercial adventure game, the *BTTF* game provides abundant textual and aural language input. When playing, the player can simultaneously hear the dialogues spoken by English native speakers and read the English subtitles (Figure 2). In addition to subtitles, all the usable objects in the game are labeled. Hovering on an item shows its name (Figure 3). The game also includes great graphic/sound effects, motivating game tasks, and simple game controls. The player can engage in different demanding tasks using the mouse, keyboard or game controller. In case the game missions are too challenging, the game also provides a list of goals and built-in hints for the player to advance the game. Because of its multiple language input (e.g., textual, pictorial, and aural) and attractive game features, *BTTF* was chosen as the instrument in this study.

Figure 1. Interacting with NPC

Figure 2. Synchronous dialogue and subtitle

Figure 3. Presentation of word and symbol

2.3 Data Collection

All participants played the *BTTF* game in a laboratory setting. Because it takes at least 8 hours to complete the whole *BTTF* game, all participants were allotted one hour to experience the game in case they might be exhausted from a long hands-on session. After one hour, all participants were asked to take a five-point Likert scale survey questionnaire. The questionnaire investigated the participants' perceptions of a) playing the game toward language learning and b) the game design of the *BTTF* game. In addition to the Likert scale survey items, the participants were required to answer some open-ended questions regarding a) possible language gains after game-playing and b) the strengths/weaknesses of the game. Both the questionnaire and the open-ended questions aimed to explore the possible language gains as well as motivating features of adventure video games.

3. Results

3.1 Perception of Back to the Future for Language Learning

The results of the survey, as presented in Table 1, showed that students considered the *BTTF* game helpful in improving their English ability (M=4.15). Among the investigated language ability/knowledge, listening (M=4.17), reading (M=4.00), and vocabulary (M=4.17) are the skills that benefited the most from game playing, whereas skills/knowledge like speaking (M=3.00), writing (M=2.78), and grammar (M=2.78) benefited less. Students also showed positive attitudes toward the helpfulness of adventure video games in English learning (M=4.12) and expressed their willingness (M=4.18) to use this kind of games for language learning.

Table 1. Mean scores for the perceptions of the game's helpfulness for language learning

Items	Mean (SD)
The game helps me improve my English listening ability.	4.17 (0.64)
The game helps me improve my English speaking ability.	3.00 (0.92)
The game helps me improve my English reading ability.	4.00 (0.64)
The game helps me improve my English writing ability.	2.78 (0.69)
The game helps me improve my English vocabulary knowledge.	4.17 (0.67)
The game helps me improve my English grammar knowledge.	3.20 (0.73)
In general, the game helps me improve my English ability.	4.15 (0.55)
I think playing adventure video games helps me improve my English ability.	4.12 (0.56)
I am willing to use adventure video games to improve my English ability.	4.18 (0.75)
Average Mean (SD)	3.75 (0.68)

Students' responses to the open-ended question regarding their language gains also reveal similar results to that of the survey questionnaire, as shown in Table 2. Students commented that playing the game could enhance their listening ability and the presentation of subtitles helped them comprehend some fast dialogues. The instructions/subtitles also helped to increase their reading ability and speed. In addition, playing the game provided a chance for them to learn many new words in a context-rich, all-English environment, and it also introduced them to some new colloquial usages. Students also reported that learning English through playing the video game enhanced their learning motivation in an enjoyable way.

While students reported some language gains after gaming, they also identified several language barriers. For example, some students complained that the fast dialogues and the fleeting subtitles caused them some problems in comprehending the text. The language used in the game, either its structure or vocabulary, was also difficult to several students. A small number of students considered the game less helpful in improving their speaking and writing ability, and expressed their doubts as to whether their English skills could benefit from game playing.

Table 2. Summary of participants' comments on language/learning gains and difficulties

		Gains	Difficulties
Language	**Listening**	• Enhances general listening ability (22) • Enhances listening with the aid of subtitles (4)	• Fast dialogue (7)
	Speaking		• Less beneficial (1)
	Reading	• Enhances general reading ability (11) • Increases reading speed (1)	• Fleeting subtitles (5)
	Writing		• Less beneficial (1)
	Vocabulary	• Enlarges vocabulary size (22) • Learning vocabulary in context (14) • Recognizing learned vocabulary (1)	• Difficult vocabulary (2)
	Grammar		

Others	• Acquiring new colloquial expressions (2) • Engaging in an all-English environment (1)	• Difficult language (1)
Learning	• Enhances learning motivation (7)	• Not educational (1)

Note. The identified times of each item are given in parentheses

3.2 Perception of the design of Back to the Future

In addition to the helpfulness of the video game, students also expressed their fondness for the game design. They liked the storyline (*M*=4.37), the graphic design (*M*=4.07), and the audio design (*M*=4.08) of the game the most. Compared to these features, students regarded the given instructions/hints (*M*=3.75), the difficulty level of the game missions (*M*=3.65) and the language level (*M*=3.52) as the less satisfactory designs of the game. Nevertheless, they still had a positive impression of the overall game design and were motivated to play the game (*M*=4.07). Table 3 shows the means and standard deviations for the perceptions of the game design.

Table 3. Mean scores for the perceptions of the game design

Items	**Mean (SD)**
The game is easy to control and gives clear/sufficient instructions/hints.	3.75 (0.73)
The game missions are moderately changeling.	3.65 (0.86)
The difficulty level of the language used in the game is moderate.	3.52 (0.77)
The storyline of the game attracts me.	4.37 (0.82)
The graphic design of the game attracts me.	4.07 (0.84)
The audio design of the game attracts me.	4.08 (0.83)
The game can attract me to keep playing.	4.07 (0.71)
Average Mean (SD)	3.93 (0.79)

As for responses to the open-ended question, students' comments of the game design echoed with the survey results. They identified the intriguing storyline of the game as one of the salient strengths, and recognized the gameplay design (e.g. game missions, character development, useful hints, multiple routes, simple game controls, and entertaining cut-scenes) as the most satisfactory design of the game. They also liked the game's delicate art design and great audio design (dubbing, sound effects, and background music). Students reported that they gained a sense of achievement after complete the game missions, and that they were engaged in playing the game for aforementioned attractive features. They also thought that the subtitles and item names provided in the game helped them to proceed in the game more smoothly.

Although many students identified the gameplay design as the most welcomed design of the game, some students also pointed out several unsatisfactory features. Among the identified weaknesses, unclear/insufficient hints and challenging game missions were the two major features that students complained about, which echoed the statistical results. Students reported that certain missions require complex procedures to accomplish, and they did not have the know-how to play this kind of game. Even though they referred to the built-in hints, they sometimes still did not know what they should do to achieve the missions. A number of students suggested that the built-in hints should be more detailed and explicit so that they can know what to do to solve the game missions. In addition, some students considered the pacing of the game somehow slow, and they also complained that there were too many dialogues in the game. A small number of students also expressed difficulties in controlling the leading character's movements. While many students liked the storyline and the art design of the game, a few students felt the storyline was boring and the graphic design was unattractive. Table 4 presents both the identified strengths and weaknesses of the game.

Table 4. Summary of participants' comments on the strengths and weaknesses of the game

	Strengths	Weaknesses
Art Design	• Delicate graphic design (12) • Vivid character depiction (2)	• Poor graphic design (4) • Poor 3D quality (3)
Audio Design	• Great dubbing (4) • Great sound effects (3) • Great background music (1)	• Poor sound effects (3)
Gameplay Design	• Intriguing storyline (24) • Engaging game missions (11) • Attractive character development (6) • Sufficient/useful hints (5) • Multiple routes (4) • Simple game controls (3) • Entertaining cut-scenes (2)	• Unclear/insufficient hints (6) • Difficult/complex game missions (6) • Slow gaming pace (5) • Lengthy dialogues (4) • Hard to control the character's movements (3) • Boring storyline (2)
Others	• English subtitles/words provided (7) • Sense of achievement (5) • Entertaining and appealing (4)	• Poor alignment of the dialogues and subtitles (4)

Note. The identified times of each item are given in parentheses

4. Discussion

This study investigated ESL/EFL students' perceptions of a commercial adventure video game for language learning. Both students' perceptions of language gains and of game design after playing the game were examined, and the study yielded the following major findings.

First, students held positive attitudes toward the use of this kind of game for language learning, and they believed this game is particularly helpful in improving their receptive language abilities/knowledge such as vocabulary, listening, and reading. The context-rich learning environment of the game allow students to receive multiple language input (e.g., textual, pictorial, and aural). For example, many students reported that the game could increase their vocabulary size by presenting new words with symbols. Picture aids have been proven to be an important factor for effective vocabulary learning and retention [23]. It is thus reasonable for students to consider the game to be beneficial for vocabulary acquisition.

Although students reported improved receptive language skill/knowledge, they thought the game was less beneficial to their productive language knowledge (e.g. writing and speaking). Students were constantly receiving language input when playing the game, but they had limited opportunities to produce language output. Even though they could choose different replies whenever encountering a dialogue tree, the pre-set, fixed replies did not allow students to reconstruct the sentence structures. Students were also not allowed to create their own sentences as replies. The lack of allowing creative language production could be seen as one of the major limitations of using commercial adventure video games for language learning. Hence, for educators who intend to facilitate language learners' speaking and writing skills through commercial adventure games, they should be more careful when adopting this type of game for language learning.

In terms of their perceptions of the game, students expressed their fondness for the overall game design and identified several attractive features that motivated their interest both in gaming and learning. To most of the students, the storyline and the challenging game missions are the two major attractive factors that motivate them to use the game. The challenging game missions, however, were also regarded as one of the major drawbacks of the game design to some students. The discrepancy in students' perceptions of the game missions might result from different language proficiency between students. As mentioned in the previous section, certain students commented that the language used in the game was somehow overwhelming to them. It is possible that they did not know how to deal with the current mission because they could not fully comprehend (or miscomprehended) the conveyed messages. They thus failed to complete certain missions and considered these game missions difficult.

In addition to language proficiency, students' gaming proficiency might also contribute to the discrepancy in the difficulty level of the game missions. Some students in this study already specified that they did not know how to play this kind of game, even though they have no difficulties in decoding the language. For less experienced adventure gamers, it might require more effort to familiarize themselves with the needed skills to complete the game missions. The lack of related gaming experiences might influence learners' perception of video games for language learning, which was already reported in Lee, Cheon and Key's study [11]. Hence, it would be worthwhile to investigate whether language proficiency and gaming experiences can positively/negatively affect learners' perceptions of adventure video games for language learning.

5. Conclusion

Generally speaking, this survey study shows that learners enjoy playing commercial adventure video games and consider these games useful tools to facilitate second/foreign language learning. For further study, it is suggested to conduct empirical research to further explore how and to what degree such games can assist learners in enhancing the previously identified language aspects. Factors such as learners' language proficiency levels and gaming experiences should also be included to determine what types of learners will benefit more from playing commercial adventure video games.

Acknowledgements

This research was supported by the National Science Council under Grant 98-2410-H-003 -079 -MY3 and supported in part by "Aim for the Top University Plan" of the National Taiwan Normal University and the Ministry of Education, Taiwan, R.O.C

References

[1] Baltra, A. (1990). Language learning through computer adventure games. *Simulation & Gaming, 21*(4), 445-452.

[2] Chik, A. (2011). Learner autonomy development through digital gameplay. *Digital Culture & Education*, *3*(1), 30-45.
[3] 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 & Education*, *59*, 661-686.
[4] deHaan, J. W. (2005). Acquisition of Japanese as a foreign language through a baseball videogame. *Foreign Language Annals*, *38*(2), 282-286.
[5] deHaan J. W., Reed, W. M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, *14*(2), 74-94.
[6] Foreman, J. (2003). Next-generation educational technology versus the lecture. *EDUCAUSE Review*, *38*(4), 12-22.
[7] Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave Macmillan.
[8] Gee, J. P. (2004). Learning by design: Games as learning machines. *Interactive Educational Multimedia*, *8*, 15-23.
[9] Gee, J. P. (2007). *Good video games and good learning: Collected essays on video games, learning and literacy (New literacies and digital epistemologies)*. New York, NY: Peter Lang.
[10] Kirriemuir, J., McFarlane, A. (2004). *Literature review in games and learning*. Bristol: NESTA.
[11] Lee, Y.-Y, Cheon, J., & Key, S. (2008). Learners' perceptions of video games for second/foreign language learning. In K. McFerrin et al. (Ed.), *Proceedings of Society for Information Technology and Teacher Education International Conference* 2008 *(pp. 1733-1738).* Chesapeake, VA: AACE.
[12] Miller, M. & Hegelheimer, V. (2006). The SIMs meet ESL: Incorporating authentic computer simulation games into the language classroom. *Interactive Technology and Smart Education, 3*(4), 311 – 328.
[13] Novak, K., & Nackerud, R. (2011). Choosing a serious game for the classroom: An adoption model for educators. In M. Ma et al. (Eds.), *Serious Games and Edutainment Applications* (pp. 291-308). London: Springer-Verlag London Limited.
[14] Peterson, M. (2011). Digital gaming and second language development: Japanese learners interactions in a MMORPG. *Digital Culture & Education*, *3*(1), 56-73.
[15] Prensky, M. (2001). *Digital game-based learning.* New York, NY: McGraw-Hill.
[16] Prensky, M. (2002). The motivation of gameplay: The real twenty-first century learning revolution. *On the Horizon*, *10*(1), 5-11.
[17] Ranalli, J. (2008). Learning English with The Sims: Exploiting authentic simulation games for L2 learning. *Computer Assisted Language Learning, 21*(5), 441-455.
[18] Rankin, Y., Gold, R., & Gooch, B. (2006, September). *3D role-playing games as language learning tools*. Paper presented at the meeting of the Eurographics 2006, Vienna, Austria.
[19] Rankin, Y., Morrison, D., McNeal, M.C., Gooch, B., & Shute, M. (2009). *Time will tell: In-game social interactions that facilitate second language acquisition*. In R. M. Young (Ed.), *ICFDG 2009* (pp.161-168). New York, NY: ACM.
[20] Reinders, H., & Wattana, S. (2011). Learn English or Die: The effects of digital games on interaction and willingness to communicate in a foreign language. *Digital Culture & Education*, *3*(1), 3-29.
[21] Szynalski, T. P. (n.d.). *Learning English with adventure games.* Retrieved from http://www.antimoon.com/how/advgames.htm.
[22] Thorne, S. L. (2008). Transcultural communication in open Internet environments and massively multiplayer online games. In S. Magnan (Ed.), *Mediating discourse online* (pp. 305-327). Amsterdam: John Benjamins.
[23] Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, *10,* 85-101.

Group Scribbles-Enhanced Collaborative Learning Improves Reading Skills: An Experiment Study in Primary Classrooms

Wenli CHEN [a*], Chiupin LIN[b], Wenting XIE [a] , Chingchun, CHU[b] & Yew Lee TAN[c]
[a]*Learning Sciences Lab, National Institute of Education, Nanyang Technological University, Singapore*
[b]*Institute of eLearning Technology, National Hsinchu University of Education, Taiwan*
[c]*Yu Neng Primary School, Singapore*
*wenli.chen@nie.edu.sg

Abstract: To improve reading skills of students in primary schools, we introduced Group Scribbles-enhanced collaborative learning into a language classroom. The quasi-experiment study conducted confirmed the positive role of technology intervention in improving learning effectiveness and improved learning attitudes. Further analysis carried out revealed the collaboration patterns emerged in the learning environment designed.

Keywords: CSCL, language learning, expository text, primary education,

Introduction

Researchers and practitioners on language learning have made great efforts to identify effective pedagogies to improve students' reading skills (Alvermann & Earle, 2003). The course of reading, in essence, is mining and analyzing information from the reading material and recalling and relating existing knowledge and past experiences to make meaning (Dole, et al., l991). This meaning-making process can be facilitated through inter-psychological interactions where meanings are negotiated, constructed and consolidated from multiple perspectives. Besides deepening comprehension, social interactions can also enhance students' motivation, interest and confidence in reading (Hollingsworth, et al., 2007; Spörer, et al., 2009). These benefits endorse the adoption of collaborative pedagogies where interactions between\among learners are pursued and promoted in language classrooms.

To better support student collaboration, network technologies are introduced into classrooms as the virtual medium is insulated from physical limitations. In our research, we are investigating how to leverage on networked technology-enhanced collaborative learning to improve students' language learning. In this study, one of our endeavors made in a primary school and its outcomes are reported. We hope the achievements and pitfalls encountered in our exploration can help the community better visualize the need and the method to engage technology intervention in collaborative language classrooms.

1. Research Background

In our school-based research, we introduced Group Scribbles (GS), a networked technology co-developed by Learning Sciences Lab and SRI International to a local primary school. Based on the metaphor of whiteboard and sticky notes for collaborative knowledge construction (Roschelle, et al., 2007), GS has been regarded as an effective and flexible tool for collaborative activity design and enactment in classroom settings (Looi, et al., 2011). Previous research has affirmed that GS enhanced collaborative learning can improve students' learning outcomes, attitudes and epistemology in learning science and math (Looi, Chen & Ng, 2010; Looi & Chen, 2011) and L2 (Chen, Wen, & Looi, In press) in primary schools. Here we translated GS-enhanced collaborative learning into L1 (Chinese language) classrooms. A quasi-experiment study was designed and implemented to examine whether the integration of online interaction could produce improved learning outcome. In the GS experiment class, student perceptions of the learning experience and interaction patterns emerged were also documented and analyzed. Through combining triple sources of data, i.e. learning performance, perception and process data, the role of technology intervention in promoting classroom collaboration and language learning could be better revealed.

2. Research Design

2.1 Participants

Two Grade 4 classes (each of 30 students) participated in our quasi-experiment design. We randomly chose one as the experiment class and the other as the control class. Students in the GS experiment class had received considerable ICT training since Grade 3 and attained good ICT literacy. After GS enculturation, they had developed satisfactory proficiency in using the technology. Students in each class were then equally distributed into 6 groups. In grouping, students of different L1 proficiency (indicated by students' Chinese language test scores before intervention) and gender were put into one group. The grouping was made so as it was proved in previous research that 3-5 students of different ability and gender composing one collaborative group benefits language learning (Salvin, 1985).

2.2 GS Technology

Unlike the control class where student collaboration was achieved through F2F interaction, students in the experiment class could engage in both F2F discussion and GS interaction in learning. In GS lessons, each student group was provided a laptop with GS installation. GroupScribbles 2.0 presents users with a two-paned interface encompassing a private working area, the "private board" in the lower section, and a public working area, the "public board" in the upper section (Figure 1). Students generate virtual pads of "scribbles" on the private board to draw, write and type in their ideas. All the actions performed and contents produced in this area are invisible to others. Scribbles are published and shared as students drag them onto the public board which is synchronized among all learning devices. Scribbles on public board can be removed, replaced, and withdrawn to private boards for editing. The essential feature of GS technology is the synergy of autonomous cognition (on private board) and collaborative cognition (on public board). The GS technology is highly customized as users can insert pictures, templates and movie clips on the public board. In our study, graphic organizers that helped students analyze text structures were incorporated.

2.3 Pedagogical Design

In our intervention, altogether 8 lessons on Expository Text reading (40 minutes per lesson, 2 lessons per week\per text) were designed and implemented. Student Teams Achievement Division (STAD) was adopted to guide the pedagogical design (Slavin, 1987). To facilitate students' meaning-making process, two reading strategies, namely text structure analysis and summarization, were employed. In text structure analysis, students were required to complete a graphic organizer for the given text by extracting the themes\topics introduced and connect ideas\information presented (Tree chart for conceptual structure; concept chart for conceptual structure; flow chart for sequential structure). In summarization, students were requested to sum up the main idea and write an abstract for each paragraph. All these reading activities could help students to better grasp the anatomy of the text and its key components. In the experiment class and the control class, identical learning activities were carried out with occasional teacher scaffolding (Table 1). To ensure consistency in instructions, lessons in both classes were delivered by the researcher.

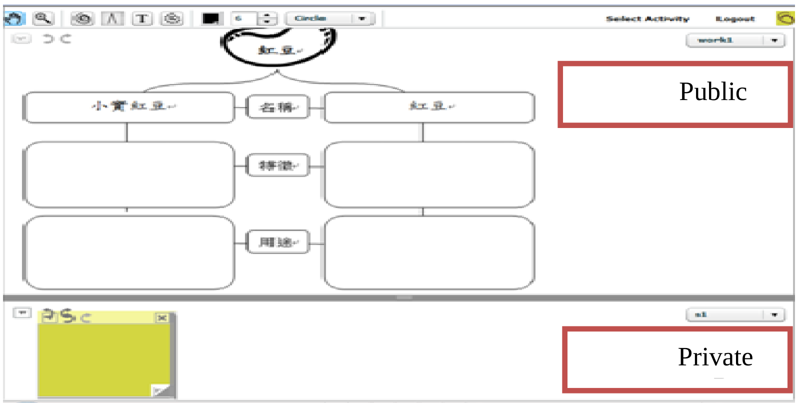


Figure 1. Group Scribbles Interface

Table 1. Collaborative Pedagogical Design

Step	Learning Activity	Description
1	Topic Introduction	Teacher introduces the theme\topic of the expository text.
2	Text Reading	Students read the text.
3	Text Discussion	Teacher proposes questions and leads student to discuss these questions.
4	Text Structure introduction	Teacher explains and elaborates on the structure of expository texts.
5	Text Structure Analysis	Students interact within the group to complete the graphic organizer provided (experiment class: F2F + online; control class: F2F).
6	Group Sharing 1	Each group presents the completed graphic organizer. Students review and comment on the graphic organizers presented. (experiment class: GS; control class: pen & paper)
7	Text Summarization	Students interact within the group to compose abstracts for each paragraph.
8	Group Sharing 2	Each group presents the completed paragraph abstracts. Students review and comment on the abstracts presented. (experiment class: GS; control class: pen & paper)
9	Reward & Round-up	Students vote for the best group work. Teacher rewards the best group and rounds-up the lesson.

3. Data Analysis & Discussion

3. 1 Performance Analysis: GS Enhanced Improvement in Students' Reading Skills

To measure students' improvement in reading skills, a reading comprehension test was developed. Three types of questions were incorporated in the test paper: 1) "Literal" questions, the answers to which can be obtained by "quoting" the text; 2) "Inferential" questions, the answers to which can be obtained by drawing inferences and implications after analyzing ideas\information embedded in the text; 3) "Integrated interpretation" questions, the answers to which can only be obtained by associating and relating ideas\information embedded in the text, student's existing knowledge and personal experiences. In the test paper, there were both multiple-choice questions (including all three question types) and open-ended questions (type 3 questions only). To ensure validity of the test paper, we invited expert teachers to review the questions constructed. A pilot test in another Grade 4 class was administered to further improve the test. According to Ebel & Frisbie (1991), the level of difficulty and discriminability of good test items should fall in the range of 0.4-0.8 and 0.4-1 respectively. After several rounds of modification, the 13 test items developed on average reached good difficulty (0.54) and discriminability (0.71). Moreover, a Pearson correlation analysis between student Chinese test scores and pretest scores was conducted. The strong correlation observed (experiment class: r= .772, p= .001; control class: r= .936, p= .000, table 2) further suggested the test paper crafted was of good validity. A pre-test and post-test design was employed to assess student learning gains. In both rounds of test, the same items were used but presented in different orders.

Table 2. Correlation analysis between Chinese test scores and pretest scores

Class	Test	Mean	SD	Correlation	Sig.(2- tail)
Experiment class	Pre-test	91.78	5.719	.772	.000
	Chinese Test	34.23	10.846		
Control class	Chinese Test	92.61	4.540	.936	.000
	Pre-test	36.97	11.577		

Considering students in two classes might vary in reading competence, when comparing their performances in the post-test, we used their pre-test scores as the covariant. The assumption of Analysis of Covariance (ANCOVA) was met as the regression coefficients between the dependent variable and covariant were consistent in the two classes ($F=2.216$, $p=.142 > .05$). Analysis shows that there was significant difference in the post test between these two classes that couldn't be explained by the discrepancy in the pre-test ($p= .045 < .05$, table 3). Therefore conclusion can be drawn that GS intervention has produced enhanced improvement in students' reading skills in the collaborative classroom.

Table 3.1 Student post-test scores

Class	Experiment class			Control class		
	M	SD	Progress	M	SD	Progress
Pre-test	51.07	16.599	21.33	54.47	17.878	17.66
Post-test	72.40	11.060		72.13	13.574	

Table 3.2 Comparison of student scores in post-test: ANCOVA analysis

Sources	Type III Sum of Squares	df	MS	F	Sig.
Pre-test	7611.080	1	7611.080	339.040	.000
Class	94.645	1	94.645	4.216	.045
Error	1279.587	57	22.449		

3. 2 Perception Analysis: GS Enhanced Collaboration & Learning Interest

In this study, both quantitative and qualitative data were mined and analyzed to probe students' attitudes toward collaborative learning and GS learning activities. The quantitative data came from a survey conducted after GS intervention. In the survey questionnaire, a 5 point Likert scale was used (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree). The higher the score was, the more students agreed with the statement given. In-depth semi-structured interviews were administered to obtain qualitative data. All students in the GS experiment class participated in the interview. The interview sessions were recorded, transcribed and translated for analysis.

3.2.1 Student Perception of Collaborative Learning

In the questionnaire, four items were generated to probe students' perception of collaborative learning. As indicated in the descriptive data (Table 4), most students held positive attitude towards classroom collaboration. In GS lessons, students shared and negotiated ideas within a group, and seek and offer help in times of need.

Table 4 Descriptive statistics of student perceptions of Collaborative Learning (N=30)

Item	Strongly Agree%	Agree %	Neutral %	Disagree %	Strongly Disagree%	Mean	SD
I would like to share ideas in my group.	42.9	28.6	25.0	3.6	0	4.11	0.916
I would like to accept different ideas in my group.	46.4	32.1	17.9	3.6	0	4.21	0.876
I would like to assist others in my group.	46.4	25.0	25.0	3.6	0	4.14	0.932
I would like to seek help form others in my group.	39.3	32.1	25.0	3.6	0	4.07	0.900

In interview, students explained and elaborated on the benefits they received from collaborative learning in the language classroom, which could be summarized as:

1) Classroom collaboration promoted confidence in students

Compared with individual learning, students were more willingly to express ideas in group discussion as they could receive peer feedback and assistance, with which they improved their answers. This promoted their confidence and encouraged their participation in class.

--"When learning in a group, we can discuss our answers. Thus I can always come up with an answer to the question, no matter how difficult it may be."(S3, S21)

--"You can ask your group members for help if you don't know the answer."(S8)

--"You can participate in the activity even when you don't have many ideas."(S21)

2) Classroom collaboration encouraged sharing and negotiation of ideas

Students had more opportunities to air their opinions in collaborative learning scenarios. In group discussion, ideas from multiple perspectives were discussed, reflected on, and synthesized, based on which answers of improved depth and breadth could be constructed.

--"In group discussion, you can express your ideas and opinions freely."(S17)

--"You can learn from different ideas proposed within the group."(S20)

3) Classroom collaboration improved collaborative skills in students

Collaborative problem solving demanded mutual engagement and coordination among all the participants. If students were more involved in group work, they could attain better skills in communication and show more respect and appreciation to others' work, all good to development in collaborative skills.

--"In group work, you will learn to collaborate with others to finish the task."(S21)

--"Apart from knowledge improvement, I have learnt to collaborate with others." (S24)

--"After group work, I am more aware of the importance of collaboration."(S10, S26)

--"You have to listen to others' opinions. You cannot only count on yourself."(S22)

4) Classroom collaboration nurtured good relationship among students

When engaged in group work, students had more opportunities to communicate and interact with each other. This helped breaking the ice among the students. Through group work, students became more aquatint with each other and had made more friends.

--“Discussing with others can develop better relationships between us.”(S15, S23, S29)
--“Group work promotes interaction among us.” (S17, S20,S23)
--“In group discussion, we develop better understanding about each other.”(S23)
--“I have more chances to communicate with the ones I am not familiar with.”(S18)

3.2.2 Student Perception of GS Learning Activities

To measure students’ attitudes toward GS learning activities, we examined how students perceived about using the GS technology and participating in GS activities. In the questionnaire, 4 questions items were on GS technology adoption. Data analysis unveiled that generally students held positive attitude towards GS. With GS, they could express their opinions and initiate discussions with ease and comfort.

Table 5 Descriptive statistics of student perceptions of using GS

Item	Strongly Agree%	Agree %	Neutral %	Disagree %	Strongly Disagree%	Mean	SD
GS is easy to use.	50.0	32.1	14.6	3.6	0	4.29	0.854
I can express my opinions easily on GS.	25.0	46.4	25.0	3.6	0	3.93	0.813
I don’t think it is difficult to discuss with others on GS.	50.0	25.0	21.4	0	3.6	4.18	1.020
I work smoothly on GS without encountering any trouble.	39.3	32.1	17.9	10.7	0	4.00	1.018

Student interview data shows that the integration of GS technology promoted student interest and motivation in learning:
--“Using computers to learn is very interesting.”(S10, S14, S23, S25)
--“It’s much more boring in traditional classrooms.”(S20)
--“You don’t have to write down your ideas. You can express them simply by typing.”(S13)

In student interview, areas for improvement were also revealed. Some students described the technical issues they experienced. There were occasions when they couldn’t type in the words or publish\move the scribbles. Sometimes, scribbles published would disappear mysteriously. As each group was only equipped with one computer, some ideas that were orally expressed were not timely and sufficiently documented in the virtual medium.

How students perceived GS activities was another important dimension examined in the survey. Data shows students in general held favorable opinions on GS activities (Table 6).

Table 6 Descriptive statistics of student perceptions of GS learning activities

Item	Strongly Agree%	Agree %	Neutral %	Disagree %	Strongly Disagree%	Mean	SD
In GS lessons, I have more opportunities to express my ideas.	50.0	25.0	21.4	3.6	0	4.21	0.917
In GS lessons, I can learn different ideas from others.	57.1	28.6	14.3	0	0	4.43	0.742
In GS lessons, I can develop better relationships with others.	67.9	7.1	21.4	3.6	0	4.39	0.956
GS lessons are interesting.	57.1	28.6	10.7	3.6	0	4.39	0.832
I prefer GS lessons.	60.7	21.4	10.7	3.6	3.6	4.32	1.056
I would like to have more GS	67.9	10.7	21.4	0	0	4.46	0.838

lessons.

3.3 Process Analysis: Group Collaboration Patterns

We also examined group collaboration patterns in GS lessons. Process data obtained included group videos and field observation notes. After analysis, 2 collaboration patterns emerged in 6 students groups (Figure 2). In interactive groups, work was distributed among group members and they took turns to control the computer. Communication and interaction was frequent in the group. Though high ability students were dominating the talk, the ones of medium and low ability also contributed their ideas and opinions. In fragmented groups, computer was under the control of a single student and some students (usually low ability ones) were isolated from group discussion and remained silent. In these groups, communication and interaction mostly occurred between two students. The fact that not all groups achieved satisfactory collaboration shows that long term efforts are needed to develop the collaborative skills in students and nurture the collaborative culture in class.

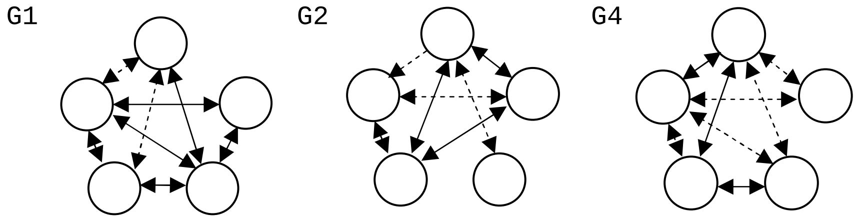


Figure 2.1 Interactive Group

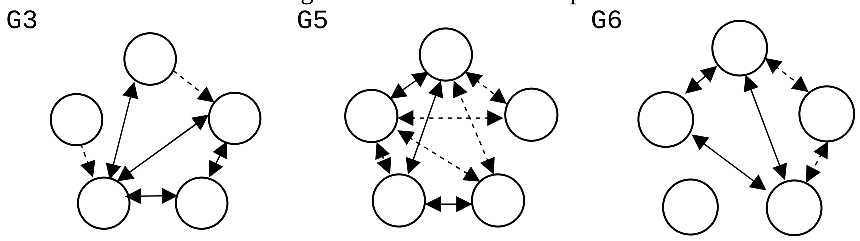


Figure 2.2 Fragmented Group

* H:high ability student; M: Medium ability student; L: Low ability student

* → : Communication of high frequency;

--->: Communication of low frequency

4. Conclusion

In this paper, we reported how we integrated network technology in a primary reading class to enhance collaborative learning. Through a quasi-experiment design, we confirmed the role of technology intervention in producing enhanced learning gains and improved attitudes. Problems encountered (concerning learning environment design and collaborative culture development) were also elaborated. However, as the study was of small scale and context specific, any application of the findings should be done with caution.

Reference

[1] Alvermann, D., & Earle, J. (2003). Comprehension instruction. In A.P. Sweet, & C. Snow (Eds.), *Rethinking reading comprehension* (pp. 12-30). New York: Guilford.

[2] Chen, W., & Looi,C.K. (2011). Active classroom participation in a GroupScribbles primary science classroom. *British Journal of Educational Technology, 42(4),* 676-686.

[3] Chen, W., Wen, Y., Looi, C.K. (In Press). Technology enhanced pedagogical innovation in second language learning. Paper accepted by Global Chinese Journal for Computers in Education.

[4] Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: research on reading comprehension instruction. *Review of Educational Research, 61(2),* 239-264.

[5] Ebel, R.L., & Frisbie, D.A.(1991). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice Hall.
[6] Hollingsworth, A., Sherman, J., & Zaugra, C.(2007). Increasing reading comprehension in first and second grades through cooperative learning, Unpublished MA Thesis. Saint Xavier University & Pearson Achievement Solutions, Chicago.
[7] Looi, C.K., Chen, W.L., & Ng, F.K. (2010). Collaborative activities enabled by Group Scribbles (GS): an exploratory study of learning effectiveness. *Computers & Education,* 54, 14–26.
[8] Looi, C.K., So, H.J., Toh, Y., & Chen, W. (2011). The Singapore experience: Synergy of national policy, classroom practice and design research. *International Journal of Computer-Supported Collaborative Learning*, *6*, 9-37.
[9] Roschelle, J., Tatar, D., Chaudhury, S. R., Dimitriadis, Y., Patton, C., & DiGiano, C. (2007). Ink, improvisation, and interactive engagement: Learning with tablets. *Computer, 40 (9)*, 42-48.
[10] Salvin, R. E. (1985). *Learning to cooperate, cooperating to learn*. New York: Plenum.
[11] Slavin, R. E. (1987). Developmental and motivational perspectives on co-operative learning. A Reconciliation. *Child Development, 58,* 1161-1167.
[12] SRI International. (2006) http://GroupScribbles.sri.com/
[13] Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction, 19*(3), 272-286.

EFL Learners’ Cognitive Load of Learning Vocabulary on Mobile Phones

Chih-cheng LIN and Ya-chuan YU
National Taiwan Normal University, Taiwan
cclin@ntnu.edu.tw

Abstract: The present study examined the effects of MMS vocabulary learning and the extent of cognitive load in different multimedia modes. This study of within-subject design recruited 32 eighth graders in central Taiwan to study 36 target words divided into four sets. Each set was presented in one of the four multimedia modes: text, text-audio, text-picture, and text-audio-picture. Immediately after learning each set all participants took a vocabulary test and a cognitive load questionnaire; and, they took a delayed test after two weeks. The results of ANOVA showed no significant differences in vocabulary gains and retention among the four modes, but in cognitive load. The text-audio-picture mode imposed lower load on our adolescents than the text mode and the text-picture mode; also, the text-audio mode induced lower cognitive load on learning than the text mode. Semantic simplicity in the target words and our learners’ repeated exposures could contribute to leveling off the differences in vocabulary gains and retention. Audio references, on the other hand, helped our learners associate forms with sounds and enhanced the input from an additional channel. The findings not only supported features of portability and immediacy in mobile learning but also confirmed the modality effect and the temporal contiguity effect of content presentations in multimedia learning.

Keywords: Cognitive load, mobile phones, multimedia messaging service (MMS), multimedia modes, vocabulary learning

1. Introduction

Mobile phones nowadays serve as useful tools for communication and for language learning. The effectiveness of SMS (Short Message Service) vocabulary learning has been widely discussed and confirmed in a number of studies (Cavus & Ibrahim, 2009; Kennedy & Levy, 2008; Levy & Kennedy, 2005; Lu, 2008; Thornton & Houser, 2005, Zhang, Song & Burston, 2011). More recently MMS (Multimedia Messaging Service) messages, involving texts, pictures, audios, and videos, offer an alternative way of vocabulary learning in both mobile and multimedia environments. While multimedia information can be beneficial, learning is likely to be less effective due to limited human cognitive capacity. Based on the Cognitive Load Theory (Pass, Tuovinen, Tabbers, & van Gerven, 2003; Sweller, van Merriënboer, & Pass, 1998), different multimedia presentations of instructional materials may impose various degrees of cognitive load, which could influence the effectiveness of learning. The purpose of the present study is to define the relationship between multimedia presentations and cognitive load in vocabulary learning on cell phones. This issue is addressed in two aspects: vocabulary gains and retention in different multimedia modes, and a comparison of the cognitive load imposed on learners by different multimedia modes.

2. Literature Review

2.1. Vocabulary learning in multimedia and mobile learning environment

Studies on multimedia annotations have been widely conducted to examine their effectiveness. The potentials of multimedia presentations of instructional materials have changed the way learners learning vocabulary. The advantages of encountering target words in versatile multimedia settings benefit language learners in retrieval and generative use. Previous research on multimedia vocabulary annotations supported the dual-coding theory (Paivio, 1986) and the generative theory of multimedia learning (Mayer, 1997). Vocabulary instructional materials presented in both verbal forms (i.e. texts) and visual forms such as pictures (Chun & Plass, 1996; Lin, 2009; Shahrokni, 2009; Yeh & Wang, 2003; Yoshii & Flaitz, 2002; Yoshii, 2006), animations, and videos (Al-Seghayer, 2001; Akbulut, 2007; Plass et al., 1998, 2003) were found to display better facilitative effects on vocabulary learning than in either verbal or visual forms only. Few studies, however, found that texts provided equally sufficient lexical information for vocabulary learning as graphics did (Acha, 2009; Yanguas, 2009).

In the mobile era, using mobile devices in language learning creates an environment without the limitations of time and space. It has extended traditional classroom learning to daily contexts. Mobile phones, compared with other mobile handhelds, have reached the highest penetration rate in most countries, including Taiwan. Considering its availability and small-sized interface, research investigating the potentials of mobile-phone assisted language learning mainly focused their efforts on vocabulary learning. In previous studies (Cavus & Ibrahim, 2009; Kennedy & Levy, 2008; Levy & Kennedy, 2005; Lu, 2008; Stockwell, 2010; Thornton & Houser, 2005, Zhang, Song & Burston, 2011), language learners considered vocabulary learning via mobile phones effective and motivating. Vocabulary lessons delivered through Short Messaging Service (SMS) were welcome among language learners primarily because of their screenful content for learning. As learners in Kennedy and Levy's (2008) study stated, their vocabulary knowledge about English was consolidated and extended in the process of SMS vocabulary lessons. The effects of multimedia annotations for vocabulary learning on cell phones were discussed in terms of their interactions and learners' cognitive processing abilities (Chen et al., 2008; Taki & Khazaei, 2011). For learners with high visual or high verbal abilities, annotations with pictures worked more effectively, while for learners with low cognitive abilities, the basic textual information was conducive to learners (Taki & Khazaei, 2011), and the pictorial one may cause cognitive overload (Chen et al., 2008).

2.2. Cognitive load

Cognitive load on learners varies with different instructional designs of materials and has a great impact on learning performances (Sweller et al., 1998; Pass et al., 2003). It can be classified as intrinsic cognitive load, extraneous cognitive load, and germane cognitive load. The learning information is referred to as the source of intrinsic cognitive load, which is determined by the level of element interactivity, that is, the degree to which the target learning element interacts and refers to other elements. Extraneous cognitive load, usually resulting from non-optimal instructional procedures, refers to the cognitive imposition on learners due to the instructional designs. Unlike intrinsic and extraneous cognitive loads that are related to materials, germane cognitive load concerns more about the use of learners' working memory resources to deal with intrinsic cognitive load. A focus of cognitive load research is to reduce the levels of extraneous cognitive load, which can be modified by various instructional designs (Pass et al., 2003; Sweller, 2010; Sweller et al., 1998).

Because cognitive load imposed on learners may positively or negatively influence their learning, Mayer and Moreno (2003) proposed important principles for designing multimedia instructional materials, two of which are related to the study. One is the modality effect, which states that when identical learning information is presented via dual modalities, visual and auditory, the capacity of humans' working memory is increased, and the cognitive load is reduced. The other principle is the temporal contiguity effect, indicating that when instructional materials in dual modalities are presented simultaneously, the cognitive load imposed on learners may be decreased.

Previous research on cognitive load on vocabulary learning seldom discussed its impact on learners nor directly measured the load on vocabulary learning. Plass et al. (2003) and Acha (2009), for instance, suggested that learners with low cognitive processing abilities would allocate more cognitive resources to handle pictures or videos. Similarly, Chen et al. (2008) indicated that learners with low cognitive processing abilities would experience cognitive overload when receiving vocabulary learning messages with both texts and pictures. Nevertheless, cognitive load discussed in the above studies was addressed based on learners' performance without further measurement to support their statements. Because learners' processing of information and learning materials are heavily influenced by the cognitive load, understanding how it works and how it is related to the learning outcomes becomes essential, especially in a multimedia learning environment, where various instructional designs can increase or decrease the extraneous cognitive load.

The present study, therefore, aims to evaluate the effectiveness of learning vocabulary with different multimedia annotations on cell phones and to examine the extent of cognitive load of different multimedia modes on learning vocabulary on cell phones.

3. Method

3.1. Participants

Thirty-two eighth graders in central Taiwan were recruited and were considered beginners of English. They had three English classes per week. All participants had their own cell phones and their parents' permissions.

3.2. Materials

Three veteran English teachers selected 36 target words from the word list of the intermediate level of GEPT (General English Proficient Test in Taiwan), including 12 nouns, 12 verbs, and 12 adjectives. Then, the target words were divided into four different multimedia mode groups: the text mode, the text-audio mode (audio), the text-picture mode (picture), and the text-audio-picture mode (combined). Each mode consisted of nine target words, that is, three nouns, three verbs, and three adjectives.

In the text mode, textual information including target words, syntactical categories, Chinese equivalents, and example sentences was offered. Words in the audio mode were presented with textual information and their audio references of target words and example sentences. In the picture mode, textual information and pictures illustrating the target words were provided. Finally, words in the combined mode included text, audio, and picture references. Figure 1 below shows examples of the four presentation modes of MMS vocabulary learning messages on different students' cell phones.

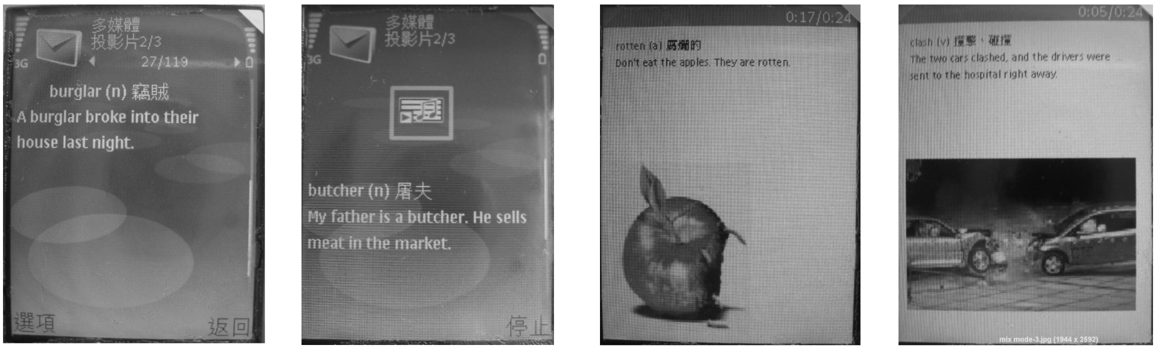


Figure 1. The four presentation modes of MMS vocabulary learning messages: (from left to right) text mode, audio mode, picture mode, and combined mode.

3.3 Procedures

In this within-subject study, all participants studied nine target words presented in one multimedia mode in one week. That required four different weeks to finish studying 36 target words. In each week, the nine target words were further divided into three groups. Each group contained one noun, one verb, and one adjective, which were composed of one MMS vocabulary lesson. In other words, it needed three MMS vocabulary lessons to finish delivering all nine target words. Because the English teacher demanded that each vocabulary item be learned twice, each MMS lesson was delivered to our participants twice every week. From Monday to Wednesday, the English teacher sent an MMS vocabulary lesson at 7 a.m. and 5 p.m., respectively. On Thursdays and Fridays, MMS lessons of jokes or short stories were delivered. The procedures of the study are shown below in Table 1.

Table 1. The procedures of the experiment

Week	Mode	Mon	Tue	Wed	Thu	Fri
1	Text	MMS 1a MMS 2a	MMS 3a MMS 1b	MMS 2b MMS 3b	2 lessons IM & CLQ	2 lessons
2	Audio & Picture	MMS 1a MMS 2a	MMS 3a MMS 1b	MMS 2b MMS 3b	2 lessons IM & CLQ	2 lessons
3		2 lessons	2 lessons	2 lessons	2 lessons Test	2 lessons DE on Text
4		2 lessons	2 lessons	2 lessons	2 lessons Test	2 lessons DE on Audio & Picture
5	Picture	MMS 1a MMS 2a	MMS 3a MMS 1b	MMS 2b MMS 3b	2 lessons IM & CLQ	2 lessons
6	Audio	MMS 1a MMS 2a	MMS 3a MMS 1b	MMS 2b MMS 3b	2 lessons IM & CLQ	2 lessons
7						DE on Picture
8						DE on Audio

Note. Words in 1a and 1b were identical and the order was different, so were those in 2a and 2b and those in 3a and 3b. IM=immediate post-test; CLQ=cognitive load questionnaire; DE=delayed post-test.

3.4. Instruments and data analysis

The instruments included an English vocabulary test on the target words and a cognitive load measurement (Paas & van Merriënboer, 1994), and a feedback questionnaire. One week before the experiment, all participants took the vocabulary test to check whether they had known any of the target words (pre-test). Then, the vocabulary test was administered after the vocabulary lesson of each mode (the immediate post-test) and two weeks after the lesson (the delayed post-test). The highest score for each set was 18 and the lowest 0. To measure their cognitive load, the participants were required to identify their mental load and mental effort on a seven-point rating scale, in which "1" referred to "very, very easy/low" and "7" meant "very, very hard/high." The highest possible score was 14 and the lowest 2. One-way ANOVA was employed to analyze the results of the pre-test and the cognitive load survey and two-way ANOVA to analyze those of the two post-tests.

4. Results

The results of the pre-test (M=0.00) showed no significant differences, suggesting an equivalence of knowing the four sets of target words. The descriptive statistics of the two vocabulary post-tests were presented in Table 2. The picture mode received the lowest scores of the four in both the immediate post-test (M=10.97, SD=4.020) and the delayed pot-test (M=9.16, SD=3.768). In the immediate post-test, our participants scored the highest on the target words presented in the audio mode (M=12.03, SD= 4.374); and in the delayed post-test, they seemed to favor the combined mode and scored the highest (M=10.41, SD= 3.723).

Table 2. Descriptive statistics for the vocabulary posttests

		Immediate		Delayed	
Mode	N	M	SD	M	SD
Text	32	11.19	3.922	9.22	3.625
Audio	32	12.03	4.374	9.63	3.415
Picture	32	10.97	4.020	9.16	3.768
Audio & Picture	32	11.88	4.172	10.41	3.723

Note. Maximum score = 18.

The two-way ANOVA revealed no significant differences among the four modes (F(3,248)=1.106, n.s.). It manifested significant differences between the two post-tests (F(1,248)=15.510, p<.05), in which the students scored higher in the immediate post-test (M=11.52, SD=0.344) than in the delayed post-test (M=9.60, SD=0.344). There were no significant interaction effects between the two independent variables (F(3,248)=.160, n.s.).

As far as the cognitive load on learners is concerned (Table 3), the text mode imposed the highest load (M=8.88, SD=1.540) while the combined mode the lowest (M=7.56, SD=1.813).

Table 3. Descriptive Statistics of the Cognitive Load Measurement

Mode	N	Mean	SD
Text	32	8.88	1.540
Audio	32	7.91	1.634
Picture	32	8.31	1.942
Audio & Picture	32	7.56	1.813

The results of the one-way ANOVA showed significant differences among the four modes ($F(3,93)=3.372$, $p<.05$). Pair-wise comparisons revealed the significant differences were between the text mode and the audio mode, the text mode and the combined mode, and the picture mode and the combined mode. Audio references seemed to help reduce our participants' cognitive load of learning new words.

5. Discussion

No significant differences of the four presentation modes were found in vocabulary gains and retention. Our adolescent learners showed no favoritism among words single-coded, dual-coded or triple-coded with multimedia information. That is, words with multimedia annotations delivered by MMS messages yielded similar learning results, as found in some previous research (Acha, 2009; Yanguas, 2009). Media-wise, the basic textual explanations of the target words found in all four modes, including Chinese equivalents and sample sentences, have already provided enough information for our participants. As Lin elaborated (2009), adolescent beginners do not need picture or audio references to learn words conveying common, daily-life concepts. A definition in L1 together with a sample sentence may have already created an essential context for our adolescent learners to learn the meanings of new words. Additional information, such as pictorial illustrations and pronunciations of words and sentences, may not be necessary for such vocabulary, which could explain why the picture mode received the lowest scores in both post-tests in the present study. The other clue for interpreting the similar results in learning vocabulary among the four modes could come from our participants' repeated exposures to the target words. The English teacher noted it was a common case that our participants would read a new MMS vocabulary lesson first and then review a previous lesson or two. Because of the screenful content of vocabulary lessons on cell phones, the participants could learn and review the lessons in a very short period. Occasions of their learning on the move included when they were waiting for a traffic light and when they were commuting. Our adolescent learners took advantages of the features of portability and immediacy found on mobile devices to help them learn. This naturalistic way of learning vocabulary found its support from many previous studies on mobile-assisted language learning (Levy & Kennedy, 2005; Lu, 2008; Stockwell, 2010). It certainly adds another piece of evidence for mobile-assisted vocabulary learning.

Significant differences in the cognitive load were found among the four presentation modes. More importantly, audio references played a crucial role in reducing our teenagers' cognitive load of learning vocabulary with multimedia annotations on cell phones. Of all the previous studies on multimedia annotations, only one study (Yeh & Wang, 2003) tackled the issue of sounds in vocabulary learning. The incorporation of the audio references was found inefficient due to the exceeding speed of reading, which their college students could not follow. The adolescent beginners of English in the present study, however, found that the audio references helped them reduce their cognitive load of learning English vocabulary with or without pictures, when compared with learning with basic textual explanations alone. They also perceived differences between the two modes with pictorial illustrations. To beginners, the pronunciation or phonetics of vocabulary is inseparable with the form or orthography (Nation, 2001). Learning new words without a demonstration of how they are pronounced in isolation and in context gives L2 beginners too much burden. Other than its position in vocabulary learning for beginners, audio references also find its support from multimedia learning. As Mayer and Moreno (2003) explain, instructional designs adopting the modality principle and the temporal contiguity principle help learners increase working memory and reduce cognitive load. Written forms of new vocabulary items in the present study must be regarded as learning targets, different from the redundant subtitles of audio clips in their studies and their models (Mayer & Moreno, 2003). Audio references for beginners of English, providing an additional channel through ears, enhance our learners' understanding of the phonetic structures represented by the written forms. This additional channel of audio input designed in both audio mode and combined mode helped our beginners of English perceive differences in their cognitive load on learning vocabulary.

6. Conclusion

The present study investigated the effectiveness of learning English vocabulary with multimedia annotations on cell phones. The similar results could be attributed to the semantic simplicity of the target words and repeated exposures to the target words. The study also examined the degrees of cognitive load imposed by various multimedia. The comparatively lower cognitive load in audio references could be derived from their letter-sound associations and their enhancement input from an additional channel. This discrepancy attributed by the audio references highlights various directions for future research. First and foremost, evidences for how audio references assist vocabulary learning and acquisition in multimedia environments are in need. Pedagogically, language practitioners long knew that phonetic structures were crucial for vocabulary learning and acquisition (Nation, 2001). Studies on multimedia annotations, however, haven't been able to support that with empirical data. This study could not provide any evidence, primarily because both modes incorporated with audio references received higher scores in different post-tests but they failed to reach any significance level. A design with more considerations, including word types, age groups, proficiency levels, and cognitive styles, is needed. Next, studies on incidental vocabulary learning on the move should be carried out to determine the role of audio references in mobile-assisted vocabulary learning. The participants in the present study were asked to study a word list delivered to their cell phones. There was no accompanied reading text in which the target words were embedded. A more naturalistic way to acquire vocabulary in the mobile age is to read texts with multimedia annotations. Whether or not audio references play a similar role in reading on the move as in paper or web-based reading is worth studying.

References

[1] Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. British Journal of Educational Technology, 40(2), 23-31.

[2] Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. Instruction Science, 35, 499-517.

[3] Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. Language Learning & Technology, 5(1), 202-32.
[4] Cavus, N., & Ibrahim, D. (2009). m-Learning: An experiment in using SMS to support learning new English language words. British Journal of Educational Technology, 40(1), 78-91.
[5] Chen, N.-S., Hsieh, S.-W., & Kinshuk. (2008). Effects of short-term memory and content representation type on mobile language learning. Language Learning & Technology, 12(3), 93-113.
[6] Chun, D. M., & Plass, J. L. (1996). Effect of multimedia annotations on vocabulary acquisition. The Modern Language Journal, 80(2), 183-98.
[7] Kennedy, C., & Levy, M. (2008). Using SMS to support beginners' language learning. ReCALL, 20(3), 315-30.
[8] Levy, M., & Kennedy, C. (2005). Learning Italian via mobile SMS. In A. Kukulska-Hulme & J. Traxler (Eds), Mobile learning: A handbook for educators and trainers (pp.76-83). London: Routledge.
[9] Lin, C. C. (2009). Learning action verbs with animation. The JALT Call Journal, 5(3), 23-40.
[10] Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. Journal of Computer Assisted Learning, 24, 515-25.
[11] Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 32(1), 1-19.
[12] Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38(1), 43-52.
[13] Nation, I.S.P. (2001). Learning Vocabulary in Another Language. Cambridge, UK, IN: Cambridge University Press.
[14] Pass, 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(1), 63-71.
[15] Pass F. G. W. C., van Merriënboer, J. J. G.., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. Percept. Motor Skills, 79, 419-30.
[16] Paivio, A. (1986). Mental representations: A dual-coding approach. New York, IN: Oxford University Press.
[17] Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. Computers in Human Behavior, 19, 221-43.
[18] Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. Journal of Educational Psychology, 90(1), 25-36.
[19] Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. Computers & Education, 34, 177-93.
[20] Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. Language Learning and Technology, 14(2), 95-110.
[21] Sweller, J., van Merriënboer, J. J. G., & Pass, F. G. W. C. (1998). Cognitive architecture and instructional design. Educational Psychology Review, 10(3), 251-96.
[22] Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. Educational Psychology Review, 22, 123-38.
[23] Taki, S., & Khazaei, S. (2011). Learning vocabulary via mobile phone: Persian EFL learners in focus. Journal of Language Teaching and Research, 2(6), 1252-1258.
[24] Thornton, P., & Houser, C. (2005). Using mobile phone in English education in Japan. Journal of Computer Assisted Learning, 21, 217-28.
[25] Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. Language Learning and Technology, 13(2), 8-67.
[26] Yeh, Y., & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. CALICO Journal, 21(1), 131-44.
[27] Yoshii, M. (2006). L1 and L2 glosses: their effects on incidental vocabulary learning. Language Learning & Technology, 10(3), 85-101.
[28] Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. CALICO Journal, 20(1), 33-58.
[29] Zhang, H., & Song, W. (2011). Reexamining the effectiveness of vocabulary learning via mobile phones. The Turkish Online Journal of Educational Technology, 10(3), 203-214.

Personalized Automatic Quiz Generation Based on Proficiency Level Estimation

Yi-Ting Huang[a*], Meng Chang Chen[b], Yeali S. Sun[a]
[a] *Department of Information Management, National Taiwan University, Taiwan.*
[b] *Institute of Information Science, Academia Sinica, Taiwan.*
*d97008@im.ntu.edu.tw

Abstract: Recent years have seen increased attention given to computer-aided question generation for language student testing and evaluation. However, this approach often directly provides examinees with exhaustive questions. This is inappropriate, because these questions are not designed for any specific testing purpose. In this work, we present a personalized automatic quiz generation model that generates multiple-choice questions at various difficulty levels and categories, including grammar, vocabulary, and reading comprehension. We combined this model with a quiz strategy for estimating examinee proficiency and question selection. The proficiency is estimated using Exponential Moving Average, combining the test responses with a student's past history. The results show that the subjects in the experimental group corrected their mistakes more frequently as well as answered more difficult questions than the control group. The experimental group also demonstrated the most progress between the pre-test and post-test. In addition, most of subjects agree the quality of the generated questions in the questionnaire analysis.

Keywords: Computer-aids question generation, personalized learning, adaptive test.

1. Introduction

Recent years have seen increased attention given to computer-aided question generation in the field of computer-assisted language learning. A growing number of studies are now available for designing different question types, such as multiple-choice test items [8-10], and cloze tests [5]. The first computer-aided question generation was proposed by Mitkov and Ha [10], and generated multiple-choice items by term frequencies or predefined syntactic patterns. Lin, Sung, and Chen [8] improved the design of Mitkov and Ha [10] by investigating the semantics of words and presenting vocabulary items, including collocation, antonym, synonym and similar word questions. The Sakumon system [5] developed a cloze test as another approach to automatically assess examinees' vocabulary and grammar skills. Beyond vocabulary assessment, Chen, Ko, Wu and Chang [3] focused on automatic grammar quiz generation. Lastly, the MARCT system [13] investigated reading comprehension and designed three question stem templates for generating questions. These studies propose methods for reducing the labor of manual question generation, instead automatically generating a list of all possible questions when given a document. However, this exhaustive list of questions is inappropriate for language learning, because it can lead to redundant, overly-simplistic test questions that are unfit for evaluating student progress. Moreover, the characteristic of items generated from these studies is insufficient. It is hard to facilitate meaningful test purpose and maximize examinees' learning outcomes.

In contrast to this approach, Item Response Theory (IRT) with computerized adaptive testing has steadily developed question selection from the relationship between an examinee's proficiency and the properties of the questions [4]. In order to create an adaptive test, IRT requires parameters, such as item difficulty parameter or item discrimination parameter, which can be determined using the items for a large number of samples first (as pre-calibration) and then manually derive the question parameters. However, their approach is more time-consuming than more automatic alternatives. In addition, in applying IRT, an examinee's ability can be obtained by observing responses during a test and then estimating using maximum likelihood estimation, maximum a posteriori or expected a posteriori [4]. While effective however, these methods only consider test responses at the time of testing, rather than incorporating this testing history.

While the studies above all investigate computer-aided question generation, little research has discussed the role of question difficulty during question generation. As a result, this study specifically examines three question types, each with various difficulty settings, relative to an examinee's proficiency level. Moreover, unlike previous research, this study incorporates an examinee's test history when estimating the proficiency. Together these improvements set this paper's model apart from the approaches briefly outlined above.

2. Automatic Quiz Generation

Figure 1 illustrates the traditional system architecture of the multiple-choice item generation process. Given the target learning material, items are created from the quiz generation process, composed of the stem generation, the answer determination, and the distractor determination. The stem generation forms a direct question or incomplete statement using predefined templates. The answer determination decides the correct answers of the question. Lastly, the distractor determination selects the plausible wrong alternatives as distractors from external resources, in order to discriminate good students from poor ones.

As shown in Figure 2, four questions (also called items) are generated from a document describing the origins of Halloween. In the document, the bolded sentences are selected by the stem generation to produce question stems, while the bold and underline word in the bolded sentences are decided by the answer determination. In the four questions, the bolded words represent stems, the bold italics are called answers, and the other plausible choices in the questions are distractors.

In this study, we generate three question types with difficulties based on the same system architecture. Those types are multiple-choice items, including vocabulary, grammar and reading comprehension. A stem is usually generated by predefined templates. In the following section, both the answer determination and the distractor determination are presented.

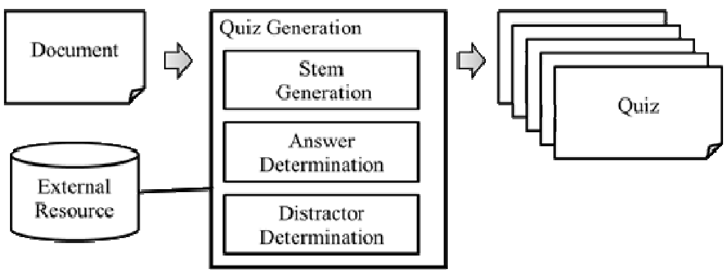


Figure 1. Overview of Traditional Automatic Quiz Generation.

2.1 Vocabulary Items

In this study, vocabulary items are generated according to an examinee's proficiency ability. The difficulty of a vocabulary question is based on the difficulty of the correct answer. We assume that if a student selects the correct answer, it is probable that he or she understood the question stem, and distinguished the correct answer from distractors. Here, word difficulty is determined by a grading word list which made by experts (http://www.ceec.edu.tw/research/paper_doc/ce37/5.pdf). When given the vocabulary proficiency level of an examinee, words with the same difficulty level in the given document are selected as the basis to form test questions.

The majority of research on vocabulary assessment extracts plausible distractors from various resources, such as a thesaurus [8] or words from the same document [10], and then designs algorithms to select the most suitable distractors for a given question. We consult the word list to ascertain vocabulary word difficulty, and then select distractor candidates of equal difficulty, part-of-speech, similar character length and small edit distance.

The first question in Figure 2 is an example of a vocabulary question. The question stem is composed from a manual template, the correct answer is found within an original document and taught in the fourth grade, and the difficulty level of the distractors is the same as the answer.

Document

Halloween, which falls on October 31, is one of the most unusual and fun holidays in the United States. It is also one of the scariest! **It is associated with ghosts, skeletons, witches, and other scary images.** …**Many of the original Halloween traditions have developed today into fun activities for children.** The most popular one is "trick or treat." On Halloween night, children dress up in costumes and go to visit their neighbors. When someone answers the door, the children cry out, "trick or treat!" What this means is, "Give us a treat, or we'll play a trick on you!" …This tradition comes from an old Irish story about a man$_1$ named **Jack**$_2$ who was very stingy. ... But he$_3$ also could not enter hell, because he$_4$ had once played a trick on the devil$_5$. **All he could do was walk the earth as a ghost, carrying a lantern**...

Quiz

1. **In the sentence "It is __________ with ghosts, skeletons, witches, and other scary images.", the blank can be:**
   (1) distributed (2) ***associated*** (3) contributed (4) illustrated
2. **In the Sentence, "Many of the original Halloween traditions __________ today into fun activities for children.", the blank can be filled in:**
   (1) ***have developed*** (2) have developing (3) is developed (4) develop
3. **The word "he" in this sentence "All he could do was walk the earth as a ghost, carrying a lantern" refer to:**
   (1) ghost (2) devil (3) witch (4) ***Jack***
4. **Which of the following statement is TRUE?**
   (1) On Halloween night, neighbors dress up in costumes and go to visit their children.
   (2) What this means is, "Give us a trick, or we'll play a treat on you!"
   (3) But the devil also could not enter hell, because he had once played a trick on the witch.
   (4) Jack was so stingy that he could not enter heaven when he died.

Figure 2. An Example in Automatic Quiz Generation.

2.2 Grammar Items

In this study, we manually predefine grammar patterns and distractor templates to generate grammar items. A set of 44 grammar patterns and distractor templates are identified from language learning textbooks. These grammar patterns are then implemented in the form of Tgrep2 patterns [7]. The difficulty of a grammar item is similar to the vocabulary item and determined by the difficulty of the correct answer. Unfortunately however, there is no predefined grammar difficulty measure available, similar to the aforementioned word list. Thus, we assigned the difficulty of a grammar pattern based on the textbook grade in which it frequently appears, which represents the age of grammar acquisition.

The second question in Figure 2 is an example of a grammar question. The target test purpose in the second question is "present perfect tense," taught in the first grade. The distractors refer to a grammar textbook to generate distractor templates about "tense." This approach helps clarify the difference between advanced students and non-advanced students. It distinguishes advanced learners who understand the implicit purpose of the question and identify the mistakes within the distractors from non-advanced learners who fail to comprehend the meaning of the question and choose the grammatical plausibility of the incorrect alternatives.

2.3 Reading Comprehension Items

In this work, we capture the relation between sentences to generate two kinds of meaningful reading questions based on noun phrase coreference resolution. Similar to Mitkov and Ha [10], who extracted nouns and noun phrases as important terminology in reading material, we also focus on the interaction of noun phrases as the test purpose. The purpose of noun phrase coreference resolution is to determine whether two expressions refer to the same entity in real life. An example is excerpted from Figure 2 (This tradition…on the $\underline{\text{devil}_5}$). It is easy to see that $Jack_2$ means man_1 because of the semantic relationship between the sentences. The following he_3 and he_4 are more difficult to judge as referring to $Jack_1$ or $devil_5$ when examinees do not clearly understand the meaning of the context in the document. This information is used in this work to generate reading comprehension questions, in order to examine whether learners really understand the relationship between nouns in the given context.

There are two question types generated in the reading comprehension item. One type is an independent referential question for the single concept test purpose, while the other follows one of the frequent question templates in [13], “which of the following statement is (not) true,” as the overall comprehension test purpose. When given a document, the coreferential relations are identified by the coreference system [11]. In the first type, noun phrases in the same coreference chain are selected as the correct answer, and noun phrases in the other coreference chains are determined as the distractors. In the second type, the correct answer is generated by replacing one noun phrase with another in the same coreference chain, and the distractors are composed by replacing the noun phrases with ones in the other coreference chain.

The difficulty of the reading comprehension questions is based on the reading level of the reading materials themselves. We assume that an examinee correctly answers a reading comprehension question because he or she could understand the whole of the story. The reading level estimation of a given document in recent years has increased noticeably. In this study, we adopt the measure of reading difficulty estimation developed by [6] to identify the difficulty of reading materials, as a difficulty measure for the reading comprehension questions.

The third question in Figure 2 is an example of an independent referential question, which assesses the concept of one entity involved in sentences. This question is answered correctly when examinees understand the reasonable semantics of the concept in the document. The fourth question in Figure 2 is an example of the overall referential question, which contains more than one concept that needs to be answered. This approach further examines the connected concepts of the given learning material.

3. PERSONALIZED QUIZ STRATEGY

In this section, the personalized quiz strategy based on automatic quiz generation is presented. This personalized quiz strategy aims to achieve the following two purposes: first, generating items depending on the proficiency level of an examinee, in order to provide an adaptive test; second, designing a suitable quiz in order to improve an examinee’s proficiency.

3.1 Proficiency Level Estimation

Let P represent an examinee's proficiency level. Proficiency level is categorized as vocabulary ability level l_v, grammar ability level l_g and reading comprehension ability level l_r, so that an examinee's proficiency level is represented as $P = \{ l_v, l_g, l_r \}$. The variables in this formula respectively represent each examinee's proficiency level, consisting of vocabulary ability, grammatical ability, and reading ability. For a given current proficiency level $P_t = \{l_{v,t}, l_{g,t}, l_{r,t}\}$ where t represents an examinee's proficiency level in time t.

To assign an examinee's proficiency level, an examinee's current proficiency score is calculated first. This score is transformed by the following formula:

$$Y_t = \sum_{i=1}^{n} u_i / n, \quad u_i = \begin{cases} 1, & \text{if an examinee correctly answered an item } i \\ 0, & \text{if an examinee incorrectly answered an item } i \end{cases} \quad (1)$$

where Y_t is the initial proficiency score in time t, I is a set of questions in an exam, i represents the i_{th} question in the exam, n represents the number of questions in the exam and u_i represents the responses of the learner in the exam. This formula represents the percentage of n items an examinee answers correctly.

We also consider an examinee's performance history and employ exponential moving average (EMA) [2] to combine it with the current initial proficiency score, transformed by the following formula: $S_t = \alpha \times Y_t + (1-\alpha) \times S_{t-1}$ (2) where S_t is the final proficiency score in time t after the combination with EMA, S_{t-1} is the past proficiency score in the time $t-1$ as history records, $\alpha=2/(m+1)$ is a constant represented as a weight, and m represents the length of the moving window. The expectation proficiency in each grade level is also measured: $E_l = \sum_{i=1}^{n} p_{i,l} / n$ (3) where $p_{i,l}$ represents the percentage of the proficiency level in which l examinees correctly answered the question i. This formula represents the average probability that the proficiency level l examinees correctly answered the test.

An examinee's proficiency level is assigned to the closest expected proficiency in grade level: $\hat{l}_t = \arg\min | S_t - E_l |, \quad l_t \in \{l_{v,t}, l_{g,t}, l_{r,t}\}$ (4) where $\widehat{l_t}$ represents an estimated examinee's proficiency level in one of the proficiency categories in time t, S_t is a learner's proficiency score in (2) and E_l is the expected proficiency score in (3).

3.2 Quiz Strategy

This section presents the quiz strategy. When given a learner's ability l, it is critical to determine how to best form a test from a series of questions which match their ability. In [1], the researchers selected history-based questions consisting of the recently used questions and correctly answered questions. In this study, a test is composed of not only fit questions (a question's level is equal to a learner's level) and history-based questions (a question's level is easier than a learner's level) but also challenging questions (a question's level is more difficult than a learner's level). Like Barla, et al. [1], we define probability values to assign questions in a test. Here, the percentage of history-based questions, fit questions and challenging questions are 20%, 60% and 20%, respectively. When fit questions are answered incorrectly, they are stored in the system. During the next iteration of the test, if there is any similar question based on the same concept, this question will be selected first. The goal of this design is to enhance learners' impression and improve their proficiency.

4. PERSONALIZED QUIZ STRATEGY

4.1 Experimental Design

The proposed methods are developed from the AutoQuiz project [8][9], which provides English language learners with automatic quiz generation. AutoQuiz is implemented on the IWiLL learning platform, which offers learners an online English reading and writing environment. The reading interface and the test interface are given in Figure 3. A total of 2,481 items, composed of vocabulary, grammar, and reading comprehension, were automatically generated based on 72 new stories as reading materials. The news articles were collected from several global and local online news websites: Time For Kids, Student Times, Voice of America, CNN, China Post Online and Yahoo! News.

The participants in this study were high school students in Taiwan, divided into two groups: a control group with general automatic quiz generation, and an experimental group with personalized automatic quiz generation. 33 students participated within the control group, while 123 students participated in the experimental group. 21 and 72 subjects in the control group and the experimental group respectively completed all phases of the research.

The experiment was held from July 1st to September 30th, 2011. During the experiment, the subjects were asked to participate in twelve activities, consisting of reading an article and then taking a test. Each test was composed of ten vocabulary questions, five grammar questions, and three reading comprehension questions. In addition, there was a pre-test and post-test for evaluating changes in learner proficiency, each with a similar degree of difficulty. The proficiency level in this study is defined from one to six, corresponding to the six semesters of Taiwanese senior high school. Finally, 30 subjects in the experimental group volunteered to fill out a questionnaire that elicited information concerning the examinee experience and the quality of the generated questions. Questions in the questionnaire were taken from [12]. A five-point Likert scale was employed.

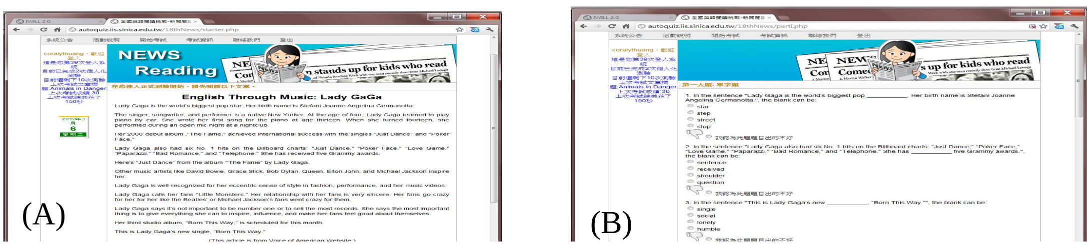


Figure 3. Snapshots of the system: (A) An example of a given reading materials from new online website; (B) An example of vocabulary items.

4.2 Experimental Results

The aim of the quiz strategy is to enhance student understanding of concepts they find unclear. We measured the rate at which students successfully corrected their mistakes on repeated concepts (denoted as the rectification rate) in the experimental group and control group, to determine the effect of generating items with repeated concepts. To make comparisons, the independent-samples t-test and the Mann-Whitney U test were both performed. Ideally, the distribution between the two groups is a normal distribution, and thereby uses a t-test. However, because of unequal sample sizes, the nonparametric method is complementary. The results suggest that the rectification rate in the experimental group was on average significantly higher than in the control group ($t=6.60$, $p<0.001$ in the independent-samples t-test and $Z=-5.97$, $p<0.001$ in the the Mann-Whitney U test). Moreover, the subjects in the experimental group ($M=0.54$, $SD=0.29$) were more than half as likely to correct unclear concepts and answer similar questions correctly. This indicates that a personalized quiz strategy would help the learners correct previous mistakes.

To further understand the influence of a personalized automatic quiz generation, the normalized score (normalized from zero to one) in the post-test between the experimental group and control group were calculated and compared in the parametric and nonparametric analysis. The results of an independent T-test (*p=0.80* in the pre-test and *p=0.46* in the post-test) and the Mann-Whitney U test (*p=0.99* in the pretest and *p=0.59* in the post-test) showed no significant effect on the post-test between the experimental group and the control group. However, the paired sample T-test and the Wilcoxon signed-rank test showed a significant effect of the pre-test and the post-test in the experimental group ($p<0.01$), while the performance of the control group had no statistically significant effect ($p>0.05$). This indicates that the personalized automatic quiz generation within the experimental group still effectively improves their own learning.

To study the performance in each difficulty level between the pretest and post-test, the number of correctly answered questions among the six difficulty levels in the pretest and the post-test were computed. The tests are comprised of 28 items among six difficulty levels (six, three, six, three, seven and three questions per respective level, corresponding to levels one through six). A Chi-Square test for homogeneity of proportions was conducted to analyze the proportion between the pre-test and post-test. Table 1 presents two contingency tables respectively in the control group and the second graders of the experimental group. The results of the experimental group ($\chi2(5)=16.24, p<0.01$) show the significant different proportions between the pre-test and post-test, while the control group ($\chi2(5)=7.46, p>0.05$) has a similar percentage among the six difficulty levels. This change reveals that the personalized test affects the ability of the students in the experimental group. To further investigate the difference in the experimental group, a posteriori comparison reveals that the number of correctly answered questions with level two and level six in the post-test were statistically higher than those in the pre-test, whereas the number of questions with level one and level four in the post-test were significantly lower than those in the pre-test. This suggests that the number questions with higher difficulty level that were correctly answered increased after the personalized quiz strategy.

Table 1. Contingency tables for the number of correctly answered questions per difficulty level in the pretest and post-test

Difficulty Level		1	2	3	4	5	6
The number of questions		6	3	6	3	7	3
Control group	Pretest	69 (23.8%)	27 (9.3%)	63 (21.7%)	36 (12.4%)	68 (23.4%)	27 (9.3%)
	Post-test	73 (21.3%)	50 (14.6%)	72 (21.0%)	33 (9.6%)	71 (20.7%)	44 (12.8%)
Experimental group	Pretest	**248 (24.8%)**	99 (9.9%)	209 (20.9%)	**129 (12.9%)**	206 (20.6%)	108 (10.8%)
	Post-test	234 (20.5%)	**147 (13.1%)**	253 (22.6%)	106 (9.5%)	236 (21.1%)	142 (12.7%)

Table 2. Questionnaire results

	Items	Mean	SD
1	The reading interface is easy to use.	3.89	0.99
2	The test interface is easy to use.	3.86	0.95
3	Taking the quiz has helped me to evaluate my strengths and weaknesses..	4.00	0.67
4	Taking the quiz has helped me to identify areas of knowledge that need improvement.	4.03	0.64
5	Taking the quiz is useful preparation for exams.	3.89	0.7
6a	I clearly understood the vocabulary questions on the quiz.	3.27	0.99
6b	I clearly understood the grammar questions on the quiz.	3.46	0.99
6c	I clearly understood the reading comprehension questions on the quiz.	3.38	0.95

In terms of evaluating the performance of the automatic question generation, six questions in the questionnaire concerning the subjects' perception were investigated. Table 2 displays the detailed questions and shows their mean score and standard deviation. From the results, the quality of the interface and the functionality of the generated questions have high agreement. Most subjects agreed that the adaptive question selection strategy could help them identify strengths and weaknesses, so that they could improve their skills and prepare well for exams. This data supports the performance of the proposed automatic question generation and represents the usefulness of the generated questions.

5. Conclusion

This paper presents a personalized automatic quiz generation model, which generates multiple-choice questions based on various difficulty levels and categories. This quiz generation technique is then paired with a quiz strategy to estimate an examinee’s ability and to select suitable questions. Compared to the control group, the results show that the experimental group corrected their mistakes more frequently, answered more difficult questions correctly, and showed significant improvement between the score of the pre-test and post-test. In addition, the questionnaire results suggest most subjects support the functionality and quality of the proposed personalized automatic quiz generation.

Acknowledgment

This work was partially supported by National Science Council, Taiwan, with Grant No. 100-2511-S-008-005-MY3 and NSC97-2221-E-001-014-MY3.

References

[1] Barla, M., Bielikova, M., Ezzeddinne, A. B., Kramar, T., Simko, M. and Vozar, O. (2010). On the impact of adaptive test question selection for learning efficiency. *Computer & Education*, 55(2), 846-857.

[2] Brown, R. G. (2004). *Smoothing, Forecasting and Prediction of Discrete Time Series*. Dover Publications: Dover Phoenix Ed edition.

[3] Chen, C. Y., Ko, M. H., Wu, T. W. and Chang, J. S. (2005). FAST：Free Assistant of Structural Tests. *In Proceedings of the ROCLING 2005*.

[4] Embertson S., and Resise, S. (2000). *Item response theory for psychologists.* New Jersey, USA: Lawrence Erlbaum.

[5] Hoshino, A. and Nakagawa, H. (2007). Sakumon: An assistance system for English cloze test. *In Society for Information Technology & Teacher Education International Conference*.

[6] Huang, Y. T., Chang, H. S., Sun Y. and Chen M. C. (2011). A Robust Estimation Scheme of Reading Difficulty for Second Language Learners. *In Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies*, 58-62.

[7] Levy, R. and Andrew G. (2006). Tregex and Tsurgeon: tools for querying and manipulating tree data structures. In Proceeding of 5th International Conference on Language Resources and Evaluation.

[8] Lin, Y. C., Sung, L. C. and Chen, M. C. (2007). An Automatic Multiple-Choice Question Generation Scheme for English Adjective Understanding. *In Proceedings of the 15th International Conference on Computers in Education*, 137-142.

[9] Lin, Y. T., Chen, M. C. Sun, Y. S. (2009). Automatic Text-Coherence Question Generation Based on Coreference Resolution. *In Proceedings of the 17th International Conference on Computers in Education*.

[10] Mitkov, R. and Ha, L. A. (2003). Computer-Aided Generation of Multiple-Choice Tests. *In Workshop on Building Educational Applications Using Natural Language Processing*.

[11] Raghunathan, K., Lee, H., Rangarajan, S., Chambers, N., Surdeanu, M., Jurafsky, D. and Manning, C. (2010). A Multi-Pass Sieve for Coreference Resolution. *In Procceding of EMNLP2010*.

[12] Wilson, K., Boyd, C., Chen, L. and Jamal, S. (2010). Improving student performance in a first-year geography course: Examining the importance of computer-assisted formative assessment. *Computers & Education*, 57 (2), 1493–1500.

[13] Yang, Y. C., Yang, C. F, Chang, C. M. and Chang, J. S. (2005). Computered-aid Reading Comprehension Automatic Quiz Generation. *In Proceedings of the ROCLING 2005*.

A Multiplayer Online Role-Playing Game for Incidental Vocabulary Learning

Ben-Gao Huang, Jie Chi Yang
Graduate Institute of Network Learning Technology, National Central University, Taiwan
hbg@cl.ncu.edu.tw, yang@cl.ncu.edu.tw

Abstract: Incidental vocabulary learning is regarded as one of the main sources of learner-centered vocabulary acquisition in authentic situations. In the past decades, extensive reading has been the main focus in incidental learning research. Recent studies have examined how information technology media can assist learners in acquiring vocabulary incidentally. More specifically, the affordances of MORPGs, such as vivid 3D simulation scenarios and players' interactions and communications, may be applied to construct an incidental language learning environment. This study is based on Krashen's Input Hypothesis to implement task-based learning in an MORPG that was suited for learners' background. An experiment was conducted for 26 sixth-grade students. Twelve target words were selected from the content text of the game to assess learners' vocabulary learning. The results from this study demonstrate that learners, especially the low academic achievers and less game experienced students, acquired the target vocabulary incidentally after being exposed to the MORPG language learning environment.

Keywords: digital game-based learning, incidental vocabulary learning, MORPGs

1. Introduction

Vocabulary ability plays a fundamental role for language learners. There are many ways for learning vocabulary. Among them, incidental vocabulary learning is the main source that learners can acquire vocabulary in relevant context and language examples at their own learning paces. Some research indicated that incidental vocabulary learning is an effective method for understanding vocabulary [1-4]. In order to achieve vocabulary learning, Krashen proposed Input Hypothesis suggesting that second language learners should have enough comprehensible input. Comprehensible input refers to learning content slightly higher than students' ability (i+1), which could help learners proceed incidental vocabulary learning [5]. For the past decades, reading has been the majority source for incidental vocabulary learning. Recently, some studies are using multimedia, CALL programs, and computer games to construct incidental learning environments [6-8]. One of the examples is the Multiplayer Online Role-Playing Games (MORPGs), which has gained much attention for its amazing features. MORPGs are claimed to be beneficial for language learning as they can provide authentic context, such as vivid 3D simulation scenarios, players' interaction and communication, and meaningful game tasks [9-10]. In such an environment, learners need to understand the narrative of game scripts and to communicate with other players. Therefore, they may learn vocabulary incidentally from the texts or conversations in MORPGs. Many studies have used MORPGs for helping learners to learn vocabulary in relaxed atmosphere [11-12]. The MORPGs are popular entertainment in new generation. Though MORPGs had the potential of benefitting vocabulary learning, and the game players were exposed to many conversations or texts incidentally, there were few studies conducted to investigate incidental vocabulary learning in MORPGs. To this end, this study aims to implement a task-based learning process in an MORPG that was suited to learners' background, with comprehensible input suggested by Krashen's Input Hypothesis, and evaluated the effectiveness of learners' incidental vocabulary learning in the MORPGs.

2. Related works

2.1 Incidental vocabulary learning

Incidental vocabulary learning occurs in learners' reading that does not provoke learners to notice the meaning of unknown words and has no specific purposes [13]. It is a learner-centered pedagogy that learners are exposed to the relevant reading context and many language examples at their own learning pace. Many studies indicate that incidental vocabulary learning is an effective method for understanding vocabulary [3, 4, 13, 14]. The situations in which incidental learning occurs are usually unlike the normal instruction with significant learning purposes. Therefore, the learning effectiveness is shown in a gradual progress. In general, incidental vocabulary learning should consider the pre-knowledge of learners. Reading materials that instructors provide for learners should fit learners' capacity [15]. Though the number of encounters may affect the words learned incidentally, the quality of the context may have a greater effect on gaining the meaning of words [16]. The Input Hypothesis indicated that second language learners could acquire vocabulary incidentally by reading proper materials. In order to construct comprehensible input contexts, the level of reading materials should be slightly higher than the learners' ability [2, 5, 15]. Huckin and Coady reviewed the incidental vocabulary learning research and argued that incidental learning was not entirely incidental [17]. In addition to the situations in relation to learners' background; instructors have to provide appropriate guidance, such as learning tasks, material presentation, and glossing method. Hulstijn, Hollander, and Greidanus used marginal glosses and the dictionary to help L2 advanced learners' reading. They found though both were available, marginal glosses were better for incidental learning, and the frequency of word occurrence had positive effectiveness [18]. Kost et al. investigated the effects of three gloss types, finding combination of text and pictures gloss was better than text gloss and pictures gloss [19]. Yoshii and Flaitz involved text and images into Web allowing learners to read English text online. They found that online learning material which was designed based on duel-code theory could help learners with incidental vocabulary learning [20]. Kuppens suggested that watching second language TV programs and playing related computer games had positive influence in incidental vocabulary acquisition [5].

Previous studies indicated that the context would affect learners' understand of word meaning during incidental learning. Providing glosses such as situational pictures, animations and other appropriate cues may assist learners to understand the word meaning. Incidental learning may occur during extensive reading, but the form of reading can be various. In order to make reading material and situations interesting, researchers have sought for other resources such as video, CALL programs, and games to enrich incidental learning context. For example, Lin investigated how news video in a CALL program could foster L2 comprehension and incidental learning and found that both low and high proficiency learners showed significant progress [8].

2.2 MORPGs for language learning

In recent years, MORPGs have become popular among young people. Many researchers have indicated that MORPGs could be applied to improve communication skills and were beneficial for learners’ language acquisition. The characteristics of MORPGs such as vivid 3D images can be used for situational simulation which allows learners to immerse in virtual reality situations. In those situations, learners must try to understand all the scenarios and narratives of MORPGs. MORPGs provide beneficial environments for second language acquisition that promote learners' motivation to learn. Learners are satisfied with the vocabulary learning effectiveness in the situation [10, 21, 22]. The main characteristics of MORPGs are communication, high-quality 3D graphic interfaces, and customizable character avatars that may support language acquisition [22]. Designing vocabulary learning tools should provide application opportunities for learners in different contexts to access target vocabulary [23]. MORPGs provide a virtual reality situation and opportunities for player interaction in a vocabulary learning environment. This study addresses the issue of how to help digitally native generation to learn language when they are immersed in an MORPG for incidental vocabulary learning.

Huckin and Coady reviewed the empirical research of incidental vocabulary learning and indicated there were some unsolved issues, such as the actual mechanism, the kinds of texts, and input modification [17]. Studies that applied MORPGs to vocabulary learning showed positive effectiveness. The vivid 3D simulation scenarios, players' interaction and communication, and game tasks can provide a vigorous learning mechanism and display learning material in edutainment. MORPGs thus may have the potential to construct an incidental vocabulary learning environment, but relative research is rare.

In order to construct an incidental vocabulary learning environment in an MORPG, this study was based on the Input Hypothesis to design game scripts adapted from a Chinese legend, with consideration of learners’ English proficiency and culture background. An experiment was conducted to evaluate the effects of incidental vocabulary learning in the MORPG. The aims of this study are: 1. the effects of implementing incidental vocabulary learning in an MORPG environment, and 2. the influences of the learners’ background (English academic performance, gender, and game experience) and the texts of narratives (word occurrence frequency) on incidental vocabulary learning in an MORPG environment.

3. Methods

3.1 Participants

The participants were 26 sixth-grade elementary school students, 13 boys and 13 girls participating in this study. All the participants had 3.5 years English lesson experience for 80-minute English classes per week, and had the basic computer skills to operate a digital game. The 26 participants were from the same class.

3.2 Instruments

3.2.1 Tests

The tests used in this study included an English vocabulary pre-tests, a post-test, a delay post-test, and a game experience questionnaire. English words can be divided into content words and function words. Function words are grammar words, while content words primarily express lexical meanings, including nouns, verbs, and adjectives that convey the real information of a statement [24]. This study selected twelve content words as the target words from the MORPG game script. They included four nouns (weather, tangerine, pineapple, and monster), four verbs (taste, find, buy, and kill), and four adjectives (cloudy, sunny, snowy, and east). These target words were never listed in participants' English textbooks they had used. In order to reduce the pre-test effect, some non-target words were added to the pre-test [13, 19]. The post-test and delay post-test have twelve questions, assessing only the target word. The test used Paribakht and Bingham's Vocabulary Knowledge Scale (VKS) five scales to measure participants' incidental vocabulary acquisition [25]. It can be used with any set of words to evaluate the depth of vocabulary knowledge [24]. The five scales from score of 1 to 5 were "I have never seen this word.", "I have seen this word before, but I don't know what it means.", "I have seen this word before, and I think it means _.", "I know this word. It means _.", and "I can use this word in a sentence: _". The questionnaire was used to investigate the game experiences of participants.

3.2.2 The MORPG

Because the participants of this study were children, the game, Rainbow Bubble (see Figure 1), adopted for this study was a lively and non-violent 3D MORPG. The character design of avatars and the narratives of game scripts are the main elements in MORPG. It is expected that the design of the learning content in narratives and tasks can enhance or arouse learners' intrinsic motivation [10,22]. In Rainbow Bubble, learners can create their own avatars to represent themselves and then to interact with Non-Playing Characters (NPCs) who are set by game designers and will assign tasks or provide problem-solving support for learners in the MORPG.

Figure 1: Screenshot from the MORPG

The pedagogical strategy adopted in MORPG was task-based. Learning content was present in game narratives and tasks. When learners finish a task, they will acquire both experience value and money for upgrading and buying equipment, which can strengthen their battle power and are beneficial to solve the learning tasks. All the game scripts and scenarios were designed in relation to participant English proficiency and their culture background which were slightly higher than the student's ability [2, 5]. In the game, a year monster of a Chinese legend would appear in Chinese New Year. To complete the main task defeat the monster players must rely on special magic which can be obtained from several magic books. In order to collect all the magic books, students must finish both series tasks and tests. During the game, the MORPG provided searching tools and word cards to support language learning, and students can help each other either by talking in classroom or discussing online. The searching tool was like a dictionary listing the vocabulary. The word cards would pop-up incidentally during the tasks, hoping to make learners have deeper impression of what they had learned. The magical book contained target words and example sentences for students to look up and review.

3.3 Procedure

The procedure of this study consists of three steps. In the first step, participants completed the pre-test and a game experience questionnaire in approximately 30 minutes. To reduce the effect of the pre-test, the pre-test was carried out one week before playing the MORPG [13, 19]. The second step was the participants playing the designed MORPG game for 40 minutes. Students did not know what vocabulary they were learning, but they were encouraged to finish each task at their best. Finally, the immediate post-test was completed in approximately 20 minutes, and the delay post-test, same as post-test, was carried out twelve weeks later.

3.4 Data analysis

The study was a quasi-experimental design which used a single experimental group. The data were collected from the pre-test, immediate post-test, delay post-test, questionnaire and students' academic score in English course of the first semester of sixth grade. In addition to descriptive statistics analysis, the study used the t-test to evaluate the differences of the scores between the pre-test, post-test, and delay post-test.

4. Results and discussions

4.1 Incident vocabulary learning effectiveness in the MORPG

The results show a significant difference between the scores of pre-test and post-test (t=-2.393, p=0.025*). The mean score was progressed 3.31 (see Table 1). This result indicates that the MORPG learning environment has a positive impact on incidental vocabulary learning. Comparing different types of target words, it is found that acquisition of nouns has improved significantly in post-test (p=0.004**) and is higher than the verb and adjective acquisition. This may be because nouns are more concrete [13]. We further divided the participants into high and low achievers by their English academic scores. The results reveal that both groups had improved in post-test. Moreover, the low achievers progressed more than high achievers in the post-test and there was a significant difference between low achievers' pre-test and post-test results. In terms of types of words, it is found that for high achievers all types of words show no significant difference between the pre-test and post-test, but for low achievers there are significant differences in both nouns and verbs (see Tables 2 and 3). These results may be due to that low achievers rely more on the visual aids while learning the words. For example, learners can learn nouns more easily with the support of 3D images and understand the meaning of the verbs more clearly as they are associated with the instructions of game tasks. Previous studies indicated that the text-based learning materials with visual support could help incidental vocabulary learning [19,20]. In an MORPG, the narratives, tasks, and scenes could just be used as learning material to construct an incidental vocabulary learning environment.

In order to examine the retaining effect, the study implemented a delay post-test after 12 weeks. There is a significant difference between the scores of the pre-test and delay post-test (t=-4.235, p=.000**); moreover, both high (t=-3.187, p=.008*) and low achievers (t=-3.207, p=.008*) had significant improvement in the delay post-test. The gain score was 4.038, which was higher than the difference between pre-test and post-test. This might be due to the fact that learners had participated in other learning activities that probably enhanced their performance within the 12 weeks.

Table 1: Paired samples t-test analysis for all students' VKS

Vocabulary	Pre-Test		Post Test			
	Mean	SD	Mean	SD	t	p
Total	31.69	16.292	35.00	15.895	-2.393*	.025
Nouns	11.19	6.487	12.81	6.020	-3.176**	.004
Verbs	10.12	6.179	10.92	5.642	-1.382	.179
Adjectives	10.38	4.309	11.27	5.096	-1.381	.179

Table 2: Paired samples t-test analysis for high achievers' VKS

Vocabulary	Pre-Test		Post Test			
	Mean	SD	Mean	SD	t	p
Total	44.54	11.759	46.92	11.019	-.985	.344
Nouns	16.31	4.662	17.77	2.743	-1.814	.095
Verbs	15.08	4.663	15.08	4.627	.000	1.000
Adjectives	13.15	3.648	14.08	4.924	-.772	.455

Table 3: Paired samples t-test analysis for low achievers' VKS

Vocabulary	Pre-Test		Post Test			
	Mean	SD	Mean	SD	t	p
Total	18.85	7.559	23.08	9.844	-3.029**	.010
Nouns	6.08	3.040	7.85	3.826	-2.714*	.019
Verbs	5.15	2.115	6.77	2.743	-3.313**	.006
Adjectives	7.62	2.959	8.46	3.573	-1.599	.136

4.2 Influences of learners' background and word occurrence frequency

The study further explored the influences of learner gender and game experience on incidental vocabulary learning in an MORPG environment. The results show that there was a significant difference between pre-test and post-test for female learners' performance (t=-2.771, p=.017*) (see Table 4). The gain score of female learners was 5, which was higher than male learners' 1.615. As regards to the gaming experience, we discussed this from two aspects. One was the number of game types that learners had played; the other was the hours the learners played per week. The results show that learners with low game experience (either in game type or play time) showed a significant difference between their pre-test and post-test performance (see Table 4). Besides, their gain scores were higher than those of the high game experience learners. It may be because low game experience learners noticed more details of scenarios or narratives of the MORPG, while those with more game experience might find the task solution from other cues and ignore the details. It can also be noted that the female learners had less game experience than male learners in term of either game types (t=2.729, p=.012*) or playing hours (t=2.702, p=.012*). Further studies on the relation between game experience and gender should be considered with more participants involved to analyze the variables.

The frequency of word occurrence may influence the effectiveness of incidental vocabulary learning [18]. However, in this study, there is no significant relation found between the frequency of word occurrence and the gain score in the MORPG. Such a phenomenon may be associated with the context where words occur [16] and the short game time.

Table 4: The gain score of VKS between pre-test and post-test

Pre-test – Post test	Mean	SD	t	p
Male	-1.615	7.411	-.786	.447
Female	-5.000	6.506	-2.771*	.017
Game types >=3	-1.091	8.006	-.452	.661
Game types <3	-4.933	6.017	-3.175**	.007
Playing hours/week > 9	-1.846	8.133	-.818	.429
Playing hours/week <=9	-4.769	5.718	-3.007*	.011

5. Conclusion

The results of this study showed that there was a significant difference between pre-test and post-test for learners' vocabulary abilities, which reveals that learners' vocabulary abilities have improved after they used the MORPG language learning environment. Learners were not only immersed in high-quality 3D virtual reality scenes and enjoyed in interesting game tasks, but also acquired vocabulary incidentally. These findings are in line with those of past studies, which showed that MORPGs can be applied to construct a language learning environment [10, 21, 22]. The target words were the investigation indicator for determining whether vocabulary could be learned when they were embedded in the narratives of the game scripts. These learning contents were based on Krashen's Input Hypothesis to implement a task-based learning. The task-based narrative MORPG has positive effects on incidental vocabulary learning. Furthermore, no significant correlation was found between gain scores and students' academic scores. To explore the performance of learners with different English proficiency, it is found that the low achievers significantly improved in the post-test, while the high achievers did not. It is necessary to provide more mechanisms for advanced learners in future design of the MORPG language learning environment.

This study also examined the effects of other factors for incidental vocabulary learning in MORPGs, such as frequency of word occurrence, different types of words, gender, and game experience for incidental vocabulary learning in MORPG environment. There was no significant difference between the frequency of word occurrence and gain score. Incidental language learning does not have explicit learning purposes as normal instruction does and thus the learning effectiveness is gradual [15]. There is a need to provide more narration in the game scripts and allow learners to play MORPGs longer in future study. In the experiment, the post-test score of nouns was significantly better than the pre-test, which might be due to the fact that the nouns used in the game were more concrete and were the basic elements of the message [13].

In this study participants' game experience varied, for instance, female learners had less game experience, in term of the playing duration and the types of game they had played. The VKS results of female learners and less game experienced learners showed significant progress, while male learners and learners with more game experience did not. The interaction between different genders and difference game experiences need further analysis if more participants are involved in the future.

As a preliminary attempt, this study integrated learning content into an MORPG game, using a familiar Chinese legend as the background story. There are still some issues of MORPGs design to be considered in the future studies, such as how to present learning material, to encourage interactions between learners, to consider other game elements, and what the role of teachers in MORPGs is. Considering the effect of incidental vocabulary learning is gradual, in the next stage, we should increase the time and participants to provide additional evidence to verify the results in this study, and thus explore more aspects of incidental vocabulary learning in MORPGs.

Acknowledgements

This study was partially supported by grant (NSC 100-2628-S-008-002-MY3) from the National Science Council of Taiwan.

References

[1] Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7*, 541–551.

[2] Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis, *Modern Language Journal 73*, 440-464.

[3] Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

[4] Rashidi, N., & Adivi, A. G. (2010). Incidental Vocabulary Learning Through Comprehension-Focused Reading of Short Stories, *Journal of English Language Teaching and Learning, 53*(217), 111-129.

[5] Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.

[6] Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. Learning, *Media and Technology, 35* (1), 65–85.

[7] Dalton, B., & L. Grisham, D. (2011). eVoc Strategies: 10 Ways to Use Technology to Build Vocabulary. *The Reading Teacher, 64*(5), 306–317.

[8] Lin, L. F. (2010). A video-based CALL program for proficient and less-proficient L2 learners' comprehension ability, incidental vocabulary acquisition. *Educational Media International, 47*(3), 199-216.

[9] Anderson, T.A.F., Reynolds, B.L., Yeh, X-P., & Huang, G-Z. (2008). Video games in the English as a foreign language classroom. In M. Eisenberg, Kinshuk, M. Chang, R. McGreal (Eds.), *Proceedings of second IEEE international conference on digital games and intelligent toys based education* (pp. 188-192). Banff, Canada: IEEE.

[10] Rankin, Y., Gold, R., & Gooch, B. (2006). 3D Role-playing Games as Language Learning Tools. *In Conference Proceedings of EUROGRAPHICS Education Program 2006. Vol. 25*. Vienna, Austria.

[11] Alemi, M. (2010). Educational Games as a Vehicle to Teaching Vocabulary. *Modern Journal of Applied Linguistics, vol2*, No. 6, 425-438.

[12] Uberman, A. (1998). The use of games for vocabulary presentation and revision. *Forum, 36*(1), 20.

[13] Swanborn, M.S.L., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*, 261-285.

[14] Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal, 11*, 9-29.

[15] Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 225-237). New York: Cambridge University press.

[16] Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*, 232-245.

[17] Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies of Second Language Acquisition 21*, 181-193.

[18] Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80* (3), 327-339.

[19] Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32* (1), 89-113.

[20] Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of picture and annotation types. *CALICO Journal, 20*(1), 33-58.

[21] Suh, S., Kim, S.W., & Kim, N.J. (2010). Effectiveness of MORPGS-based instruction in elementary English education in Korea. *Journal of Computer Assisted Learning, 26*(5), 370-378.

[22] Peterson, M. (2011). Massively multiplayer online role-playing games as arenas for second language learning. *Computer Assisted Language Learning, 23*, 429-439.

[23] Wood, J. (2001). Can software support children's vocabulary development? *Language Learning & Technology, 5*, 166-201.

[24] Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press. Edited work.

[25] Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal, 11*, 9-29.

Situated design of computer assisted language learning (CALL) to support an endangered language

Virginia WESTWOOD
Murdoch University, Australia
admin@proteatextware.com.au

Abstract: This paper describes two unique partnerships between two Australian Aboriginal communities and a commercial computer assisted language learning (CALL) development- publishing company. The partnerships are conducting community–based projects of CALL development to support the continuation of the communities' languages. The project aims to produce a new model of situated design of CALL which will be available as a template for other communities who are interested in language continuation. The researcher is working at the cultural interface between Western colonising academia and Indigenous knowledge systems and lifeways. The CALL shell and editing template will be sustainable only with the support of a linguist and training in each community.

Keywords: Endangered languages, CALL, situated design, Australian Aboriginal languages

Introduction

The invasion and subsequent colonisation of Australia by European settlers cost the lives of many of the estimated 250 (Walsh, 1991) original Australian Aboriginal languages. Earliest contact with European settlers occurred in the eastern states, while contact in the north-west of Australia first occurred in the 1830s (Vigilante, 2001), and some Aboriginal clans only came out of the desert as late as the 1960s (Pindan, 2012). This project is located in the west Kimberley region of Australia (see Fig. 1.) and is concerned with the design of a computer assisted language learning (CALL) shell program through collaboration of two language communities and a successful CALL design-development-publishing company. The community languages are Nyikina, a non-Pama-Nyungan River language and Walmajarri, a Pama-Nyungan Desert language, thus representing both major (linguistic) families of Australian Aboriginal languages. The company (Protea Textware Pty Ltd) is sponsoring a PhD student (the author) who has 20 years of CALL design and development experience. The project was initiated when the Nyikina community approached the company in 2009 for a CALL resource to support their language continuation efforts. The project will deliver at least one CALL program for each language which the communities will own, a shell program for further CALL programs and an editing template for the content.

The intent is to reuse documentation materials created by independent linguists working with Aboriginal communities, including existing digitised dictionaries and records as well as grammars and orthographies and Protea Textware's language teaching approach. Previous attempts to draw on existing materials, such as digitised dictionaries, have been unsuccessful (Yang & Rau, 2005). The CALL program will have a phonemic module, an interactive picture dictionary module, communicative whole language activities based on longer texts and ideas for authentic (real-life) learning activities plus cultural information.

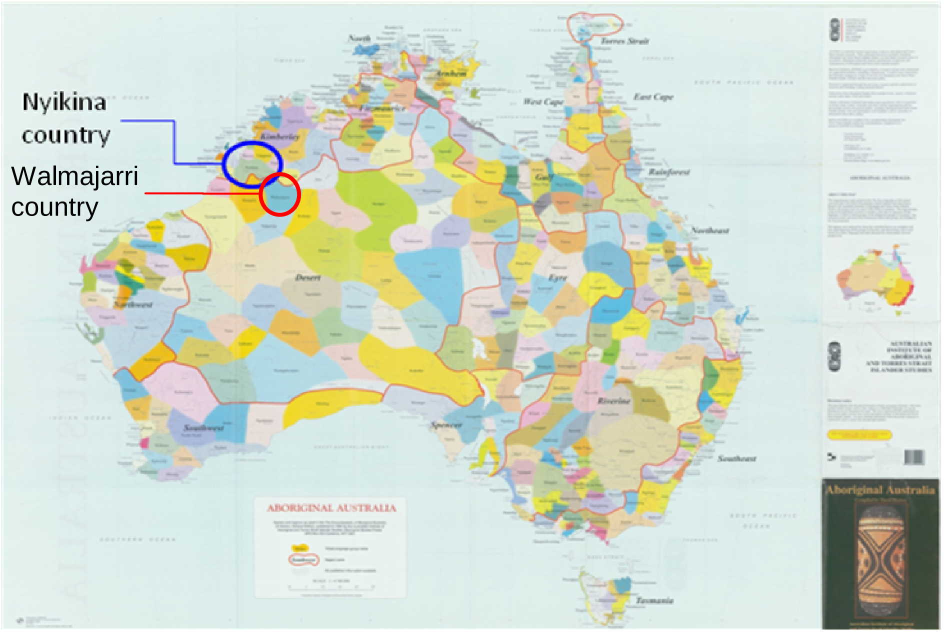


Figure 1. Nyikina Country and Walmajarri Country in Australia (AIATSIS, 2008)

1. Theory

1.1.1 Cultural Interface

The researcher is working at the cultural interface (Durie, 2005) between Western colonising academia and Indigenous knowledge systems and lifeways. Tuhiwai Smith (1999) notes that the research activity is transformed when Indigenous people become the researchers. In this project, this is reflected in the methodological processes, the outcomes, feedback of information to the communities and ownership of the outcomes.

The project was initially planned to work in the Nyikina community with the Nyikina language as a prototype for the template. However, contact with interested persons from a second community led to a second Aboriginal language (Walmajarri) being included in the project. Formal procedures such as the Intellectual Property (IP) agreement, permissions and invitations were already in place before work actually started on country for Nyikina language. Meetings were organized with Walmajarri community members with language and project experience and skills on country at Djilimbardi on the edge of the Great Sandy Desert in north-west Australia to negotiate formal permissions and agreements.

1.1.2 Situated Design

The concept of situated design in the context of endangered languages is here proposed to mean taking place on country with the community fundamentally involved in the design. Usually, situated design describes user-centred design (Wilson, 1995) meaning design by designers taking into consideration the needs and proficiencies of the end-user. I propose that situated design be taken to mean that the end-users are an integral part of all stages of the design (Cárdenas-Claros & Gruba, 2009). Previous attempts at developing CALL programs for endangered languages have not been designed on country in partnership with communities. They have instead been designed by designers, teachers or programmers in academia or commercial offices with varying success (Nathan, 2006; Ward, 2001).

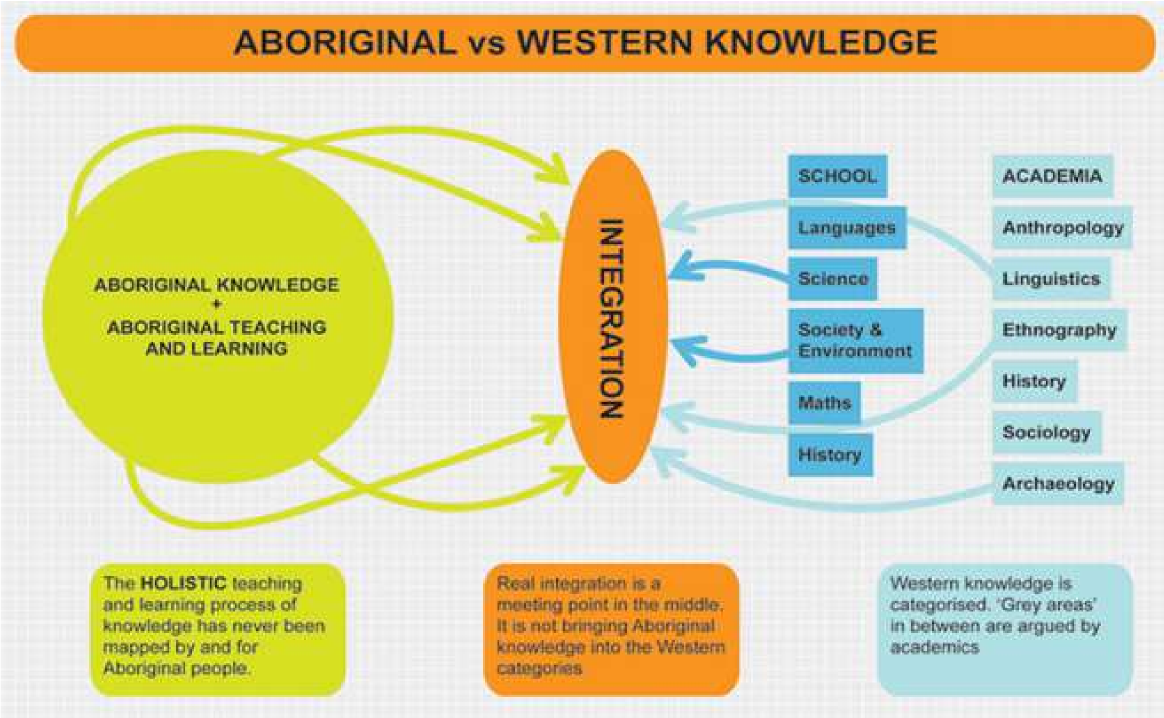


Figure 2. The cultural interface (Bedford & Casson, 2010)

1.1.3 Actor-Network Theory (ANT)

Actor-Network Theory (Latour, 2005) proposes that country, culture and language, as well as speakers, designers, technicians and programmers are all 'actors' in the design process. Without all of them, the process will fail.

2. Methodology

The three theories are actualised as they intertwine and reinforce each other through the design and development of the program, carried out entirely on country, and led by community protocols and traditions. Working with two different languages simultaneously, more rigorous evaluation of the template can be undertaken than was at first proposed.

1.2 Design-Based Research

The research approach in this study is that of design-based research (DBR) (as described by Collins et al (2004) as design research). In the context of instructional design, DBR is described as extending instructional designers to create practical and transferable design principles (Wang & Hannafin, 2005). Design-based research starts with a problem and then a theoretical solution is proposed, the principles for design formulated, the design implemented and evaluated (in a process similar to Critical Participatory Action Research (Kemmis, 2008)) and this cycle is repeated until a satisfactory solution is obtained.

1.2.1 Phase I

The **problem** of appropriate design for CALL for an Indigenous language was explored both from an instructional design point of view and from the Indigenous perspective, from the literature and from consultation. Draft design principles were established based on Aboriginal relationality, situated design, ANT and Protea Textware's philosophy of CALL design (Westwood & Kaufmann, 1999).

1.2.2 Phase II

The **preliminary solution** proposed an organized, structured series of meetings on a planned timeline. After consultation with the communities it was decided to follow an Aboriginal way of making decisions by all interested persons in the community meeting somewhere on country for a few days to talk, listen, sing and dance about the preliminary design issues. A smaller design group was set up for each language for ongoing reference, feedback and redesign. The solution also called for a CALL program that could be reused for different endangered languages, and so the content should be separated from the programming. Thus the CALL program is designed as a shell which reads all content data from underlying xml files, with a separate editing interface for the xml files.

1.2.3 Phase III

The **design** followed the draft principles with a number of aspects running concurrently: (a) create a program to read in existing documentary files from Lexique Pro (dictionary software – http://www.sil.org) to output in xml files; (b) contact, inform and seek support from senior community leaders; (c) establish trusting relationships with communities; (d) arrange intensive workshops; (e) design a CALL prototype shell as a starting point for discussion; (f) form smaller design groups for ongoing participation; and (g) design an editing interface for the xml files. The content (linguistic and contextual) and protocols desired by the community are decided at the intensive workshops, while redesign is ongoing with the smaller design group, based on evaluative feedback. The author will create structured exercises and activities with the content using Protea Textware models.

In the **implementation** stage, the program to read Lexique Pro files was created in Adobe Flex by the author. It outputs xml files that the CALL shell can read and display as an interactive program module. This has a simple push button interface. The author then wrote the remainder of the CALL shell program in Adobe Flex (AIR project) based on Protea Textware technologies and using the refined design principles arising from consultation with the communities as the design/program progressed. Technology access in the community is poor, leading to the decision to create a desktop-based program which would be accessible by all. Consultation determined that the program should be suitable for all ages and literacy levels. Implementation of the program will take place in after-school language classes (voluntary), libraries, resource centres, on home computers and on the project laptops taken to meetings with community.

Evaluation will draw on recorded open yarning sessions with users of the software and the consultant linguists and TechSmith Morae screen recording software. This will be collated and summarised before each meeting of the respective design groups.

1.2.4 Phase IV

Redesign and refinement of design principles will take place at each design group meeting.

3. Outcomes

The project will result in the formulation of a process for designing and developing a CALL program for endangered Australian Aboriginal languages and a construct which can be used as a template for other communities interested in language continuation. The xml editing interface will be developed in tandem with the CALL program and completed after the CALL program is finished.

To produce further CALL programs from the templates, training in use of the template will be required along with a linguist in the community.

Acknowledgements

I thank all the people of the Nyikina and Walmajarri communities for inviting me to their country and engaging with this project. In particular I would like to acknowledge the tremendous support of Dr Anne Poelina, Mr R. Craig Pindan and Ms Colleen Hattersley.

References

AIATSIS (Cartographer). (2008). Aboriginal Australia.

Bedford, P., & Casson, S. (2010). Conflicting knowledges: Barriers to language continuation in the Kimberley. *Australian Journal of Indigenous Education, 39S*, 76-86.

Cárdenas-Claros, M. S., & Gruba, P. A. (2009). Bridging CALL & HCI: Input from participatory design. *CALICO Journal, 27*(3), 576-591.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13*(1), 15-42.

Durie, M. (2005). Indigenous Knowledge Within a Global Knowledge System. *High Educ Policy, 18*(3), 301-312.

Kemmis, S. (2008). Critical theory and participatory action research. In P. Reason & H. Bradbury (Eds.), *The SAGE Handbook of Action Research: Participative Inquiry and Practice* (2nd ed.). London, UK: SAGE Publications Ltd.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory* Oxford: Oxford University Press.

Nathan, D. (2006). A talking dictionary of Paakantyi NSW. In L. E. Dyson, M. Hendrix & S. Grant (Eds.), *Information Technology and Indigenous People* (pp. 200-204). Hershey PA: Idea Group Inc.

Pindan, R. C. (2012). [Personal Communication].

Tuhiwai Smith, L. (1999). *Decolonizing methodologies: Research and Indigenous peoples*. Otago, NZ: Zed Books Ltd.

Vigilante, T. (2001). Analysis of Explorers' Records of Aboriginal Landscape Burning in the Kimberley Region of Western Australia. *Australian Geographical Studies, 39*(2), 135-155. doi: 10.1111/1467-8470.00136

Walsh, M. (1991). Overview of indigenous languages of Australia. In S. Romaine (Ed.), *Language in Australia*: Cambridge University Press.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development 53*(4), 5-23.

Ward, M. (2001). *A template for CALL programs for endangered languages.* Masters Dublin City University. Retrieved from http://www.computing.dcu.ie/~mward/mthesis.html

Westwood, V., & Kaufmann, H. (1999). Making a multimedia title. In S. Hawthorne & R. Klein (Eds.), *Cyberfeminism: Connectivity, critique & creativity* (pp. 352-367). North Melbourne, Vic. Australia: Spinifex Press.

Wilson, B. G. (1995). Situated instructional design: Blurring the distinctions between theory and practice, design and implementation, curriculum and instruction *Proceedings of the 1995 Annual National Convention of the Association for Educational Communications and Technology (AECT).*

Yang, M.-C., & Rau, D. V. (2005). An integrated framework for archiving, processing and developing learning materials for an endangered aboriginal language in Taiwan *Proceedings of the ALR-2005, Jeju Island, Korea* (pp. 32-39).

A Social and Cultural Analysis of Computer-mediated Communication Dialog in Asian Contexts

Pin-hsiang Natalie WU & Shih-min LIU
Applied Foreign Languages Department, Chien-kuo Technology University, Taiwan
Department of Education, National Kaohsiung Normal University, Taiwan
natalie@ctu.edu.tw

Abstract: The present study analyzes how social and cultural contexts influence the CMC interactions in a cross-cultural English collaborative learning project for university participants from Taiwan, Japan, and Korea. The main purpose of the study is to understand how social and cultural differences, manipulated by one's social and cultural context and interest toward international affair, influence the CMC dialogues among students of Asia, and how attitudes differ from one cultural group to another. The web platform communication tool used in this study was an online classroom established by Nicenet organization (www.nicenet.org). The project employed a qualitative research methodology using NVivo 9 digital software to analyze social and cultural aspects of the CMC dialogues. The results show that intergroup climate does play an important role in cross-cultural asynchronous CMC projects. Unfamiliarity among participants can present a constant influence on the flow of interaction, and feedback giving, which decreases motivation for participants from different cultural groups.

Keywords: Computer-mediated communication, cross-cultural, social context, Asia, second language learning

Introduction

The widespread use of Computer-Mediated Communication (CMC) to support foreign language learning in higher education is now very common around the world. The process of incorporating CMC into English learning is complicated because of variables associated with Second Language Learning (SL). Willingness to Communicate (WTC), Intercultural Communication Competence (ICC), and Computer-mediated Communication Competence (CMCC) will all influence the learning process and lead to differences in learning achievement. MacIntyre, Clemént, Dörnyei, Kimberly & Noels (1998), used a complex heuristic model of variables influencing L2 willingness to communicate and concluded that social context is the principal factor governing willingness to communicate.

Motivation has been recognized as a central factor of intercultural communication competence in communicative interactions (Kupka et al., 2009). Likewise, in the analysis of Yashima (2002), the participants' attitudes toward the International Community, or International Posture (IP), measuring attitudes toward Interest in International Activities, Intercultural Friendship, and Interest in Culture, was considered influential on their motivation when measuring their willingness to communicate (Wu & Kawamura 2012). Wu & Kawamura found that international posture had a direct influence on motivation regarding cross-cultural, online communications. In Vygotsky (1978), Fernandez, Wegerif, Mercer, and Rojas-Drummond (2011), Paulus (2007), and Ocker & Yaverbaum (1999), reciprocal and responsive collaboration, built on a good intergroup climate, is likely to increase peer interaction in the CMC process. Among the eight possible factors associated with CMC proposed by Andrew Tolmie and James Boyle (2000), the factors Size of Group and Knowledge of Other Participants clearly indicated that familiarity and acquaintance between participants increases the likelihood of success of online interaction.

This study analyzed how social and cultural contexts influence CMC interactions in a cross-cultural English collaborative learning project for participants from Asia. The main purpose of the study was to understand how social and cultural differences, manipulated by one's social and cultural context, and particularly in this study by intercultural posture, manipulated ongoing of CMC dialogs among students in Asia, and how their attitudes differed among three cultural groups. Intergroup attitudes influencing CMC were examined for: 1) level of social interaction displayed, 2) frequency of message;, and 3) message function.

Sixty-four university-level participants from Taiwan (26 students), Japan (29 students) and South Korea (9 students) were recruited for this study. The web platform communication tool used in this study is an online classroom established by Nicenet organization (www.nicenet.org), a formal Internet service that provides asynchronous computer-mediated communication. This web-learning platform is nonetheless user friendly and instructionally transparent for asynchronous CMC. Asynchronous CMC was considered a better choice compared with synchronous communication in the current study because asynchronous interaction allows more time and greater opportunity for one to reflect on one's own ideas, as well as on comments made by others (Paulus, 2007; Chou, Chen, and Hsieh 2009).

The project employed a qualitative research methodology using NVivo 9 Digital Software to analyze social and cultural aspects of the CMC dialogs. Different levels of social interactions displayed in the messages and message functions were coded and analyzed in order to analyze the relationships among all possible factors influencing the CMC interaction. Because participants from different cultural groups did not know each other prior this project, expressing social etiquette was considered environmentally friendly and a way of reinforcing social interaction that would contribute the success of the CMC dialogs. Other than that, introducing oneself to all participants was considered an even higher level of environmental friendly act. The following examples contributed by a Korean participant best portray such an attitude:

> Hello! My name is Yonghyun Kwon. My English name is Kate. You can easily [editor's note: simply] call me Kate! I'm [a] Korean student. Now I'm studying hard for English in U.S.A. I don't speak English very well. But I'll do my best. We are [editor's note: can] study together!
>
> Hi, My name is Yu Gyeong Kim. Call me Runa. I'm [a] Korean student. Now I'm studying in America. Difficult, but it's funny.

The above welcoming speech was not required by the teacher and therefore was considered a strong motivation to integrate others into the cross-cultural experience. In the process of analyzing Nicenet comments, expressions regarding social interactions were coded and then compared among participants of different countries. However, how participants socially interacted was also compared to the time and length of the messages they wrote and their responses to others.

Based on the frequency of comment writing and feedback giving, Taiwanese participants appeared to have the strongest motivation compared with learners in the other two countries, as shown in Figure 1. But the frequency of writing comments and replies does not completely demonstrate the intergroup climate in detail. Based on coding for frequency of social etiquette or motivation, almost all Asian participants included a certain level of social interaction. However, participants from different countries varied. Japanese participants usually started their comments with a short introduction that gave information about themselves, but initiated no further social interactions. For example:

> Hi! I'm Hyodo Toshiaki, a student of Kwansei Gakuin University in Japan and I'm a first-grade student. Nice to meet you all.

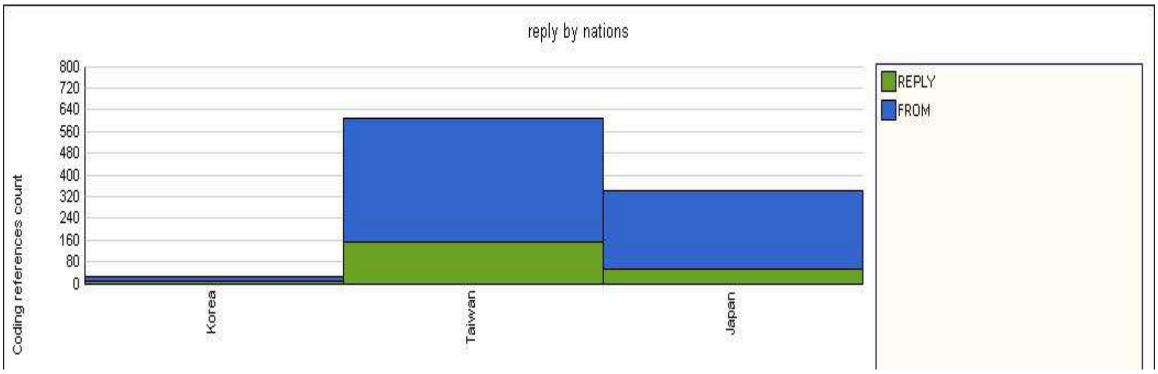


Figure 1: Frequency comments and replies from participants in the three countries

Korean students generally not only showed a stronger motivation for friendship but also shared English learning experiences. For example:

> Hi, My name is Yu Gyeong Kim. Call me Runa. I'm Korean student. Now I'm studying in America. Difficult, but it's funny … I need to study English hard.

Taiwanese participants were mostly task-focused and usually wrote what they thought was the "answer" to the questions without self-introduction or greetings unless they were making a reply to a certain person online.

These examples illustrate the learner's attitudes in the different cultural groups in this study. Taiwanese students tended to focus only on tasks whereas Japanese and Korean learners saw computer-mediated communication as a venue where a certain type of "social encounter" could happen. Among them, Japanese learners expressed social etiquette in the online CMC but not a desire for social interaction, whereas Korean participants wanted to build friendly relationships among the EFL learners. The social etiquette displayed in the Nicenet computer-mediated communication for different cultural groups is presented in Figure 2. Even though the social etiquette of Taiwanese students had the highest frequency, as seen in Figure 2, most of the instances were social etiquette appearing only in feedback given to specific people.

A closer examination of the way participants gave feedback provides a higher level of understanding of the attitudes toward cross-cultural interaction for these Asian participants in this study. Even though Japanese participants always introduced themselves in their first comments, which might be taken as a desire for acquaintance, most of the replies they made were to people in their own cultural group and were not intended as cross-cultural communication. Taiwanese participants expressed a very similar attitude, giving feedback mostly to other participants in Taiwan. For Korean participants, even though the number of messages was small, feedback was also given to other Korean learners instead of Japanese and Taiwanese participants. Social interactions among the three countries were shown in the Figure 2.

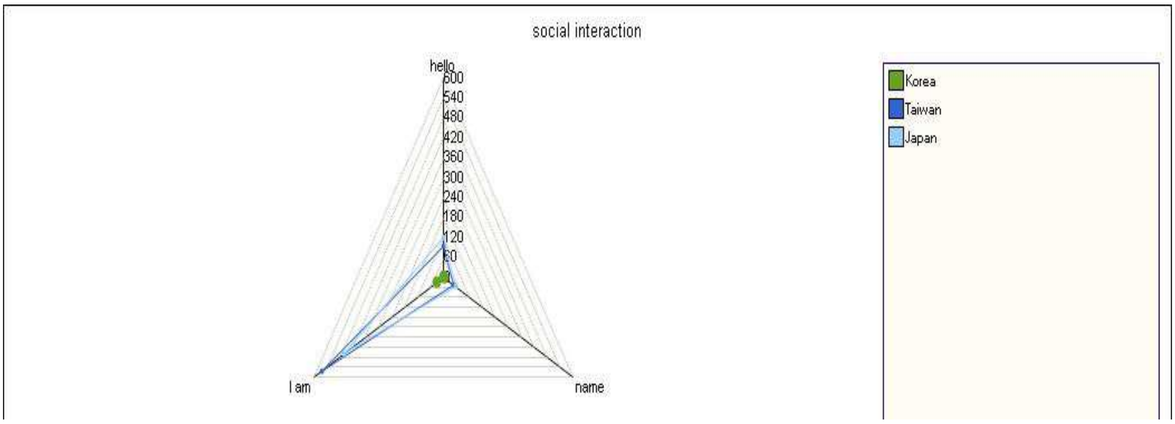


Figure 2: Social etiquette displayed by different cultural groups

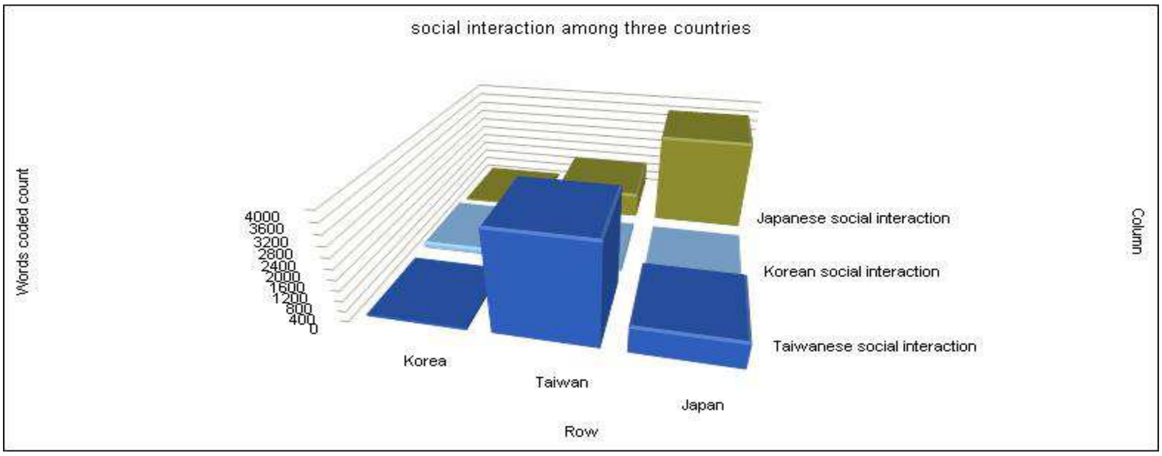


Figure 3: Social interactions among three countries

The function of cross-cultural CMC is to promote engagement among users of different countries in order to create productive discussions, which is considered beneficial for English learning in multicultural backgrounds in the current study. Based on the results of this study, it is clear that participants actually differ from one cultural group to another in terms of the nature of their social interactions.

Japanese participants usually included social etiquette in writing and meant to write for people of the same cultural group. Also, the Japanese style of social etiquette usually included only a short self-introduction which did not strengthen the interaction with other people. Taiwanese participants usually did not make cross-cultural interaction a social event. They were basically task-focused but would address someone when giving feedback. Korean students liked to share their English learning experiences with others, and were believed to be motivated to social interaction in a stronger way. But many Korean students' failure to use Nicenet actively was considered to be another lack of CMC motivation.

Overall, the weak social interaction among participants of the three countries resulted in feedback that was often too short, on many occasions only one or two sentences. Not knowing each other, participants tended to reply more to people of the same cultural group, or to their own classmates, which decreased the role of the multicultural experience in English learning. The primary lesson learned from this study was that intercultural social interactions will not happen automatically in CMC contexts. The study suggests that students are less likely to communicate with others whom they do not know. Unfamiliarity weakened the flow of interaction among all participants. As a result, instructors need to begin intercultural CMC with friendship building activities and exercises first, followed only later by the curriculum-based learning assignments. When the participating students become friends first, they will then be less likely to communicate only within their cultural group. Intergroup climate does play an important role in cross-cultural asynchronous CMC project. Unfamiliarity among participants will constantly influence the flow of interaction, feedback giving, and will decrease motivation for participants from different cultural group. But teachers have the ability to shape the intergroup climate into a setting in which students feel safe and comfortable in sharing ideas outside their cultural groups.

References

[1] Fuchs, C. (2007). Student language teachers as intercultural learners in CMC-based project work. *Journal of Intercultural Communication, 13*.

[2] Garrett, N. (2009). Computer-assisted language learning trends and issues revisited: integrating innovation. *The Modern Language Journal 93*(1), 719-739.

[3] Kawamura, M., Wu, P.H.N. & Jung, J. (2010) Collaborative teaching methodology: Global approach to teaching English as a foreign language. *Asian EFL Journal International Conference*. Taichung: Providence University.

[4] Levy, M. (2009).Technologies in use for second language learning. *The Modern Language Journal 96*(1), 769-782.

[5] Tolmie, A., & Boyle, J. (2000) Factors influencing the success of computer mediated communication (CMC) environments in university teaching: a review and case study. *Computers & Education, 34*, 119-140.

[6] Wrench, J. S. & Punyanunt-Carter, N. M. (2007). The relationship between computer-mediated-communication competence, apprehension, self-efficacy, perceived confidence, and social presence. *Southern Communication Journal 72*(4), 355-378.

[7] Yashima, T. (2002) Willingness to communicate in a second language: the Japanese EFL context. *The Modern Language Journal, 86*(1), 55-66.

[8] Yashima, T, Zenuk-Nishide L., & Shimizu, K. (2004) The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning, 54*(1), 119-152.

[9] Ziess, E., Isabelli-Carcía, C.L. (2005). The role of asynchronous computer mediated communication on enhancing cultural awareness. *Computer Assisted Language Learning 18*(3), 151-169.

[10] Wu, P.H.N. & Kawamura, M. (2012) Willingness to communicate online—Asian students in a cross-cultural English learning project. A.M. Stoke (Ed.) JALT 2012 Conference proceedings.

Collaborative Languaging for L1 Learning in a CSCL Classroom via Group Scribbles: An Exploratory Case Study

Wenli CHEN*, Wenting XIE & CheeKit LOOI
Learning Sciences Lab, Nanyang Technological University, Singapore
*wenli.chen@nie.edu.sg

Abstract: A micro-generic case study has been conducted to explore students' collaborative languaging practices in a CSCL L1 classroom. Descriptive analysis shows that students were highly engaged in the collaborative task and their group work was efficiently coordinated, during which required linguistic forms of accuracy and appropriateness were constructed, so did the related linguistic knowledge. This reveals the significance of collaborative languaging and technology intervention to enhance learning in L1 classrooms.

Keywords: Collaborative languaging, L1 learning, CSCL, case study

Introduction

Social-Cultural Theory endorses the significance of language in human cognitive development. Complex cognitive functions (e.g. problem-solving) are mediated by semiotic tools, among which language is the predominant [4]. Recent literature advocates conceptualizing learner language use as both the "process" and the "product" of cognitive activities [19]. When producing language, learners are engaged in meaning making, constructing tangible artifacts for further reflection [13]. Swain [13] proposed the concept of "languaging", "a dynamic and never ending process of making meaning and shaping knowledge and experience through language use", to describe cognition involving language. Thoughts are emerged, expressed and transformed in words [18]. Rendering thoughts into language is a process where thinking reaches a new level of articulation [11]. As evidence accumulates, languaging being a legitimate source for learning is recognized. In existing literature, self-explanation has proved enhancing scientific concept learning [2]. Collaborative dialogues and private speeches in natural conditions are positive to L2 development, so do the elicited "self-dialogues" in forms of immediate reports, think aloud, and stimulated recalls [15]. Written languaging (including typed texts) also fosters learning [12]. Yet, in a CSCL L1 classroom, students' languaging practices are inadequately examined. This study is intended to contribute to this topic.

1. Research Context

In our design-based research, we are investigating how to improve students' L1 (English) learning in secondary schools via integrating a network technology, Group Scribbles (please refer to [10] for GS introduction). As learning is socially grounded and mainly internalized via language [8], collaborative languaging about language, i.e. "learners work together to solve linguistic problems and co-construct language or knowledge about language" [17], has been identified as the pathway to L1 learning and encouraged in our intervention. Collaborative languaging is also the target for analysis in that cognition which is dialogically derived can be observed directly in learners' linguistic interactions in problem-solving tasks [5] and that it can be captured as natural and intact in real classroom settings [17]. This paper explored students' collaborative languaging via GS (i.e. what did students language about in collaboration? how did students language in collaboration?) and the role of networked technology intervention via a micro-genetic case study.

Our partner school provides 1:1 networked computing environment and students are comfortable and competent with ICT-mediated learning. English GS lessons are implemented in a Grade 2 class of 22 students who were randomly distributed into 5 groups each of 4-5. In GS lessons, students each was provided a MacBook and seated face-to-face in physical proximity. Interaction over dual spaces (online + F2F) was supported. The lesson reported here was on the module of Persuasive Writing where students wrote an argumentative essay on the topic of Cyber Bullying. The GS activity designed focused on essay planning (Table 1). Considering task difficulty and student capacity, scaffolding prompts in a graphic organizer were provided on Group Public Board (Figure 1). The except chosen for examination was on Content Organizing and Linearizing, an important and difficult phase in argumentative essay writing [3]. We selected one group (Figure 2), whose group product was assessed as the best, as the case for analysis. The data included group artifacts, group audio and video transcripts, individual student Morae transcripts. Student GS act and verbal talk was comprehensively and chronologically documented in transcripts.

Table1. Collaborative Learning Activity: Persuasive Writing

Activity	Description	Time
Introduction	Teacher shows a video clip about Cyber Bullying	5mins
Content Generation	Intra-group interaction: brainstorm to mine ideas\ arguments \ examples	10mins
	Inter-group interaction: Gallery walk-visit other groups' boards for more ideas	5 mins
Organization & Linearization	Intra-group interaction: 1) select, categorize, synergize and arrange contents 2) devise thesis statement \topic sentences	20mins
	Inter-group interaction: Presentation-present group work and offer comments	10mins

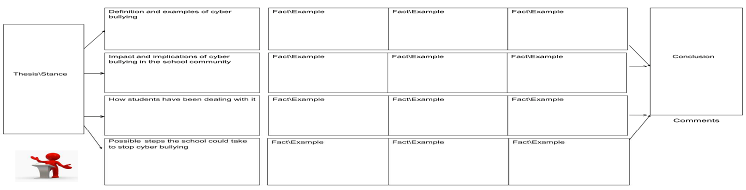


Figure 1. Graphic Organizer on Group Board

Fiona	Kim
Michael	Peter

Figure 3. Group sitting arrangement

2. Data Analysis and Discussion

The focus of analysis was on the ***Area*** and the ***Type*** of collaborative languaging occurred. For Area, we examined the problems the group encountered and executed via languaging: whether they were engaged in constructing linguistic forms\knowledge (Language Bound Languaging, LBL), Group Coordination, or Off-Task interactions. As LBLs are beneficial to language development [16], we further studies the Type of LBLs (how languaging was achieved) and its relation to the interactional medium. The categorizing framework used in Kuouzi [7] in examining individual learners' private speech in learning a grammar concept (Voice in French) was translated into this study (Table 2). One turn in conversation or GS text on one scribble was coded as a unit. Each unit was then tagged by ***Medium*** and ***Area***. LBL unites were further coded by ***Type***.

Table 2.1 Area of collaborative languaging

Category		Description
Language Bound Languaging (LBL)	Language (LBLL)	Group members language to construct the group product (e.g. topic sentence). e.g. -- *Cyber bully has negative impacts on the…*
	Language Knowledge (LBKL)	Group members language to construct language-related knowledge (vocabulary, schematic, etc) that helps constructing the group product. e.g. --*Overlook means you just ignore. Oversees means take charge.*
Group Coordination Languaging (GCL)		Group members language to coordinate group work (e.g. negotiation of working procedures or inviting for assistance). e.g. -- *Skip that first.*
Off-Task Languaging (OTL)		Group members language about off-task topics. e.g.--*What?…(laughing)*

Table 2.2 Type of collaborative languaging (LBL)

Category	Description
Repetition (R)	Group members repeat (parts of) the previous neighboring languaging unit. e.g. -- *On school. <u>On the school community</u>.* --*<u>On the school community</u>*
Integration (I)	Group members, apart from repeating (parts of) the previous neighboring unit, introduce new thoughts. e.g. --*Is like, <u>I think it's like</u>…* --*<u>I think it's like</u> the summary*.
Elaboration (E)	Group members, repeat the previous unit(s) (not the neighboring one)and introduce new thoughts e.g. -- *Cyber bullying has a lot of <u>negative</u> impacts* … -- *But you emphasize on the <u>negative</u> part ok, because…*
Formation (F)	Group members introduce new thoughts not occurred before in the languaging sequence. e.g. -- *I think we should separate both terms out , cyber and bully. Cyber is the virtual world-*

2.1 Area of collaborative languaging

In the 20-minute activity, the collaborative task was completed successfully with 11 sequences constructed (1 Thesis, 4 Topic sentence, 6 Example\Fact) (Figure 3). In total, 209 languaging units were observed and only 2 (1%) was on Off-Task interaction, suggesting active participation and equal contribution of students in group work (Table 3).

In solving the complex linguistic problem, students came across difficulty and divergence in developing conceptual knowledge on stylistics (*thesis statement*) and vocabulary (word connotations: *ability vs skill vs knowledge vs maturity; overlook vs oversee*). Through 26 languaging units (12.4%), they constructed and improved understanding on these concepts. In Vygotskian interpretation of mind, scientific concepts develop from spontaneous concepts (existing forms) and the merging of the two underlines concept maturation and cognitive development [9]. In L2 learning, languaging on individual basis will help learners mediate between known linguistic forms and the conceptual system [7]. In our study, languaging in a collaborative manner also contributed to the evolution of scientific concepts in L1. As scientific concepts transform spontaneous ones, making them structured and conscious [8], the attainment of linguistic concepts will bring about improved language use.

The other LBL units (141, 67.5%) were devoted to constructing linguistic forms. Through F2F and online languaging, ideas on Cyber bullying were organized and translated into language. Languaging in collaboration enhanced language learning by improving accuracy in grammar (e.g.:-Fiona: GS text <u>Cyber bullying has a negative impact **to schools** and the school should take action</u>*;* -Peter: "*On school. On the school community.*") and vocabulary (e.g.:–Kim: "*Some students are able to deal with the problem, its deal or-*"; Peter: "*Deal with, deal with the problem.*"), all good to language development.

In the learning activity, group coordination was efficiently achieved through verbal talk. Analysis of the 40 units (19.1%) showed that students were brave to express "Don't Knows" (e.g.-Fiona: "*So what…I didn't know how to write a thesis statement.*") and willing to seek peer assistance (e.g.-Fiona: *"Can you take it done?"*). Scaffolding prompts were found useful in directing group focus. One reading over a prompt made the group immediately focus on that particular piece of problem and initiated the pooling of ideas (e.g.-Kim: "*Ok, never mind. Let's come back. Possible steps that school should take to stop cyber bullying*").

2.2 Type of collaborative languaging

Further analysis of LBL reveals the prevailing repetition of previous neighboring unit (Repetition + Integration=126 units, 75.49%) (Table 4). This was probably attributed to the vast adoption of F2F languaging at intra-group level. As verbal talk is of temporal logic, the included old information served as the anchor in conversation, enabling students to track and continue the topic. Besides enhancing collaboration, repetition, which denotes a deep level of reprocessing as learners realize that the langauging introduced is not thoroughly comprehended, also contributes to improvement in L1 proficiency (though learners are not involved in cognitively complex work) [7]. Apart from absorbing ideas from others, contributing own ideas was also found necessary for effective collaborative languaging. This was indicated by the titanic number of Integration units (117, 69.6%), compared with the sheer amount of pure Repetition (10, 6%). Though differ in L1 proficiency, personality and preference of interaction medium (Unlike others inclining to verbal talk, Michael worked primarily online), students made active contribution to the construction of language of validity, accuracy and persuasiveness in collaboration.

2.3 Medium for Collaborative Languaging

F2F interaction was the dominant (92.8%), both in group coordinating and LBLs (especially in LBKL). It was probably due to the strategy adopted in group work. In the beginning, students agreed to firstly verbally discuss over and decide on the linguistic form, and then to use GS texts to record it. The strategy was chosen very likely by taking into account of proximity in physical distance, richness of social cues and lightweight interaction in F2F. Yet, the spontaneous production of verbal talk often resulted in "slips of tongues" and its logic of temporality frequently caused production blocking and called for clarification, all slowing down the working process. So when the time issue arose (e.g.: Fiona: "…*we are too slow, really slow*…"), online interaction increased as it permits paralleled expressing and processing of ideas. Moreover, when producing written language, one is more conscious about the accuracy and appropriateness of the language, reducing the time for correction. Thus besides recording and consolidating collective knowledge, the new medium is also good for presenting ideas, especially when the time constraint exits. This was supported by the dominating Formation in LBL units via GS (GS:72.7%; F2F: 7.2%).

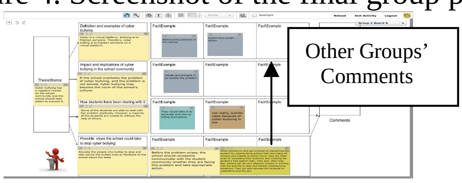
Table 3. **Area** of collaborative languaging units

Area	LBL		GCL	OTL
	LBLL	LBKL		
F2F	60.3%	12.4%	19.1%	1%
GS	7.2%	0%	0%	0%

Table 4. **Type** of Language Bound Languaging units

Type	R	I	E	F
F2F	6%	64.7%	14.4%	5.9%
GS	0%	4.79%	0%	4.2%

Figure 4. Screenshot of the final group product



3. Conclusion

In this paper, a micro-genetic case study has been performed to explore students' collaborative languaging practices in a CSCL L1 classroom. In analysis, students were observed as actively participating and engaged in collaborative learning where they successfully construct linguistic forms and knowledge, all good to L1 development. From this analysis, we better understand the significance of enriching collaborative languaging for L1 development and engaging networked technologies to foster languaging in classroom settings. However, considering the scope and specificity of the study, any application of the conclusions to other learning scenarios should be done with caution.

References

[1] Brooks, L., Swain, M., Lapkin, S., & Knouzi, I. (2010). Mediating between scientific and spontaneous concepts through languaging. *Language Awareness*, *19*(2), 89-110. doi:10.1080/09658410903440755

[2] Chi, M., de Leeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

[3] Coirier, P., Andriessen, J. E. B., & Chanquoy, L. (1999). From planning to translating: the specificity of argumentative writing. In J. E. B. Andriessen & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 1–29). Amsterdam, The Netherlands: Amsterdam University Press.

[4] Cole, M. (1996). *Culture in mind*. Cambridge, MA: Harvard University Press.

[5] Donato, R., & Lantolf, J. P. (1990). The dialogic origins of L2 monitoring. *Pragmatics and language learning, 1,* 83-97.

[6] Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*(3), 463-486. doi:10.1016/j.chb.2004.10.038

[7] Knouzi, I., Swain, M., Lapkin, S., & Brooks, L. (2010). Self-scaffolding mediated by languaging: Microgenetic analysis of high and low. *International Journal of Applied Linguistics*, *20*(1), 23-49.

[8] Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.

[9] Lave, J., &Wenger, E. (2005). Practice, person, social world. In H. Daniels (Ed.), *An introduction to Vygotsky* (pp. 149–156). London: Routledge.

[10] Chen, W., Looi, C.-K., & Tan, S. (2010). What do students do in a F2F CSCL classroom? The optimization of multiple communications modes. Computers & Education, 55(3), 1159–1170. doi:10.1016/j.compedu.2010.05.013

[11] Smagorinsky, P. (1998). Thinking and speech and protocal analysis. *Mind, Culture, and Activity, 5,* 157-177.

[12] Suzuki, W. (2009). *Languaging, direct correction, and second language writing: Japanese university students of English.* Unpublished Doctor's Thesis, University of Toronto.

[13] Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In Heidi Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). New York: Continuum International Publishing.

[14] Swain, M. (2010). Languaging: Private speech and collaborative dialogue. In M. Swain, P. Kinnear, & L. Steinman. (Eds.), *Sociocultural theory in second language education: An introduction through narratives* (pp. 34-50). New York: Multilingual Matters.

[15] Swain, M. (2010). Talking it through: Languaging as a source of learning. In R. Batstone (Ed.), *Sociocognitive perspectives on second language learning and use* (pp. 112-130). Oxford: Oxford University Press.

[16] Swain, M., Lapkin, S., Knouzi, I., Suzuki, W., & Brooks, L. (2009). Languaging: University Students Learn the Grammatical Concept of Voice in French. *The Modern Language Journal, 93*(1), 5-29. doi:10.1111/j.1540-4781.2009.00825.x

[17] Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. Annual Review of Applied Linguistics, 22, 171-185.

[18] Vygotsky, L. (1987). *Problems of general psychology*. New York: Plenum.

[19] Wells, G. (2002). Learning and teaching for understanding: The key role of collaborative knowledge building. *Social Constructivist Teaching, 9*, 1-41.

Technology Enhanced Movie Presentation with Focus on Foreign Language Anxiety and PBL Skills

Yuichi ONO[a*], Manabu ISHIHARA[b] & Mitsuo YAMASHIRO[c]
[a]*Foreign Language Center, University of Tsukuba, Japan*
[b]*Department of Electrical and Computer Engineering, Oyama National College of Technology, Japan*
[c]*Information Science and Manufacturing Engineering, Ashikaga Institute of Technology, Japan*
*ono.yuichi.ga@u.tsukuba.ac.jp

Abstract: This paper presents a new pedagogical model of movie presentation as a foreign language course employing existing technologies to enhance motivation to communicate actively in a foreign language. The "Technology Enhanced Language Learning (TELL)" technologies adopted in our study are Computer Assisted Language Learning (CALL) system, presentation or editing software, learning management system (Moodle), and so on. After illustrating the outline of the system conducted in our study and description of the lesson plan, this paper argues that the proposed course was helpful as a new approach to academic presentation course in a foreign language teaching. A questionnaire research was conducted in our study concerning students' awareness of their "Project-based Learning (PBL)" skills and foreign language anxiety. The results of our factor analyses of significant items showed that three factors were abstracted as underlying affecting structures; they are (i) *Anxiety in the Classroom*, (ii) *Information and Computer Skills*, and (iii) *Class Engagement*.

Keywords: CALL, Presentation, Digital Storytelling, Foreign Language Anxiety, PBL

Introduction

The importance of providing students with opportunities to reflect themselves on how they progress or achieve toward their goal is emphasized in various educational settings, and this is true in the field of foreign language teaching. Previous study [1] argues that the movie production activity motivates students with great anxiety toward speaking in public on the basis of the case study carried out in a national college of technology in Japan. Some students who are fearful of speaking a foreign language in front of people sometimes prefer non-linguistic PBL activities like searching the net, discussing the issue, using the computer, designing the layout, and so on. This paper presents a course model introducing TELL and PBL and examines how their awareness of PBL skills and foreign language anxiety changes after movie production course.

1. Video Production in Foreign Language Teaching

1.1 TELL and The Role of Digital Storytelling in Foreign Language Teaching

Our course involves the use of technologies in many ways and computer-assisted teaching method is well encouraged. Recent Computer Assisted Language Learning (CALL) System is designed in terms of implementation of socio-constructivism as an integrative model ([2][3]). The merits of conducting digital storytelling in a classroom are mentioned in studies like [4][5][6][7] and [8]. As for the incorporation of computer-assisted language learning and the digital storytelling, [8] provides a detailed description on an integrated course involving technology enhanced language teaching and digital storytelling in a Japanese university. There are some studies concerning the effects of introducing digital storytelling into course curriculum on the oral proficiency and motivation to learn a foreign language ([6]). [7] and [9] carried out a collaborative digital storytelling project and showed that students' motivation increased as a result of engaging in the project. In these studies, however, each *group* creates one movie in a collaborative work, which sometimes leads to a situation where only some students did well in a group. The CALL room will make it possible to conduct individual movie making. But very few studies have dealt with the effect of digital storytelling in a CALL environment.

1.2 Foreign Language Anxiety

Anxiety has been shown to negatively influence foreign language learning including achievement ([10] and [11]) and one of the most widely used language learning anxiety indices is the Foreign Language Classroom Anxiety Scales (FLCAS). Reading, writing, as well as listening in a second language can trigger anxiety, but speaking seems to be most anxiety-provoking ([10] and [12]), perhaps because of the requisite immediacy of the response. In the case of academic-level oral PowerPoint presentation in front of people, some unaccustomed presenters speak, or even "read", an academic level English script holding a piece of paper and operate PowerPoint at the same time, which leads to panic in front of the audience. Students who prefer movie presentation to oral presentation in general say that the course is unique and interesting, and they want to try it once more.

1.3 Course Plan

The class was conducted in a CALL-utilizing compulsory class called "Integrated English" for first-year students of a Japanese university in 2011. The CALL model of the course is described in Figure 1 below. The main aim of the course is awareness of cross-cultural understanding. Since there was a flood disaster and the capital was greatly damaged in October in 2011, we picked up "Thailand" as a topic for project work in 2011. Each student made a two-minute movie about Thailand on the basis of their research on what happened in Thailand and their various impacts on Japan and the whole world. They are encouraged to include its cultural and historical aspects in their presentation. Each student went through process of group discussion and peer evaluation. The course plan is given in Table 1 below.

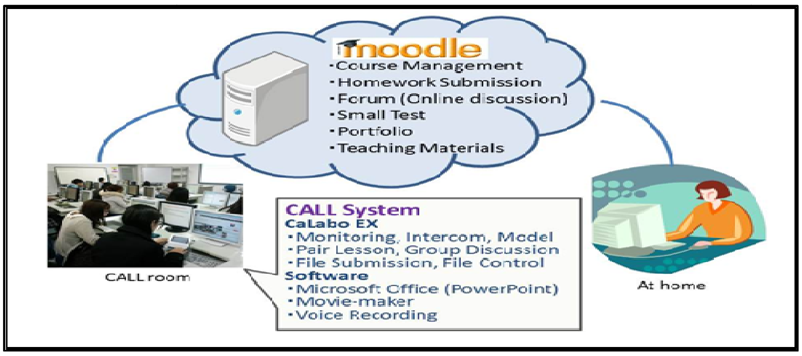


Figure 1. Outline of Our Teaching Model

Table 1. Course plan

Week	Lesson	Content	Activity	Technologies
1	Introduction of Project Work	Understanding of backgrounds of Thailand Watching news on youtube	Discussion on Thailand Decision of title	Internet browser, On-line discussion, File download
2	Searching, Organization	Completion of organization worksheet PowerPoint creation	Discussion on organization Individual work	File control/submission, Office (Excel, PowerPoint), On-line discussion
3	Recording, Edition	Sound recording Movie edition	Creation of first movie Peer-evaluation on first movie	File submission, Text-to-speech, Sound recorder, Movie maker, On-line discussion
4	Evaluation (1)	Evaluation	Completion of evaluation sheet	File control/submission, Office(Excel), CALL(Income)
5	Evaluation (2)	Evaluation	Completion of evaluation sheet	File control/submission, Office(Excel), CALL(Income)

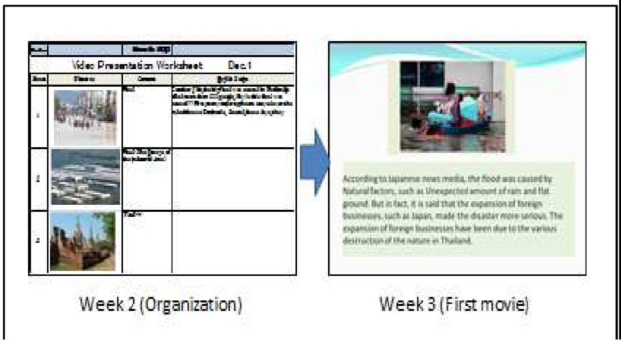


Figure 2. Development from Week 2 to 3

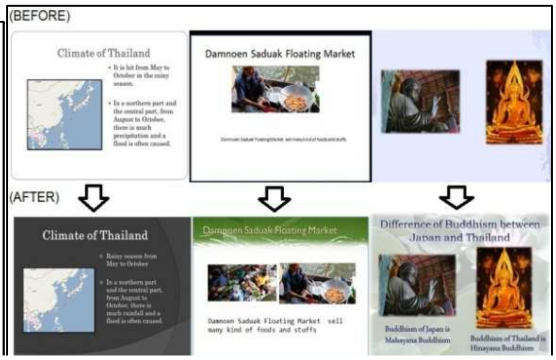


Figure 3. Students' Revisions

Collaborative session is an important component in our course plan, because their ideas are clarified, infelicities of their work or pronunciation are noticed through this session. Functions of intercom and pair/group work of CALL helped a lot to facilitate their collaborative learning. Pictures in Figure 3 above show the examples of changes of student's work before and after collaborative sessions.

2. Research

This paper tries to investigate how their awareness toward PBL skills and foreign language anxiety changed before and after the course and what factors are working on their change. A total of 40 first-year students at a national university in Japan participated in the study. The total number of students analyzed is less than this number because of absence or other reasons. All the students completed our course of movie production. The study was carried out for five weeks from November to January in 2011-2012 based on the lesson plan given in Table 1 above. The materials included instruments that have been well validated and widely used in the language learning literature. Minor modifications were made for the Japanese EFL context.

The material used in the study consists of two sections:

(1) Instrument of asking students their PBL skills used by [13] was employed in our study. The items were translated into Japanese. Question items are related to subcategories like searching, computer literacy, cooperation, planning, achievement, creativity, and so on.

(2) A Japanese version of FLCAS ([10] and [14]) was employed for this study. Participants were presented with various statements and asked to indicate on a five-point scale (from 1: It does not apply to me at all to 5: It applies to me completely).

3. Results

All the 64 question items (30 items for (1) and 34 items for (2)) were analyzed and went through a t-test to see if there is a significant difference between pre- and post- surveys. The result was that only three out of 30 items on presentation skills showed a tendency of significance, all the items of which, interestingly, are all related to searching process with the use of computers. Moreover, about half of the items of foreign language anxiety showed significance or tendency of significance. Then, we conducted a factor analysis to observe interrelationship among clustered items in each factor. All these significant items were subject to a principal axis factor analysis with Promax rotation. Examination of the scree plot and various different solutions resulted in three factors accounting for 65.7% of the total variance (KMO=0.601) The items with factor loadings (>|.35|) are shown on Table 2. Factor intercorrelations between Factors 1 and 2, Factors 1 and 3, and Factors 2 and 3 are 0.08, 0.20, and 0.44, respectively.

Table 2. Factor loadings of significant items

		Factors		
		1	2	3
44	It embarrasses me to volunteer answers in my English class.	-.891	-.089	.487
51	I can feel my heart pounding when I'm going to be called on in English class	.727	.062	.101
64	I get nervous when the English teacher asks questions which I haven't prepared in advance.	.721	-.217	.411
46	I get upset when I don't understand what the teacher is correcting.	.620	-.101	.067
60	I get nervous when I don't understand every word the English teacher says.	.491	.191	.125
5	skill in clarifying problems	-.051	.805	-.018
10	skill in sorting information and necessary data	-.008	.786	.054
3	technical skills with computers	.069	.537	.274
48	I often feel like not going to my English class.	.124	.010	.596
38	I keep thinking that the other students are better at English than I am	-.024	.045	.580
36	It wouldn't bother me at all to take more English classes.	.023	.273	.511

As shown in Table 2, Factor 1 contains items concerning anxiety of speaking or understanding English in the classroom. "Anxiety in the Classroom" was labeled. Factor 2 contains items related to information skills (clarifying problems and sorting information) and computer skills. We named Factor 2 "Information and Computer Skills". Factor 3 contains attitudes toward English class and inferiority feelings among students. Since improvement of these items may result from more awareness toward class engagement, we labeled Factor 3 "Class Engagement".

4. Discussions

Factor 1 seems to imply that movie presentation did not cause students pressure or anxiety compared with speaking in public. However, a closer examination of Factor 1 reveals that there is one item with high negative value, which means that answering in public negatively affected students after the digital storytelling. It might be possible that the vocabulary used for video became too difficult for less-confident learners to cover and that the course did not change their attitude toward active communication in public. In order to overcome this situation, it might be necessary that the new model include Q-A practice sessions using the vocabulary used in the video.

Factor 2 implies that the CALL-based PBL activities worked well in terms of computer literacy and information skills. However, other components of PBL skills like planning, cooperation, achievement, and creativity were not affected by this course. Group movie presentation might have resulted differently. The study [14] suggests that development of self-understanding depends on theme content. We need to study further on the examination of effects possibly resulting from kinds of project theme

5. Conclusions

This paper presented our technology enhanced pedagogical model employing CALL for individual-based video presentation course. The study investigated students' awareness toward PBL skills and anxiety and showed that the course model reduced foreign language anxiety and enhanced information skills and computer literacy for the project. Decreasing foreign language anxiety is expected to enhance learner's motivation to use a foreign language. There is a high possibility that CALL-based movie production is another way to get access to academic English presentation, or improve learner's PBL skills especially for learners who do not have confidence to speak in front of people.

References

[1] Ono, Y. (2008). Podcasting system-wo katsuyou-shita English presentation-no jyugyo (English presentation class utilizing podcasting system). *Information Communication Technology: Practice & Research 2008*, 245-253.

[2] Warschauer, M., Shetzer, H. and Meloni, C. (2000). *Internet for English Teaching*. Longman, New York.

[3] Takeuchi, O. (2008). *CALLjyugyou-no tenkai (Development of CALL classes)*. Shohakusya, Tokyo.

[4] Bell, N. (2008). Visualizing and Vervalizing for Language Comprehension and Thinking. Download on April. 15, 2012, from http://www.psllcnj.com/visualizing_and_verbalizing.htm.

[5] Dal, M. (2010). Digital video production and task based language learning. Ráðstefnurit Netlu – Menntakvika 2010. Menntavísindasvið Háskóla Íslands. Sótt af http://netla.khi.is/menntakvika2010/alm/021.pdf

[6] Kimura, M. (2010). Visualizing and Verbalizing by Digital Storytelling in English Reading class in Japanese University. In D. Gibson & B. Dodge (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010*, 1479-1486.

[7] Agawa, T. (2012). Cooperative learning in digital storytelling: A way to raise university English learners' motivation. *THE LANGUAGE TEACHER: 36.1*. 11-16.

[8] Obari, H. (2009). Integration of e-learning and m-learning in teaching EFL in Japan. *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009*, 1009-1015.

[9] Ono, Y., Ishihara, M. and Ideo, M. (2009). A Design and a Practice of ELDP Project━LMS, Podcasting, and English Presentation Class. *Proceedings of the 5th International CDIO Conference, Singapore Polytechnic*. Bloomington, IN: Indiana University Press.

[10] Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *70*, 125- 132.

[11] Young, D. (1986). The relationship between anxiety and foreign language oral proficiency ratings. *Foreign Language Annals*, *19*, 439- 445.

[12] Price, M. L.. (1991). The subjective experience of foreign language anxiety: Interview with highly anxious students. In: Horwitz, E. K., Young, D. J. (Eds.), *Language anxiety: From theory and research to classroom implication. Englewood Cliffs*, NJ: Prentice Hall.

[13] Miyaji, I. (2011). Effects of creating digital storytelling by three kinds of themes, T. Hirashima et al. (Eds.), *Proceedings of the 19th International Conference on Computers in Education. Chiang Mai, Thailand: Asia-Pacific Society for Computers in Education*, 531-538.

[14] Yashima, T., Noels, K., Shizuka, T., Takeuchi, O., Yamane, S., and Yoshizawa, K. (2009). The Interplay of Classroom Anxiety, Intrinsic Motivation, and Gender in the Japanese EFL Context. *Foreign Language Education Studies*, Vol. 17, 41-64.

A Study on the Use of Prepositions mediated by an ICT Tool

Suriati ABAS
Beacon Primary School, Singapore
suriati.abas@bcps.sg

Abstract: This paper describes how an English language teacher from a primary school in Singapore employed an ICT tool to understand the use of prepositions amongst nine-year-old children. Using the written composition of 28 primary three pupils in a Modular lesson that she has taught, the teacher created a learner corpus, named it CLAL (Corpus-based Learning about Language), and used Ant Conc, a corpus analysis technology, to identify patterns in her pupils' usage of prepositions, 'in', 'on' and 'at'. Drawing from the preliminary findings, the teacher then attempted to adopt the data-driven learning (DDL) approach in the teaching of grammar, harnessing on similar technology. She concludes by showing two angles in which Ant Conc could be used in classrooms.

Keywords: learner corpora, corpus analytic toolkit, prepositions, ICT tool

Introduction

The use of prepositions in English language is a concern. Among the twenty most frequently used words, eight are prepositions: of, to, in, for, with, on, at and by (Kucera & Francis, 1967). Leech and Svartvik (1975) defines prepositions as words which connect nouns or noun phrases with other structures in a sentence. Most prepositions are simple, 'short, invariable forms' (Biber, Johansson, Leech, Conrad & Finegan, 1999) such as, at, for, in, into, on, off, to and with, but, there are also others which consists of multi-word units. Some examples of such complex prepositions include, apart from, because of, such as, in addition to, and in spite of. Although prepositions are said to appear as soon as a child can produce two word utterances (Kochan, Morgenstern, Rossi & Sekal, 2007), pupils generally encounter difficulties in the use of even one-word prepositions.

1. Studies conducted on learners' use of prepositions

Learners of varying age group had problems in using prepositions. After a freshman writing exam, Scarcella (2002) reported that approximately 60% of the university students failed and had to attend a remedial freshman writing course despite their previous schooling experiences. To add on, one of the grammatical difficulties which they experienced in their writing include the use of prepositions. The prepositions are often either absent or, used incorrectly. Meanwhile, in another study conducted by Connors and Lunsford (1998) on college students' writing, prepositions surfaced as the list of frequency of errors made by students. The findings clearly indicate that prepositions are one of the language areas that should be addressed in classroom teaching. At the secondary level, Silayong (1984) affirmed that Thai students encountered problems in the use of prepositions in English due to interference from their mother tongue language. In similar vein, Mariano (1984) highlighted that the fourth grade students of Juan Sumulong Elementary School in

Philippines made mistakes when using simple prepositions such as *in, on, over, beside, under* and *behind,* because they had a hazy concept of the meanings. Furthermore, the multiple personalities taken by prepositions, which may be inconsistent and capricious (Wahlen,1995), contributed to their non-standard use. Additionally, it is possible for several prepositions to be used for similar purposes. For instance, *in* the afternoon, *on* Thursday afternoon and *at* night, are used to indicate time, albeit their differences.

Retrospectively, while non-corpus based studies as mentioned above have shown that prepositions are one of the problematic areas (Lindstromberg,1991;Capel,1993), previous corpus-based research in similar language area which involved English language learners merely focus on using a concordancer to teach prepositions and compositions.In the teaching of prepositions, Daud & Abusa (1999) claimed that the use of concordance output helped learners discover the use of in, on and at, in multiple contexts. On the other hand, through corpus consultation, in a research which involved Korean in-service teachers, Lee, Shin & Chon(2009) discovered that there were significant improvements in their use of vocabulary, but not for grammar.When the compositions were rated, those from the post-training writing tasks reveal problems in prepositions and collocations. Likewise, in the written assessment of native speakers at the masters and undergraduate level, O'Sullivan and Chambers (2006) gathered that prepositions and word choice are one of the most common problems possibly due to native language interference between English and French. Thus far, in the local context, only Mei Ling (2007) has used corpus-based materials to teach prepositions. Her findings suggest an alternative approach which could possibly benefit teachers and pupils in language learning.

2. An ICT tool that helps to understand pupils' use of prepositions

Drawing my knowledge from previous studies on the use of prepositions, while at the same time, leveraging on the ubiquitous use of computers in my school, I attempted to use Ant Conc. Ant Conc is a freeware concordance program. It could be obtained from Laurence Anthony's website at http://www.antlab.sci.waseda.ac.jp/software.html. My intent in employing such ICT tool is mainly to identify and understand the patterns of usage in the prepositions, 'in', 'on' and 'at' from my pupils' written compositions since the freeware facilitates the identification of 'linguistic and situational co-occurrence patterns' (Reppen, 2010). AntConc is basically a corpus analysis toolkit. A corpus is a large, principled collection of naturally occurring texts (written or spoken) stored electronically. By 'naturally occurring texts', I am referring to diverse language that is obtained from authentic language situations such as daily conversations, meetings, letters, class assignments and books, rather than made-up language (Reppen, 2010). Over the years, with the advent of technology, Ant Conc has proved to be effective in classroom context (Noguchi, 2004). For instance, in the learning of vocabulary, learners were found to acquire new words by looking at huge examples of varied natural contexts (Cobb, 1999). Meanwhile, in the teaching of grammar, the use of corpora enables teachers and pupils to observe nuances of usage (Hunston, 2002) to better understand specific language items.

3. Features of Ant Conc

For this particular study, the freeware application sufficiently affords a set of features for analysing a small-sized corpora. In fact, by analysing a corpora, it is possible to gain insights into the typical linguistic contexts of a word (Hunston, 2002) instead of simply relying on how people normally think it should be used. However, since the corpus of words was meant to “highlight the regularities which are hidden from the naked eye” (Lorenz, 1999) in two written compositions of a group of 28 pupils, therefore, I created a learner corpora, which I named it as CLAL (Corpus-based Learning about Language).

4. Data collection

In CLAL, I managed to input 62 texts with an average of 200 words per composition, making up a corpus of 10 000 words. Although the corpus is very small as compared to most present-day corpora, as Leech (1991) argues, size is not important. In this case, the small corpora serve as a sample for a specific investigation (Gavioli, 1997:88) of which, I decided to focus on the use of prepositions. Nonetheless, a small corpus such as CLAL cannot be considered a representative sample of the written compositions of primary three pupils in Singapore.

5. Analysing the patterns of use

In order to investigate the patterns of use for the prepositions, ‘in’, ‘on’ and ‘at’, in primary three pupils’ written compositions, I input CLAL into the concordance tool in AntConc. The prepositions (‘in’, ‘on’ and ‘at’) were typed into the search term, individually, to retrieve related sentences. Since the patterns of a word is defined as “all words and structures which are regularly associated with the word and which contribute to its meaning” (Francis & Hunston, 2000), therefore, I examined the concordance lines which were automatically generated from the search results. From the concordance output, it is possible to find out the word use and how the same word can have multiple meanings (Reppen, 2010). I then grouped these lines as a set, to explain the patterns of their usage in my pupils’ written compositions.

6. Preliminary findings

From the CLAL output in Ant Conc, I observed that ‘in’ has the highest frequency of 187, followed by ‘at’ with a total frequency of 116 and ‘on’ which has 83 of such occurrences in the two written compositions of the primary three pupils. Generally, the prepositions, ‘in’, ‘on’ and ‘at’ tend to be used most commonly in the pupils’ writing for describing a position and expressing time. In terms of position, pupils have been using ‘in’ to describe the state of being in a three-dimension enclosed space such as ‘in his house’. They use ‘at’ to denote a point in a space, such as ‘at the right hand corner’, and ‘on’ to tell the position on a surface which could be ‘a table’ or ’chair’. With regards to time, pupils use ‘in’ to tell part of a day as in ‘in the afternoon’, whereas ‘at’ has been used to indicate the specific time on a clock, ‘at 4pm’ as well as the time of the day, ‘at night’. Meanwhile, the preposition ‘on’ is solely used to state a particular date. Having analysed instances of how the pupils used ‘in’, ‘at’ and ‘on’ in their compositions using AntConc, those that are atypical from a teacher’s intuition were grouped into sets. Each of the concordance lines in these sets were compared with the written subcorpora in ICE-GB since the use of one variant over another could not be reliably predicted from our intuitions (Biber, Conrad & Reppen,1998). Concomitantly, Sinclair’s (1991) upward and downward collocates were applied to analyse the collocational pattern which might prove that the non-standard use of prepositions could be accepted.

While downward collocation involves a more frequent node (the word being studied), a, with a less frequent collocate, b, the reverse occurs in an upward collocation, which tends to show a weaker relationship as it pertains to grammatical property.

7. Linking main findings to previous studies

Putting together evidences from previous research on the problems students encountered in using prepositions (Mariano, 1984; Silayong,1984, O'Sullivan & Chambers, 2006), as aforementioned, it appears that similar findings could be deduced from the primary three pupils' compositions; In a nutshell, their non-standard use of prepositions lie on the fact that they had a hazy concept of prepositions, and that they could have experienced 'language interference' (Dulay, Burt & Krashen,1982) from their mother tongue, aside their inability to use them in a semantically inappropriate manner. Yet, to deduce that these children are not able to use the prepositions correctly because of several non-standard instances in their writing is irrational, for, what could possibly count in the prepositional meanings is how the object is construed, for the purposes of speaker and learner (Clark, 1993).

To gain a deeper understanding of my pupils' work, I then used Ant Conc to investigate whether the use of prepositions in their writing reveals further findings with regards to their language development. In order to do so, I had to generate the number of prepositions that were used by my own pupils in CLAL and compare them with those of the learners in the ICE-GB corpus. Since the size of the corpus is far smaller than the ICE-GB, therefore, I had to normalise it as shown in Table 1b below.

Table 1a Prepositions, 'in', 'at' and 'on' in CLAL & ICE-GB (written subcorpora)

Prepositions	**CLAL** 10 000	**ICE-GB(written)** 423,702
At	116	2128
In	187	8564
On	83	2925

Table 1b Normalised figures per million words

Prepositions	**CLAL** 10 000	**ICE-GB(written)** 423,702
At	116 x 100 = 11 600	2128/ 423 702 x 1million = 5022.4
In	187 x 100 = 18 700	8564/423 702 x 1million = 2012.3
On	83 x 100 = 8 300	2925/423 702 x 1million = 6 903.4

From Table 1b, it was evident that the nine-year-olds in this study used the prepositions, 'at', 'in' and 'on', more than the writers in ICE-GB. In fact, the frequency of use is twice or more in the three prepositions. This could be due to the differences in age group between the learners in CLAL and ICE-GB. Comparatively, it is partly the cognitive maturity that enables the older learners in ICE-GB to compose more abstract language than the nine-year-old children in the study (Taylor, 1974). Yet in reality, there are also exceptional cases of younger learners who acquire language proficiency earlier and are as competent as the older ones.

8. Conclusion

There are indeed two ways in which Ant Conc can be used in classrooms. Firstly, it is a useful ICT tool to inform language teaching. From the analysis, as a language teacher, I could gather that there are spatial and temporal patterns in the ways the primary three pupils use in, on and at. However, instances of their non-standard use of prepositions do not reveal any patterns as they are semantically inappropriate. Clearly, I was made aware that the pupils could have been confused by the fact that moments in time could simultaneously be construed as being analogous to points in space. At the same time, their mother tongue languages could have contributed to their non-standard usage of prepositions. Lastly, I speculated that the high frequency of prepositions used in the children's might be due to their lack of lexical knowledge.

Secondly, Ant Conc provides opportunities for using the data-driven approach to teach prepositions. Adopting DDL, I have introduced the corpus-analytic toolkit into my classrooms and train my pupils in their use. An advantage in using DDL is that it provides opportunities for 'grammatical consciousness raising' (Rutherford, 1987) by presenting learners with evidences and asking them to identify patterns and make generalisations about language form and use (Johns, 1986) from a set of texts or corpus which serves as a source of data (Doyle, 2007). In all, the teaching of prepositions mediated by an ICT tool such as Ant Conc, has made it possible for any language teacher, to gain insights in his or her classroom experiences.

References

[1] Biber, D., Conrad, S. & Reppen, R. (1998).Introduction to Corpus Linguistics. Cambridge:Cambridge

[2] Biber,D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). Longman grammar of spoken and written English. Pearson Education:China.

[3] Capel, A. (1993). Prepositions. London: HarperCollins Publishers.

[4] Cobb,T. (1999).Breadth and depth of lexical acquisition with hands-on concordancing.*Computer Assisted Language Learning*.12,4,345-360.

[5] Doyle, P. (2007). Teaching grammar with a learner corpus. In T. Ruanni, F. Tupas, Yuan Yi and Christianty Nur (eds.), *Changing Perspectives on Pedagogical Grammar*, Singapore: Singapore Association for Applied Linguistics, 77-86.

[6] Dulay, H., Burt, M. & Krashen, S. 1982, *LanguageTwo*. New York: Oxford University Press.

[7] Francis, G. & Hunston, S. (2000). Pattern Grammar. Amsterdam:John Benjamins.

[8] Johns, T. (1986). Micro-concord, a language learner's research tool. System, 14(2), 151-162.

[9] Kucera, H. & Francis, W.N. (1967). Computational Analysis of Present-Day English. Providence:Brown University Press.

[10] Kochan, A., Morgenstern, A., Rossi, C. & Sekal, M. (2007). Children's early prepositions in English and French: A pragmatic device. Proceedings of LingO. Oxford Postgraduate Conference in Linguistics.

[11] Lee, M., Shin, D., & Chon, Y.V. (2009). Online corpus consultation in L2 writing for in-service teachers of English. *English Teaching,64*(2), 233-254.

[12] Leech,G. & Svartvik. (1975). A communicative grammar of English.London:Longman.

[13] Leech,G. (1991). The state of art in corpus linguistics. In K.Aijmer & B. Altenberg. English corpus linguistics – Studies in honour of Jan Svartvik. London:Longman.

[14] Lindstromberg, S. 1991. (Re) teaching prepositions. *English Teaching Forum*. 29(2), 47-50.

[15] Mei Ling, A. (2007). Using corpus-based materials as a teaching resource for the teaching of prepositions in a Singapore primary school. Unpublished masters thesis. Retrieved March 1, 2010 from National Institute of Education Digital Repository (http://repository.nie.edu.sg/jspui/handle/10497/9).

[16] Noguchi, J. (2004). A genre analysis and mini-corpora approach to support professional writing by nonnative English speakers. English corpus studies, 11, 101-110.

[17] O'Sullivan, I., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing, 15*, 46-68.

[18] Reppen, R. (2010). Using Corpora in the language classroom. In Richards, J.C. (Ed.) United States of America:Cambridge University Press.

[19] Scarcella, R. (2002). Effective Writing Instruction for English Language Learners. California English, 7(4), 6-13. Retrieved March 18, 2010 from Education Research Complete database.

[20] Silayong, D. (1984). A comparative study of English and Thai prepositions with some suggestions for pedagogical application at secondary levels. SEAMEO Regional Language Centre: Singapore.

[21] Sinclair, J. (1991). Corpus, concordance, collocation.Oxford University Press: Hong Kong.

[22] Taylor, B. (1974). Towards a theory of language acquisition. Language learning. 24, 23-36.

Moving from the Conventional to the Virtual Classroom

Sarah Hsueh-Jui LIU , Yao-Ting SUNG, Kuo-En, CHANG, & Yu-Ju LAN*
National Taiwan Normal University, Taiwan
*yujulan@gmail.com

Abstract: This paper describes the field experience of two prospective teachers of Chinese as a Second Language (CSL) in the teacher-training programme, regarding the differences in pedagogic skills between the face-to-face and the virtual world of Second Life. The student teachers recounted how they made pedagogical changes in relation to the extent of the consciousness raising, decision making and perception towards the different environments. A qualitative approach was primarily adopted to elicit this information from the data that consisted of post-teaching interviews and reflection reports. The preliminary results confirm that, to some extent, the prospective teachers' pedagogical knowledge was influenced by their previous experience (e.g., foreign language learning). The findings derived from a self-rating scale further reveal that the student teachers felt confident about their pedagogical knowledge and skills in the face-to-face context, e.g., conducting activities, but not confident in the virtual classroom, e.g., managing students' conduct. Both the student teachers opined that teaching in the virtual world was feasible to engender the involvement and interest of learners. The teachers' knowledge transferred from the conventional classroom to the virtual world has implications for teacher education.

Keywords: Second Life, Chinese as a Second Language, pre-service teacher, reflective practice

1. Introduction

Recent years have seen an increase in interest in emerging technologies or online education to allow access to education to those who cannot attend courses or classes that take place in face-to-face traditional classrooms [2]. In light of this, a three-dimensional multi-user virtual environment (3D MUVEs), Second Life (SL) acts as an important platform to simulate real life to provide an immersive and social environment [3]. SL has a huge impact on second or foreign language (L2/FL) education [6], and the game-like environment engages learners in authentic or meaningful interaction.

For teaching on SL, although there are some commonalities between the traditional and the virtual contexts, the pedagogical knowledge and skills of the latter, such as the online tools and organising the learning activities needed for teaching, differ significantly from those required in the former [2]. However, helping pre-service or in-service teachers to develop technical skills consciously to deal with pedagogical and technological difficulties is crucial [2], but research of pre-service teacher training in language education in both physical and on-line classrooms has long been overlooked [4]. Hence, this study attempts to bridge the gap by examining the pre-service CSL teachers' pedagogical knowledge, whereby they modified their instructional use according to the constraints of the virtual environment, e.g., lack of verbal or physical clues. The purpose of this current research is to give the pre-service teachers a hands-on experience and to report on how they felt about teaching in the traditional and the virtual classroom. There are two research questions to be answered in this paper: 1) How does the previous learning experience relate to the pedagogical knowledge of CSL student teachers during the teaching practice? 2) How does the instructional awareness through reflective practice determine the differences in the two environments?

2. Methodology

2.1 Participants

Two female student teachers of CSL, namely, Jessica and Joan (pseudonyms were used to protect their privacy), were involved in this study. The 21-year olds had been studying at the Department of Applied Chinese Language and Literature at National Taiwan Normal University (NTNU) and spoke Chinese as their mother tongue. Also participating in this study were 17 preparatory academy students from NTNU, aged between 20 and 26, with multi-cultural backgrounds, such as Vietnam, Japan and Indonesia with an elementary level in Chinese. They were randomly divided into two groups (A and B) and this resulted in 8 students as Group A in the conventional classroom and 9 students as Group B in SL.

2.2 Design

The design of the pre-service teacher training included two stages. In the first stage, the participating student teachers attended 8 weekly face-to-face workshops, each lasting two hours, making 16 hours in total; this helped them become familiar with the two environments (e.g., practise in skills for conducting pair/group work and on SL tools like making/sending a ‘notecard’). Furthermore, during this stage, the instructor shared her past teaching experiences and ideas with the student teachers as well as gave them support in designing lessons and activities. They were also taught using L2/FL teaching approaches, e.g., the task-based approach. The purpose of this was to assist them in organising the sequences of a lesson. In the second stage, each student teacher taught both Groups A and B one after the other and employed a given teaching approach to deliver similar teaching contents in the two environments. More importantly, before teaching each lesson, the each student teacher demonstrated their teaching to two experienced tutors in the relevant field, who in turn provided the student teachers with feedback on their teaching contents or approaches. The demonstration of teaching engaged the student teachers in peer discussions.

2.3 Instrument

The research contained two main instruments: 1) student teachers’ written reflection reports with detailed descriptions of the implementation steps and objectives including teaching activities and their own opinions/comments on the activities; and 2) post-teaching interviews with an array of semi-structured questions, such as their own previous learning or teaching experience during this training programme.

2.4 Procedure and data collection

There were eight lessons in total, with four lessons each taking place in the conventional classroom (CC) and using Second Life (SL) between December, 2011 and January, 2012; each lesson had one topic and one teaching approach. For instance, one of the lessons, consisting of the topic ‘How to get there’ and the teaching approach of Audio-lingual method (ALM), was taught by Jessica. As a result, Jessica and Joan each taught twice in an alternating manner during the second stage; one first taught in CC and then in SL during weeks 1 and 3 while the other did the teaching during weeks 2 and 4. Immediately after each lesson, each prospective teacher was asked to fill in the self-reflection report. When all teaching sessions had been completed, both student teachers were interviewed individually by one of the researchers of this current study. Regarding the participating students, the conventional classroom group attended the CSL lessons on campus at NTNU whereas the other group of students accessed SL from their home, the library or the laboratory at NTNU.

3. Results and Discussion

This study examines the prior learning experience of the prospective student teachers in relation to their newly acquired teaching knowledge and how the prospective CSL teachers consciously made changes between the CC and SL. Due to the constraint of space in this paper, we present and discuss only key results from the collected data to answer the research questions that were given previously.

In answer to Research Question 1, the participating student teachers had had scant teaching experience in either a face-to-face classroom or an online environment, but had been learning English as a core subject in formal education since they were 9 years old. Apart from that, both of them had had some experience of learning another foreign language, such as French or Japanese. Jessica commented that learning French had had a huge impact on her pedagogic knowledge growth, and she had been greatly influenced by the language activities conducted in the class. She was even aware of what teaching approaches her French teacher had employed and what activities she performed, stating, '*I like one of my French teachers who happened to be Taiwanese… When in the class, we were instructed to use French as much as possible but only when we came across some new words or unfamiliar expressions in the language we used our own language…. I think she used some kind of task-based approach…. I remember an activity we did was called exchange of information. It's not only useful to learn new vocabulary words, but also to get us to speak French when completing a language task. We also practised dialogues or performed role plays… I really liked her activities… So, when teaching, I had made some changes to those activities to fit my lessons, using pictures to help students practise vocabulary… I truly aspired to her teaching enthusiasm and professional knowledge….*'

Unlike Jessica, Joan rarely talked about any teaching methodologies or activities that her teachers had employed in her Japanese class. However, she did mention that a lack of prior experience in observing how others taught online and insufficient knowledge regarding online resources in Second Life had made it difficult for her to manage teaching in the virtual environment. She commented: '*During the teaching practice, I felt totally lost because I couldn't see students' facial expressions or their body language that would give me a hint of whether they understood me… In one situation where I teleported them to a place for a task, only 3 students were there, and the other 4 students disappeared… I kept moving my avatar here and there to look for the missing ones… perhaps, I've never had experience for online courses or classes... If I got some chances to see how others teach online, I'd have been able to teach more successfully*'.

The preceding paragraphs show how the student teachers' previous learning experiences influenced their current practice of teaching, e.g., how they designed their classroom activities. This, in turn, helped them integrate what they had learned into the new teaching context(s). To some extent, this proves that the teacher-training programme helped them to develop their knowledge and skills through carrying out classroom activities; the implicit knowledge was made explicit via the personal reflections. Besides, the aforementioned discussions remind us that like experienced teachers reflecting on their teaching, novice teachers like Jessica and Joan, while not having many teaching experiences, still have the 'faculty to reflect' [1]. Such newly integrated knowledge built upon the prospective teachers' tacit knowledge (past learning experience) and their exploration of the theoretical aspects (their own ideological ideas) of the professional initiatives during the teaching practice.

In answer to Research Question 2, both student teachers reported that they were confident about their teaching in the conventional classroom in terms of a) tasks, skills and resources, including questioning and eliciting techniques, presenting and explaining a topic and etc.; b) adoption of teaching methodologies, e.g., TBA and TPR; and c) classroom management, such as giving instructions and managing students' conduct (see Table 1). In CC, both student teachers rated two components of (a) and (c) at nearly 4, apart from (b), whereas they rated the three aspects with around a score of 2 in SL.

Table 1 - self-rating of pedagogical skills

	Tasks, skills and resources (A)		Teaching methodologies (B)		Classroom management (C)	
	CC	SL	CC	SL	CC	SL
Jessica	4.2	2.2	3.7	2.3	3.8	2.4
Joan	3.5	1.8	3.0	2.0	4.1	1.5
Average score	3.9	2.0	3.3	2.2	4.0	2.0

(Note: the numbers indicate a 5-point rating scale: 1-5, from extremely unconfident to extremely confident)

In addition to what has been discussed so far, the two student teachers made some remarks in their self-reflection reports regarding the differences, difficulties and constraints they encountered in SL and the involvement of the participating students between the two environments. Jessica said, '*In the traditional classroom, I know that when things go wrong I can tell straight away from students' body language or verbal expressions. I remember that in my first time teaching, I gave instructions for an activity: 'giving directions'. Students were paired up to complete the task… Perhaps my instruction confused them and that caused them to start having a chat with their partners... However, in the virtual classroom, my first time teaching was chaos… Many students experienced technical problems, their voice breaking down and some of them couldn't display my power point slides. Suddenly, I got myself into panic… Indeed, I'd found teaching in Second Life more challenging than teaching in the conventional classroom. Virtual teaching not only requires technology skills, like the knowledge of fixing the technical difficulties, but also the knowledge of the language that I am teaching. However, the effects of learning in Second Life were quite straightforward as students were actively engaged in the virtual scenes.*

To echo the point just made by Jessica, an example given from Joan's reflection illustrated that when teaching students how to order food in Chinese, she noticed that students' responses were very different in both pedagogical environments. Whereas students in the traditional classroom were likely to be quiet and their responses tended to be short, students in Second Life actively responded to her questions and their utterances were longer with more new vocabulary: '*When I asked students some questions (in the traditional classroom), such as 'Did you try watermelon juice before?', though they seemed to understand my questions, they had come up with very short answers. But, students in SL were more active to give various answers and even give me some new vocabulary words… They were very responsive…*'. Joan went on to say, '*I felt… umm, perhaps, in the virtual environment, students weren't afraid to say anything they wanted… but in the face-to-face classroom, students might feel afraid of making errors when speaking…*'.

In the above discussions, noticeably, in accordance with their early language learning experiences, the student teachers may have been accustomed to a face-to-face context where a teacher was in control of the class, and so, when they first came across teaching in SL, they might have felt a loss of control. Such a feeling is not unusual, especially for those who are used to face-to-face classroom teaching or learning. However, community building and socialisation, whereby groups of members in the community share common beliefs and principles, are needed whether in a face-to-face classroom or an online environment [2] [5].

Conclusion

In this pilot study, we introduced the background of online education as well as the pedagogical benefits of using such a 3D virtual environment for language teaching and learning. Subsequently, the methodology of this study was presented in terms of the research design, including the participants, data, and procedures of data collection. Finally, the preliminary results derived from the post-teaching interviews and the self-reflection reports were given to answer two research questions. This study has some implications for teacher training. Teacher-training programmes are necessary to provide pre-service teachers with practical experiences whether in traditional or virtual settings, and the demand for learning in the virtual world places an emphasis on the need for teachers' knowledge and skills to be appropriate to teaching virtually. To conclude, the scope of this study is confined to describing the experiences of the two participating student teachers teaching in the two different pedagogical contexts; thus, the results of this study cannot be generalised due to the small sample size and inherently qualitative characteristics.

Acknowledgment

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract Nos. NSC 101-2511-S-003-031-MY3 and NSC100-2511-S-003-018. The authors also thank the Ministry of Education of the Republic of China, Taiwan, for financially supporting this research under Aim for the Top University Plan.

References

[1] Almarza, G. G. (1996). Student foreign language teacher's knowledge growth. In D. Freeman, & J. C. Richards (Eds), *Teacher learning in language teaching* (pp. 51-78). Cambridge, England: Cambridge University Press.

[2] Bennett, S., & Marsh, D. (2002). Are we expecting online tutors to run before they can walk? *Innovation in Education and Teaching International*, *39*(1), 14-20.

[3] Chiang, T. H. C., Tsai, T. T. C., Li, I. Y. S., & Hsiao, I. Y. T. (2011). Technology enhanced language learning in virtual worlds. In S. Graf, F. H. Lin, Kinshuk, R. McGreal (Eds), *Intelligent and adaptive learning systems: Technology enhanced support for leaners and teachers* (pp. 293-309). IGI Global.

[4] Grable, C., Kiekel, J. K., & Hunt, A. (2009). Digital pre-service teacher education: Field experiences as a possible augmentation to the traditional brick and mortar field experience. *Journal of Education, Informatics & Cybernetics, 1*(1), 1-6.

[5] Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning*, *18*(4), 311-326.

[6] Jarmon, L., Traphagan, T., Mayrath, M., & Trivedi, A. (2009). Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. *Computers & Education*, *53*, 169-182.

Evaluating Potential Effects of Digital Storytelling Websites for Promoting EFL Young Learners' Writing Skills

Feng-lan Kuo[a]*, Heien-kun Chiang[b], Yue-ru Lin[a], Ya-han Cao[a], & Huei-hsun Yen[c]
[a]Graduate Institute of Children's English,
[b]Department of Information Management,
[c]English Department
National Changhua University of Education, Taiwan
***laflkuo@cc.ncue.edu.tw**

Abstract: In this paper, based on the evaluation criteria proposed by Chapelle (2001) the five researchers evaluated three digital storytelling websites for promoting writing skills for EFL young learners. Among the evaluated websites of *StoryJumper*, *Tikatok*, and *Storybird*, *Storybird* was found to be the most appropriate website due to better learner fit, better teacher fit, and the extra provided functions of "Class Library" and "Discussion".

Keywords: Digital storytelling, writing skills, EFL young learners, *Storybird*

1. Introduction

Recent CALL research has suggested that digital storytelling may facilitate language learning (Mead, 2010; Oakley, 2011; Yuksel, Robin & McNeil, 2010). According to Barrett (2006), digital storytelling can be described as the convergence of four student-centered learning strategies including student engagement, reflection for deep learning, project-based learning, and the effective integration of technology into instruction. These student-centered learning strategies may bring the outcomes of enhancing student learning, student motivation, student engagement, building technology skills, and is more effective than paper-based reflection. As an innovative pedagogical approach and a helpful educational tool, digital storytelling can further encourage peer collaboration and peer communication, and foster learners' higher order thinking and deep learning (Smeda, Dakich & Sharda, 2010). Digital storytelling can also foster a sense of ownership in learners, let learners have a deeper understanding of the text, and facilitate learners to have a longer retention rate of the text (Mead, 2010). Yuksel, Robin and McNeil (2010) showed that most of their respondents indicated that it was useful to teach language arts by using digital storytelling. Digital storytelling can be used to improve learners' writing skills with modeled writing of digital texts (Oakley, 2011). When teaching young learners, it is important to engage the learners during the story writing process to facilitate sight words learning and text comprehension. Furthermore, digital storytelling writing instructions share some similar features of blogs. For instance, they provide more writing opportunities for learners, foster a sense of audience, gain feedback during writing (Liou & Peng, 2009), and have interaction between teacher-student and peer-to-peer. Another benefit of digital storytelling is that it can construct the learners' experience in the content, facilitate collaborative activities, promote in-class discussion, motivate learners to learn critical thinking skills, and foster understanding of complex ideas. During the process of making a short story in small groups, learners can learn problem-solving skills, enhance their information gathering, learn to work collaboratively, engage community, and interact with others. However, though digital storytelling has great promise nowadays, many digital storytelling websites are set up with uneven quality. Thus, appropriate evaluations of related digital storytelling websites are needed. To evaluate and choose the more appropriate digital storytelling websites for particular group of learners, teachers should pay attention to different website evaluation criteria and use the more appropriate ones based on their objectives and focus. Since the present study focuses on language learning, the five researchers thus implemented Chapelle's 2001 CALL criteria to evaluate three popular digital storytelling websites and choose the most appropriate digital storytelling website among the three which can potentially promote writing skills for EFL young learners.

2. Evaluation of the Websites

This section presents the evaluation of the three websites by the five researchers. Following Chapelle's 2001 evaluation criteria, each website was analyzed according to: language learning potential, learner fit, meaning focus, authenticity, positive impact, and practicality.

2.1 Evaluation of StoryJumper (http://www.storyjumper.com/)

The tasks provided in *StoryJumper* present sufficient opportunity for beneficial focus on learners' writing skills. Learners can do the writing practices without restrictions of time and space. Nevertheless, there is no comments function. Neither teachers nor learners have chances to give feedback. Furthermore, learners cannot see their classmates' work in the classroom environment. There is no interactional modification or modification of output.

Regarding the learner fit feature, there is no suggested level according to the users' age or language proficiency, thus teachers cannot choose appropriate tasks for the students according to their individual differences in linguistic ability level or non-linguistic characteristics. In contrast, users can upload their own pictures to make their storybooks. As a result, the learner fit feature is limited. Regarding meaning focus, the researchers considered that learners' primary attention is directed toward the language meaning by using *StoryJumper*. To make their own storybooks, the students need to learn more vocabulary items and sentence patterns to accomplish the task. Since the students can use the vocabulary items and sentence patterns to communicate with others outside the classroom, so the feature, authenticity, is taken into consideration. However, since there are not many different topics included in the practices, the connection between the CALL task and tasks outside the classroom is also limited. Concerning the positive impact, learners will potentially learn more about the target language and about strategies for language learning through the use of the task. To accomplish the tasks, the students need to learn more vocabulary items and sentence patterns and even some strategies. Since the storybook-making process is fun, both learners and teachers can have a positive learning/teaching experience with technology through the use of the task. Finally, since the hardware, software, and personnel resources are sufficient to allow the CALL task to succeed, *StoryJumper* thus has the feature of practicality.

2.2 *Evaluation of Tikatok* (http://www.tikatok.com/)

To evaluate the language learning potential feature, activities of *Tikatok* provide learners great chances to receive language learning. Focusing on writing skills, *Tikatok* helps learners focus on form when making their storybooks. While students create their own stories with various expressions, the teacher has opportunities to choose appropriate activities and to monitor students' works to help them focus on form. However, the peer-correction function on *Tikatok* is limited, so students didn't have chances to share their works with peers. Therefore, only teachers are able to monitor learners' works, but peers cannot share their works in the classroom environment.

Toward learner fit, *Tikatok* is suitable for learners of every age level. There are three entries for different-age users, so learners can choose the appropriate entry for creating stories. Moreover, the "Teachers" function allows teachers to give instruction or writing assistance to every student individually indicating that *Tikatok* is designed to fit the individual differences in linguistic ability level and non-linguistic characteristics. Regarding the meaning focus of *Tikatok*, learners' primary attention is directed toward the language meaning--to accomplish their storybooks. To complete a storybook, learners need to output the target language. Furthermore, learners can write different content by following different tips provided on the website, which allows learners to have more opportunities to write various works. Therefore, the element of meaning focus is taken into consideration. Concerning authenticity, the writing task in *Tikatok* is relevant to learners' language use beyond the classroom. They can create personalized books with different purposes. Therefore, learners can use the target language for specific purposes. Additionally, positive impact is guaranteed in using *Tikatok* as teachers' supplementary material. *Tikatok* not only provides colorful and vivid pictures for learners to create stories, but gives learners chances to upload their personal pictures to do storybook making. This function improves learners' learning interest in writing because they can share their imaginary thoughts with others. Students may also feel free to make notes according to their imagination. Finally, for the practicality of *Tikatok*, it is easy for learners and teachers to integrate this website into a class or language program. Only when teachers set up the classroom environment do they need to read the guidance and detailed information about the online classroom environment. For students, the creating storybooks operation interface is easy to learn and use.

2.3 Evaluation of Storybird (http://www.storybird.com/)

As for the language learning potential feature, the task conditions in *Storybird* present sufficient opportunity for beneficial focus on form. The language learning activities mainly consist of creating storybooks. Young learners can write along with their caretakers at home or with their teachers at school. Furthermore, students can interact with others through the use of written language in *Storybird*. There are "comment parts" in every storybook, and readers can give comments after they read the story. The caretakers or teachers can also monitor students' output. There is no time pressure, so the writing activities could be kept going until the teacher and students feel satisfied.

To evaluate learner fit, the *Storybird* tasks provide learners opportunities to work with a range of target structures appropriate to their level, and it depends on whom the teacher teaches in the classroom. Beginning learners can use *Storybird* to improve spelling skills, whereas the advanced learners can practice writing by group works or as individuals. Learners with different characteristics can use this website. Creative learners have enough opportunities to create their own stories during the writing process, while shy learners not brave enough to communicate with others face to face can use the "comment parts" to share their opinions with other users. Regarding meaning focus, since the website provides vivid pictures for users to create storybooks and share their own works with others, the learners' attention can be directed toward language meaning--writing a meaningful story. Moreover, to create the story lines and write the appropriate description about each picture, learners can communicate and work together to accomplish the task. They use language purposefully for constructing and interpreting meaning. Concerning authenticity, storybooks provided in the website are one of the authentic reading materials. Learners can read storybooks in the classroom, and this reading experience could potentially improve their reading motivation not only in the classroom but also outside the classroom. Additionally, during the creating process, learners can practice their writing skills. Therefore, the writing task is a language task. As for positive impact, by using *Storybird*, learners' interest may be engaged in the target culture in a way that will help develop their willingness to seek out opportunities to use the target language. The vivid pictures and writing activities are more interesting than normal classroom tasks, so learners' learning motivation could be enhanced. Finally, regarding practicality, *Storybird* is easy for the learners and the teacher to implement the writing tasks. The fundamental functions in the website are free, so users only need to become their members and sign up onto the website.

2.4 Summary

Among the three evaluated digital storytelling websites, they shared the same functions that users can create their own stories with spelling check function, which can potentially increase the opportunities of English writing practice. They are equipped with many vivid pictures to provide writing clues, so learners can follow the images to write stories. The three websites provide teachers the classroom setting function, thus learners' writing processes could be monitored. In addition, website users have opportunities to share their works with other website users. Finally, the three digital storytelling websites provide a learning environment without restrictions of time and space, thus they are appropriate to be used as teaching activities either in the classroom or as students' homework.

In contrast, there are some differences. As for the creating function, *StoryJumper* has basic writing functions, including adding pictures, writing the content, checking the spelling and sharing the works; however, no age level function is provided. In addition to the above functions, *Tikatok* has three extra kinds of writing works for different age-level users to create stories, while *Storybird* has seven suggested age levels for story makers to set to fit different-age audiences. Secondly, both *Tikatok* and *Storybird* have the giving comments function, while *StoryJumper* doesn't. Thirdly, regarding the rights the teachers have in the classroom setting, teachers in *StoryJumper* can neither edit students' works nor give comments. *Tikatok* gives teachers the right of editing and commenting on students' works, while only the teachers can give feedbacks to their students. *Storybird* provides the teachers with the rights of not only editing and commenting on the story contents, but also provides the commenting function between the student peers. Therefore, the whole class can share and comment on every student's work in the classroom setting. In addition, *Storybird* has the "Class Library" and "Discussion" functions for teachers and students to share works and discuss related topics, which may potentially increase students' writing opportunities.

The five researchers thus determined the appropriateness of the aforementioned websites for promoting EFL young learners' writing skills based on the six criteria proposed by Chapelle (2001) using scales ranging from 1 point to 5 points. One point was assigned to signify very inappropriate, whereas five points were assigned to signify very appropriate feature. Table 1 shows the mean score and standard deviation of every characteristic of the evaluated websites. With the highest average mean score ranging from 4.60 to 5.00, the five researchers considered *Storybird* as the most appropriate and functional digital storytelling website among the three evaluated websites for promoting learners' writing skills.

Table 1 *Mean Scores of Each Criteria for the Three Websites*

Criteria	***StoryJumper***		***Tikatok***		***Storybird***	
	M	**SD**	**M**	**SD**	**M**	**SD**
Language Learning Potential	3.20	0.45	3.20	0.45	4.80	0.45
Learner Fit	3.00	0.00	4.20	0.45	4.80	0.45
Meaning Focus	4.20	0.45	4.40	0.55	4.80	0.45
Authenticity	3.60	0.55	3.80	0.45	4.60	0.55
Positive Impact	4.40	0.55	4.80	0.45	5.00	0.00
Practicality	3.40	0.55	4.00	0.00	4.80	0.45

3. Conclusion and Pedagogical Implication

Based on Chapelle's evaluation criteria, the five researchers considered the three websites, *StoryJumper*, *Tikatok* and *Storybird* to be appropriate in providing supplementary materials especially for promoting EFL young learners' writing skills. However, *Storybird* provides better and private classroom functions, including "Assignment", "Library" and "Discussions." Additionally, *Storybird* provides seven age levels to fit different learners' proficiency levels. Furthermore, *Storybird* has the functions of peer collaboration and peer correction which are crucial elements in language learning. As a result, among the three websites, the researchers consider *Storybird* as the most appropriate website for enhancing young EFL learners' writing skills.

References

[1] Barrett, H. C. (2006). Researching and evaluating digital storytelling as a deep learning tool. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2006* (pp. 647-654). Chesapeake, VA: AACE.

[2] Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research.* Cambridge: Cambridge Univesity Press.

[3] Liou, H., & Peng, Z. (2009). Training effects on computer-mediated peer review. *System, 37*, 514-525.

[4] Mead, K. (2010). *Digital storytelling as an assessment tool.* (Doctoral dissertation). Retrieved from http://fullerton.academia.edu/KristiMead/Papers/1197690/Digital_Storytelling_as_an_Assessment_Tool

[5] Oakley, G. (2011). *Pre-service teachers creating digital storybooks for use in early childhood classrooms*. In M. B. Nunes and M. McPherson (Eds.), *Proceedings of the IADIS International Conference e-Learning 2011* (pp. 84-88). Rome, Italy: International Association for Development of the Information Society.

[6] Smeda, N., Dakich, E., & Sharda, N. (2010). Developing a framework for advancing e-learning through digital storytelling. In M. B. Nunes and M. McPherson (Eds.), *Proceedings of the IADIS International Conference e-learning 2010* (pp. 169-176), Freiburg, Germany: Victoria University Research Online Publications Store.

[7] Yuksel, P., Robin, B. R., & McNeil, S. (2010). Educational uses of digital storytelling around the world. Retrieved from http://digitalstorytelling.coe.uh.edu/survey/SITE_DigitalStorytelling.pdf

Design Principles for a Pedagogically-Sound Mobile-Enabled Language Learning System

Agnieszka PALALAS
George Brown College, Toronto, Canada
apalalas@georgebrown.ca

Abstract: This paper provides a brief overview of a longitudinal Design-Based Research study conducted with ESP students and practitioners at a Canadian college. The study resulted in two key outcomes, namely the Mobile-Enabled Language Learning Eco-System and corresponding design principles. The design guidelines presented in this paper focus on the pedagogical aspects of the mobile-enabled intervention viewed through an ecological lens. The purpose of this paper is to introduce practical guidelines for language teachers and instructional designers who seek to develop m-learning solutions.

Keywords: Mobile-Enabled Language Learning, design principles, DBR research

1. Introduction

This paper summarizes the key findings of an interdisciplinary study which investigated the design of a Mobile-Enabled Language Learning (MELL) solution. The MELL Eco-System (MELLES) which resulted from this Design-Based Research (DBR) study addressed the problem of inadequate aural skills acquisition for college ESP (English for Special Purposes) students. Drawing on previous studies (Palalas, 2009), the project looked into employing mobile technologies to expand ESP learning beyond the classroom.

The main purpose of the DBR study, thus, was to generate design principles for a valid MELL intervention addressing the development of listening skills amongst ESP adult students. The eighteen-month process of data collection and analysis resulted in a conceptual model and design principles for a MELL solution supporting flexible authentic ESP practice. Consequently, the MELLES system was designed, developed and tested through multiple iterations involving 163 students and ten experts from a variety of related fields. MELLES provided a framework and a prototype mobile solution for teaching and learning listening while integrating the other language skills. In addition, the study enhanced the understanding of the broader context of MELL. The role of the learning environment was also examined resulting in the Ecological Constructivist framework proposed as a learning theory appropriate for any MELL context (Hoven & Palalas, 2011).

This paper focuses on the key results of the final phase of the three-phase DBR study, namely the pedagogical design guidelines for generating an effective MELL system. These guidelines encapsulate the essential characteristics of MELLES (substantive emphasis) and the strategies needed to realize those features (procedural emphasis).

2. Background and Statement of Problem

Listening is a language skill which, next to speaking, was identified as the area requiring the most remediation amongst second language speakers at George Brown College in

Toronto, Canada (Palalas, 2009). Studies conducted at the college between 2007 and 2009 indicated that students needed flexible learning solutions to accommodate their busy schedules and provide additional language practice. Our learners appreciated the portability and convenience of mobile technologies, and the personalized learning across real-life language situations. Participants observed that interactive and engaging MELL activities were an effective approach to learning listening. Several researchers demonstrated that mobile technologies were deemed appropriate for language teaching and learning (Demouy & Kukulska-Hulme, 2010; Kukulska-Hulme & Shield, 2008).

It was thus necessary to investigate what kinds of MELL activities would promote aural skills acquisition and to determine the technical and pedagogical requirements of such educational intervention. A systematic process was needed to design, develop and test appropriate MELL instruction including a prototype system as an instantiation of the theory. No guidelines or standards were available for creating this type of mobile-technology-based educational intervention. Hence, the three-phase DBR research study aimed at the outcomes described in the following section.

3. Research Question and Outcomes

The study was guided by the main question which inquired into the salient characteristics of an effective MELL solution. Auxiliary questions were also generated at various stages of the study congruent with the research activities and feedback gathered at each phase.

3.1 Overarching Research Question

The research question driving the DBR study was derived from an investigation of the problem depicted in the previous section and related literature. The question inquired: What are the characteristics of an effective, pedagogically-sound MELLES for students' mobile devices, through which adult ESP students in a community college enhance listening skills, while expanding their learning outside of the classroom?

3.2 MELLES Intervention and Design Principles

The DBR research, guided by the overarching question, produced two key outcomes – a MELL intervention prototype and design principles (intervention theory). The study aimed to develop a practical, innovative, m-learning intervention to teach listening skills to adult ESP students. Based on participant feedback the MELLES solution was generated to support flexible language practice situated in the real-world language speaking context, including the streets and landmarks of Toronto. The MELLES tasks combined individual focused practice with group activities contextualized in dynamic communicative situations challenging learners to converse with native speakers. Several consecutive versions of stand-alone mobile applications and individual mobile-enabled listening tasks were produced before a more systemic framework resulted in integrating these tasks into a complete solution - the MELLES system. Its portal, namely the mobi-english.mobi website, was used for tests and summative evaluation of the MELLES design guidelines. The many cycles of the solution redesign coupled with the evolution of thinking resulted in a design framework guided by Ecological Constructivism (Hoven & Palalas, 2011). The conceptualization and development of the successive prototypes were driven by the DBR feedback and design guidelines emerging progressively from each cycle. Ultimately, these principles evolved into what the pilots demonstrated to be pedagogically useful guidelines.

4. Methodology

Consistent with the ecological lens applied to the study, the DBR approach provided a broader and more systemic methodology for the design and research processes completed in-situ. It also accommodated the evolution of design constructs and conceptual thinking. Moreover, this interventionist approach allowed for experts in the field of mobile learning, software design, wireless technologies and language learning to collaborate and contribute their input. Both students and practitioners worked together to generate a solution to an educational problem they had been experiencing at the college.

This study stretched from June 2010 to December 2011 encompassing three phases: Informed Exploration, Enactment, and Evaluation (Bannan, 2009). Informed Exploration included literature review, audience characterization, investigation of comparable design solutions, and qualitative data collection via interviews with experts and language teachers as well as student focus groups and a mobile device usage survey (n=191). Mixed data was collected from students (n=21) and practitioners (n=7) from the Schools of Business, Design, and Technology as well as Communications/ESL. The first phase produced a theoretical *ideal* to guide the design experiment.

The two main outcomes of the Enactment phase were the prototype and the redefined design principles. During this phase, prototype MELL conceptual models were proposed by the Design and Technology students. Subsequently, the constant evaluation of the numerous design ideas combined with the new ecological paradigm led to a more holistic solution, namely, a functional MELLES prototype: a mobile web-based system. MELLES was designed, developed and tested by the researcher in collaboration with students and practitioners. In total, the Enactment phase feedback was collected from 41 students and six experts. Mixed data were gathered through assignments, designer logs, focus groups, meetings and correspondence, as well as researcher observations and reflections.

Evaluation encompassed implementation, testing, evaluation, and the refinement of the MELLES design. It involved testing of MELLES by L2 students and their professors in a real-life setting in Toronto. Five groups of intermediate-level ESP students completed eight listening tasks using their mobile phones. Feedback was collected from 109 respondents through successive surveys, interviews, focus groups and communication via the Wiggio site. The multiple cycles of coding resulted in recurring themes which indicated the design features considered by participants as crucial for the desired MELL intervention. These themes were organized into two super-categories: Pedagogy and Technology. They were then validated by the quantitative data. The ensuing findings were encapsulated into the final design principles examples of which are presented below.

5. Findings: Design Principles

A set of interconnected design principles was extracted from the research findings and analysis. The ten *pedagogical* essential characteristics distilled from the feedback include:

1) Balanced combination of individual and collaborative (group work) tasks;
2) Learner-generated linguistic artefacts (audio, video, photos, images);
3) Game-like real-life communicative tasks;
4) Expert facilitation: scaffolding, feedback, and coordination;
5) Feedback mechanism (immediate and delayed);
6) Focus on authentic listening tasks in dynamic real-world communicative situations;
7) Support of self-paced individual audio tasks feeding into/preparing learners for the real-life tasks;
8) Integrate all four language skills but focus on listening outcomes;

9) Linguistic resources (task-related): relevant vocabulary, dictionaries, pronunciation, clear task directions and explanations, examples of language usage;
10) Support of out-of-class learning with in-class (f2f) instruction and practice (a blend of in-class and out-of-class context).

These key features of an effective MELL solution need to be supported by corresponding procedural strategies. Examples (10 of 72) of substantive and procedural guidelines are mapped out in Table 1 to demonstrate their relationships.

Table 1. Examples of pedagogical MELLES design principles (substantive-procedural map)

Strategy (*Procedural Emphasis*)	**Essential Characteristic (*Substantive Emphasis*)**
1) Ensure communication and interaction with others in-person and via mobile-enabled channels	1) 1, 4
2) Include discourse with diverse interlocutors including L1 speakers	2) 1, 3, 5, 6
3) Incorporate language problems requiring negotiation of solutions	3) 1
4) Inject fun, enjoyment and challenge	4) 1, 3
5) Ensure dynamic meaning-making and negotiation	5) 1
6) Maintain regularity of group/class activities	6) 1
7) Build individual tasks to feed into the group tasks	7) 1
8) Include audio recordings (video, images, photos) created by students in response to communicative tasks	8) 2
9) Share and showcase learner-generated linguistic artefacts	9) 2
10) Provide tools for recording, editing, upload and viewing/listening on-the- go (or demonstrate device built-in tools)	10) 2, 6, 7, 10

6. Discussion

A MELL system developed following the above design principles, allows for the development of new knowledge and language skills based on the interactions among content, pedagogical procedures, mobile technology, learners and experts, as well as the context of learning. Consistent with Ecological Constructivism, MELLES relies on collaboration in the real-world context which mediates communication and, thus, language learning. While the real-life language situations create communicative challenges, the linguistic affordances offered by the environment provide support for meaning-making which, in turn, stimulates and generates language acquisition. In addition, the MELLES network of peers, experts and authentic language speakers facilitates learning by way of authentic discourse, feedback, resource sharing and social support.

Mobile devices are the tool which enables such communicative exchanges and the interaction over the MELLES web. They also help point to contextual affordances and capture linguistic evidence by way of learner-generated artefacts (Hoven & Palalas, 2011). Accordingly, learning supports come from MELLES resources accessed via the device. Mobile technology is thereby the enabler and one of the interconnected components of the MELLES learning network (Figure 1).

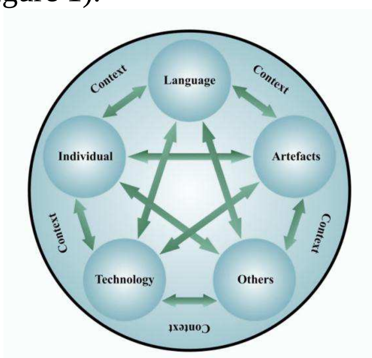


Figure 2. Interconnected elements of the MELLES learning context

All in all, MELLES enables practice of listening as part of a *whole* language experience that integrates all four language skills in a *whole* context of real-world communication out and in the class. While students complete eight integrated-skills tasks, regular in-class instruction should be combined with the out-of-class practice and linked into a cohesive learning experience by way of the MELLES platform and its tools. Furthermore, offering on-demand links to the system promotes social, cognitive, teaching, and emotional presence (Swan et al., 2008). This results in a collaborative network which has become the predominant structure of the recommended intervention. Adult learners engaged in meeting the demands of everyday life, need the support of a learning community to provide help and motivation. Considering how significant the notion of motivation was to participants, MELLES encourages continuity of practice through its emphasis on collaboration and communication as well as the design of its interlinked learning modules. MELL modules integrate time- and place-flexible individual listening tasks with collaborative game-like tasks completed in the dynamic language setting. These situated group tasks occur regularly throughout the semester to provide face-to-face peer interaction and motivate learner engagement. Finally, continuity of practice is also encouraged by means of feedback and expert facilitation which are offered in-person and digitally through the MELLES portal. All study participants, based on their perceived learning and positive learning experience, considered MELLES to be an effective approach to learning listening skills.

7. Conclusion

This eighteen-month-long DBR study incorporated voices of learners and practitioners resulting in two major outcomes, specifically (1) replicable MELLES design principles and (2) a prototype of a learner-centred MELL system that facilitates acquisition of language skills. Other key outcomes of the study included (3) an enhanced understanding of the broader context of learning ESP using mobile devices and (4) the role of the constituent elements of the learning environment. In addition, the study contributed the real-life praxis of (5) the Ecological Constructivist framework and the (6) DBR methodology which suited the dynamic character of the intervention under study. Future research is recommended to optimize the MELLES functionality and sustainability, as well as its applicability in other learning contexts.

References

Bannan, B. (2009). The Integrative Learning Design Framework: An illustrated example from the domain of instructional technology. In T. Plomp & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 53-73). SLO: Netherlands Institute for Curriculum Development.

Demouy, V., & Kukulska-Hulme, A. (2010). On the spot: Using mobile devices for listening and speaking practice on a French language programme. *Open Learning: The Journal of Open, Distance and e-Learning, 25*(3), 217-232.

Hoven, D., & Palalas, A. (2011). (Re)-conceptualizing design approaches for mobile language learning. *CALICO Journal, 28* (3)

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. ReCALL, 20(3), 271-289.

Palalas, A. (2009). Using mobile devices to extend English language learning outside the classroom. In D. Metcalf, A. Hamilton, & C. Graffeo (Eds.), *mlearn2009: 8th World Conference on Mobile and Contextual Learning. Proceedings* (pp. 179-183). University of Central Florida.

Swan, K., Shea, P., Richardson, J., Ice, P., Garrison, D. R., Cleveland-Innes, M., & Arbaugh, J. B. (2008). Validating a measurement tool of presence in online communities of inquiry. *E-Mentor, 2*(24), 1-12

Engagement with Livemocha as an Informal Learning Resource - Initial Findings from a Technology University Reading Course in Central Taiwan

Gloria Shu-Mei CHWO[a*], Yu-Shih LIN[b], Po-Jung CHEN[b], Guan-Wen LAI[b], Huei-Chin LIU[b], Chiao-Ching HO[b] & Yi-Ping WANG[b]
[a] *Associate Professor of Applied English Dept., HungKuang University, Taiwan*
[b] *Undergraduate School of Applied English Dept., HungKuang University, Taiwan*
*cho2.cheng@msa.hinet.net

Abstract: This study assesses whether learning style and learning strategy impact on our technology university EFL learners' degree of satisfaction, learning attitude, and other general perceptions with respect to Web 2.0 informal learning with the *Livemocha* program. A mixed method study was conducted using interviews and a questionnaire based survey. The interview results showed that writing, chatting and speaking are the most attractive functions for our EFL learners. From the survey, no significant correlation was found between learning style or strategy and attitudes to *Livemocha*. However, significant correlations were found among the four measures of attitude to and perception of *Livemocha*: participants' learning attitude, degree of satisfaction and judgment of visual design and website function. Pedagogical significance is discussed for future implementation.

Keywords: *Livemocha*, visual interface design, website function design, degree of satisfaction, learning attitude, learning strategy, learning style

1. Introduction

Among all the modalities, computer assisted language learning (CALL) has been acknowledged to have the greatest potential to accommodate diversity and enhance collaboration in language instruction settings (Keobke, 1998; Kim, 2009; Chen, 2002). Likewise, Web 2.0 continues to gain prominence as a result of the need to facilitate and promote communication and sharing among users worldwide (Gonzalez & Louis, 2008; O'Reilly 2005). Gonzalez & Louis (2008) further elaborate that 'Negotiation of meaning is a must for language learning to take place, and this is the only possibility when there is interaction through real communication' (p. 29). However, such claims need to be assessed for our context. Specifically, we need to consider how these resources can be effectively integrated into the language learning experience of Taiwanese technology students. For these reasons as teacher-researcher, we undertook the present research investigating our students' perceptions of *Livemocha* Web 2.0 resources in their own context.

The *Livemocha* system supports an interactive online community driven by the power of social networking. This community was the first of its kind: a web-based language learning solution integrating online instructional content with a global community of language learners (Liaw, 2011). *Livemocha* is available in thirty-eight languages in the free version and eleven languages in the paid version- English, Spanish, French, German, Mandarin Chinese, Hindi, Japanese, Icelandic, Italian, Portuguese (Brazil), and Russian. Three key learning features are (a) available lesson content (b) access to a global community of learners and native speakers and, in the paid version, tutors (c) motivational benefit. The operation of *Livemocha* can be briefly described as follows.

The *Livemocha* system requires learners to register at no charge. The website is user friendly with readily accessible Help videos. *Livemocha* is organized into four pedagogical areas: Home, Learn, Practice, and Share. The lesson structure of the Learn page in the paid version, which was used in this study, consists of seven activities from which to choose: (1) Learn (2) Reading (3) Listening (4) Magnet (5) Writing (6) Speaking (7) Dialogue. On the Share page, learners can submit writing and speaking samples which tutors can correct.

The essence of *Livemocha* is to help build a social network of foreign language learners from all over the world and to provide a learner community which allows users to share their language learning process through interaction. In conformity with socio-cultural learning theories, *Livemocha* is predicated on the assumption that human cognitive development is highly associated with the social context; learning takes place not only on the individual level but significantly involves social interaction with others.

One of the most important features of the program is appropriate feedback which corresponds to the spirit of community. Receiving feedback on specific content helps to heighten learners' ability to notice not only errors in their language use but also linguistic features to be learned that have not been noticed before. It must be mentioned however that the quality and accuracy of the feedback from peers with low proficiency remains questionable.

The program is designed to enhance learner autonomy as users look for learning opportunities outside the classroom, engage in the language learning community in a self-paced, self-access language course, and integrate their linguistic knowledge into their personal framework facilitated by social relations with people who share the same interest.

Intercultural understanding is another area to be established through the ongoing real life conversations with informants from other countries in the chat program. Making friends with native members of the target culture can open up additional channels for intercultural communication for English language learners in English as a foreign language context who normally do not have direct contact with native speakers.

Learners of a given language benefit from the community and contribute to the community as native speakers of another language. This approach supports the idea that L2 learners should be considered as active language users rather than passive learners and that their interlanguage is of great value in their learning process, rather than a faulty version of the target language.

The *Livemocha* website is still in short supply of tutors and learning materials for learners of varying proficiency levels. A disparity also exists between the nature of the exercises with respect to what the students actually do versus the skill that the exercises are supposed to promote.

Learners can improve their learning autonomy and motivation by checking their progress regularly and by selecting and doing the activities whenever they desire. The new concept of a learning community can serve as an alternative approach for those who want to learn a language while they are interacting with friends all over the world.

2. Research Questions

1. What is the general perception and liking that our EFL learners have of *Livemocha* website learning?
2. Is there any correlation between the EFL learners' satisfaction with the program and attitude to learning with it and their judgment of the visual interface design and website function design of it?
3. Is there any correlation between our EFL learners' individual learning style and learning strategy on the one hand and their survey responses concerning visual interface design, website function design, degree of satisfaction, and learning attitude on the other?

3. Methodology

Thirty free Active English *Livemocha* accounts were provided by a professor in a national university in Central Taiwan. After a brief orientation, students volunteered to participate in a trial period of use of the *Livemocha* program. An agreement form was signed with individual participants giving their permission to access their accounts and practice records for research purposes. *Livemocha* served as informal learning outside class time.

3.1 Participants

Thirteen day-school Taiwanese English major sophomore students volunteered to participate in the six month *Livemocha* trial. They attended a required associated course entitled Advanced Reading and Discussion in a Technology University in Central Taiwan.

3.2 Instruments

A learning strategy survey (Yang, 1992; Oxford & Burry-Stock, 1995; Oxford, 1990; Yang, 1999), learning style survey (Reid, 1995), were used together with a *Livemocha* survey targeting visual design (Chen, 2007), website function (Chen, 2007), degree of learning satisfaction (Shi, 2003; Huang, 2000; Wu & Chan, 1992; Chen, 2004; Feng, 2004; Chen, 2003), and learning attitude (Lin, 2003).

3.3 Procedure

The *Livemocha* trial was launched in September 2011 and lasted for six months in conjunction with the Advanced Reading and Discussion course, as an informal learning resource. Initially four tutor sessions introducing *Livemocha* were provided in the self-access learning centre to ensure a smooth start for those who volunteered. The instructor reserved ten accounts each for elementary, intermediate and advanced level students. Seventeen students volunteered for the *Livemocha* trial, but within the first four weeks, four students decided to drop out, so a total of thirteen students completed the trial. Students were encouraged to explore the functions of the *Livemocha* program at their own pace and provide their learning logs for instructor's and researcher's reference. At the end of the trial, two interviews and the questionnaire survey were conducted to investigate our EFL learners' perception of various aspects of the *Livemocha* program. Learning style and learning strategy surveys were also administered afterwards in order to gauge whether individual variation can be accommodated by *Livemocha* learning.

4. Results and Conclusion

In answer to the first research question, writing, chatting and speaking were the most attractive functions for our EFL learners. In answer to the second research question, learning attitude is significantly correlated with visual design and website function ($p < 0.01$), degree of satisfaction is significantly correlated with visual design ($p < 0.01$), website function ($p < 0.05$) and learning attitude ($p < 0.01$). In answer to the third question, no significant correlation was found between learning style or learning strategy and the measures of perception of and attitude to *Livemocha*.

The fact that attitudes to *Livemocha* do not correlate with learning strategy and learning style appears to suggest that the program equally suits all types of student. This is a major advantage since we would not want Web 2.0 CALL only to be attractive to students with certain styles or strategies for learning. On the other hand the fact that learning attitudes and degree of satisfaction both significantly correlated with website function and visual design suggests that the latter are critical for the success of any program such as *Livemocha*. Future study should further explore all these variables in order to enable teachers and web designers to optimize learning results and achievements.

References

[1] Chen, J-C. (2002). *Investigating distinct instructional mythology: Comparing website learning and traditional learning* (MA dissertation in Chinese). Jung-Shan University, Taiwan.

[2] Chen, J-T. (2007). *Learning theory on website teaching design and learning satisfaction* (MA dissertation in Chinese). Jung-Yun University, Taiwan.

[3] Chen, S-L. (2004). *The impact of learning motivation, learning strategy and knowledge orientation and culture on e-Learning effect* (MA dissertation in Chinese). National Jung-Shan University. Kaoshiung, Taiwan.

[4] Chen, Y-M. (2003). *The impact of learning style and mode on the junior high students' electronics learning achievement: An example of 240 students in Yun-Ling Jing-Tong junior high school* (in Chinese). Retrieved March 12, 2012 from http://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi?o=dnclcdr&s=id=%22091CCU00396012%22.&searchmode=basic

[5] Feng, D-S. (2004). *Motivation and learning effects via website by on job training for elementary and junior high school teachers* (MA dissertation in Chinese). National Taipei Normal University, Taipei, Taiwan.

[6] González, D. & St. Louis, R. (2008). The use of Web 2.0 tools to promote learner autonomy. *Independence*, *43*, 28-32.

[7] Huang, W-J. (2000). *The impact on learning achievement via interactive website teaching environment* (MA dissertation in Chinese). National Jun-Shan University, Taiwan.

[8] Keobke, K. (1998). Computers and collaboration: Adapting CALL materials to different learning styles. In J.M. Reid (Ed.), *Understanding Learning Styles in the Second Language Classroom* (pp. 46-51). Upper Saddle River, NJ: Prentice Hall, Inc.

[9] Kim, I-S. (2009). The relevance of multiple intelligences to CALL instruction. *The Reading Matrix, 9*(1), 1–21.

[10] Lin, S-J. (2003). *Learning style, learning strategy and website learning achievements: An example of program design* (in Chinese). Retrieved March 12, 2012 from http://libproxy.ncue.edu.tw/cgi-bin/cdrfb3/gsweb.cgi?o=dstdcdr

[11] Liaw, M. L. (2011). Review of Livemocha. Language Learning & Technology 15(1),36-40.

[12] O'Reilly, T. (2005). *What is Web 2.0: Design patterns and business models for the next generation of software*. Retrieved May 25, 2012 from http://www. oreillynet.com/pub/a/oreilly/tim/news/2012/05/25/what-is-web-20.html

[13] Oxford, R.L. (1990). Strategy inventory for language learning (SILL). In R.L..Oxford (Ed.), *Language Learning Strategies What Every Teacher Should Know* (pp. 293-300). Boston, MA: Heinle & Heinle Publishers.

[14] Oxford, R.L., & Burry-Stock, J.A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the strategy inventory for language learning (SILL). *System*, *23*(1), 1-23.

[15] Reid, M.J. (1995). Perceptual learning styles preference questionnaire. In M.J. Reid (Ed.), *Learning*

Styles in the ESL/EFL Classroom (pp. 202-207). Boston, MA: Heinle & Heinle Publishers.

[16] Shi, H-J. (2003). The impact of learning style and methodology based on interaction mode (MA dissertation in Chinese). Jung-Yen University, Taiwan.

[17] Wu, J-J., & Chen, B-L. (1992). Amendments on learning strategy charts (in Chinese). *Chinese Assessment Association Annuals*.

[18] Yang, N.D. (1992). *Second language learners' beliefs about language learning and their use of learning strategies: A study of college students of English in Taiwan* (Doctoral thesis). University of Texas at Austin, USA.

[19] Yang, S.C. (1999). An overview of research on language learning strategies. *The Journal of National Institute for Compilation and Translation, 28*(1), 307-328.

Through the online platform with expert and peer review task to improve EFL student writing

Wen-Chi Vivian Wu, Hsin-Yu Fan & Wan-Yu Liu
Department of English Language, Literature and Linguistics, Providence University, Taiwan
wcwu@pu.edu.tw

Abstract: Even though peer reviews and expert reviews are a commonly explored subject in the field of second/foreign language writing, few studies have investigated the function of follow-up discussion between feedback givers and receivers on the received feedback in the review process. In this study, both experts and English majors of EFL students in Taiwan served as reviewers to provide feedback and comments on the EFL students' writing. In view of this, this study aimed to examine to what degree discussion forum could improve the students' writing in terms of organization, logical development of ideas and style and quality of expression via an online platform designed by the researchers.
Keywords: second/foreign language writing, online platform, peer review writing tasks.

Introduction

Although studies over the past decade on ESL/EFL peer review writing tasks in the CALL context have been widely discussed, few of them have dealt with how a follow-up discussion between student writers and student reviewers over the received feedback would impact writing outcomes. Theoretically, receiving comprehensible feedback might further increase the amount of the feedback that the student writers adopt in revising their writing drafts. In view of this, the discussion forum in this study functioned to create a cyberspace for the student writers to clarify the unclear feedback generated by experts and cross-institutional peers until they fully comprehended the received feedback. Such technology enhanced language learning (TELL) has been widely used in collaborative learning, with students work together as members of a learning community, working on problem-solving tasks by questioning each other, discussing and sharing information.

Peer assessment, also called peer evaluation or peer review, is a process wherein peers evaluate each other's work, usually along with, or in place of, an expert (e.g. [12], [22]). Opportunities for students to revise their written work based on feedback have long been seen as critical to improving their writing skills ([10], [11], [16], [18], [21]). When students revise with feedback, they may not only improve the current piece but also develop general writing skills and learn to self-evaluate their writing [17].

Experts possess plentiful domain-specific knowledge that is highly organized, whereas novices have loosely organized knowledge [5]. Experts are faster in detecting problems because they need fewer cues [6], respond automatically rather than in a controlled way [13], access their own memories rather than focusing on the task at hand which is often external to their mental representations [3], use heuristics rather than exhaustive search [15], recognize data patterns [4], and use compiled rules and response plans [1].The expert and peer review (EPR) task in this study demanded the students to conduct two rounds of news reflective writing, with each round including the drafting and revising processes.

Both the purpose and the primary research question of this study was to examine to what extend students writing improved as a result of the feedback from experts and peer reviewers, in the writing areas of organization, logical development of ideas, and style and quality of expression.

Method

The participants in this study were forty-four English majors who enrolled in two intermediate-level writing courses at two distinct school sites in this study. A total of 22 experts, including the instructors of two writing courses, evaluated the students' writing in this study. Students were required to produce two rounds of writing reflections about the news. Each writing task used the drafting and revising process and each was assigned at least two experts and two cross-institutional peers for feedback randomly and anonymously via an online learning platform.

In order to examine the differences resulting from the expert and peer feedback, this study was designed with two reflective writing tasks for two different issue topics. The processes of the tasks were carried out by 1) providing prompts an topic suggestions for the responses during class time, 2) writing the reflection (first draft), 3) performance of the expert and peer reviews, 4)receiving the feedback, 5) revising the writing (revised draft) and finally 6) repeating the expert and peer reviews with the revised version draft. Steps two through six were completed in the online platform that was designed by one of the researchers.

Data analysis

These two writing tasks (including the reviewed feedback and the participants' drafts and writing products) were posted on the online platform. The data of this study consisted of the participants' two reflective writing assignments and the related feedback from the peers and expert reviewers. The analyses of the participants' writing improvement were according to the scoring rubric used in the peer and expert reviews. Each writing task were had results consisting of comments and reviewed feedback for the first draft and the final draft from two experts and two peers. The researchers; comments were analyzed and the difference of the drafts were compared to examine the effectiveness of the comments and feedback from the peers and experts. Hence, the methodology used in this study was qualitative.

Results

In this study, there were two rounds of writing tasks and each task had a first draft and a revised draft. In addition, each draft had four reviewers (two peers and two experts) to provide the comments and revision suggestions via the online platform. The researchers want to know the improvement between the first draft and revised draft from different aspects according to the writing scoring rubric which included organization, logical development of idea (content), and style and quality of expression.

The present study compared and analyzed the first and revised draft difference according to the criteria in the scoring rubric. The results showed that the participants' writing improved within the three categories in the scoring rubric. Most of the drafts had more improvement on the organization and quality of expression but less improvement on the logical development. However, in general, the authors of most of the drafts followed the comments and feedback from the experts and peers in writing the revised draft.

Discussion and Conclusions

The students' reactions provided not only useful insights into the course design and the learning project but also valuable suggestions for improving them. The students unanimously liked the course design and the class activity. The questionnaire findings showed that students clearly preferred the expert comments compared to their peers. The data helped the authors understand why. Despite the fact that students clearly preferred the experts' comments, there are some roles that can be identified from this study. From this study, the students' revisions were supposed to focus primarily on the organizations and style and quality of expression. Therefore, the experts and the peers w2ere instructed to focused more from these two parts.

This study provided a discussion forum that functioned to create a cyberspace for the student writers to clarify the unclear feedback received from the experts and cross-institutional peers until they fully comprehended the received feedback. However, there were still several limitations. First, some students received less than two items of feedback from two experts and two peers. Second, though this study lasted for a whole semester, it was still insufficient. Third, the website was self-designed and the actual process of logging in to edit and provide peer review was time consuming, which may be why, in second draft, some students received less than four comments. The authors recommend that future researchers might make the time longer for the study, and adding functions to the website.

Acknowledgements

This research was partially supported by the National Science Council in Taiwan through Grant NSC 101-2511-S-126-007.

References

[1] Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

[2] Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-148.

[3] Besnard, D., & Bastien-Toniazzo, M. (1999). Expert error in trouble-shooting: an exploratory study in electronics. *International Journal of Human Computer Studies*, 50, 391-405.

[4] Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.

[5] 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.

[6] Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 18, 257-273.

[7] Cho, K. & MacArthur, C. (2010) Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 238-338.

[8] Cho, K. & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline. *Computers and Education*, 48, 409-426.

[9] Davies, P. (2006). Peer assessment: Judging the Quality of Students' work by Comments rather than Marks, *Innovations in Education and Teaching International*, 43(1), 69-82.

[10] Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research*, 57, 481-506.

[11] Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Reading, writing, and language processing. Advances in applied psycholinguistics,* Vol. 2 (pp. 176-240). Cambridge, UK: Cambridge University Press.
[12] Rada R. & Hu K. (2002) Patterns in student–student commenting. *IEEE Transactions of Education* 45, 262– 267.
[13] Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. detection, search, and attention. *Psychological Review*, 84, 1-66.
[14] S. Graham, C. MacArthur, & J. Fitzgerald (Eds.), *Best practice in writing instruction* (pp. 141-162). New York: Guilford.
[15] Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
[16] MacArthur, C. A., Schwartz, S. S., & Graham, S. (1991). Effects of a reciprocal peer revision strategy in special education classrooms. *Learning Disabilities Research and Practice,* 6, 201-210.
[17] MacArthur, C. A. (2007). Best practice in teaching evaluation and revision. In
[18] McCutchen, D., Hull, G. A., & Smith, W. L. (1987). Editing strategies and error correction in basic writing. *Written Communication*, 4, 139-154.
[19] Miao, Y., Badger, R. & Zhen, Y (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, 179–200.
[20] Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms, *Language Teaching Research*, 8(1), 31–54.
[21] Sommers, N. I. (1980). Revision strategies of student writers and experienced writers. *College Composition and Communication,* 31, 378-387.
[22] Topping K.J. (1998) Peer assessment between students at colleges and universities. *Review of Educational Research,* 68, 249-276.
[23] Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing,* 9(2), 147-170
[24] Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *The Internet and Higher Education*, 11(3–4), 186–193.

Assessing English as a Second Language: From Classroom Data to a Competence-Based Open Learner Model

Susan BULL[a*], Barbara WASSON[bd], Michael KICKMEIER-RUST[c], Matthew D. JOHNSON[a] , Eli MOE[d], Cecilie HANSEN[d], Gerhilde MEISSL-EGGHART[e], Klaus HAMMERMUELLER[e]

[a] *Electronic, Electrical and Computer Engineering, University of Birmingham, UK*
[b]*Department of Information Science & Media Studies, University of Bergen, Norway*
[c]*Knowledge Management Institute, Technical University of Graz, Austria*
[d]*InterMedia, Uni Health, Uni Research AS, Norway*
[e]*Talkademy, Austria*
*s.bull@bham.ac.uk

Abstract: With the increase of ICT in classrooms comes much data that can be used for evidence-based assessment. We focus on harnessing and interpreting this data to empower teachers in formative assessment. We describe e-assessment of English as a Second Language and illustrate how we move from data collected in classroom activities, through an automated assessment method, to visualising competence levels in an open learner model.

Keywords: second language, evidence-based formative assessment, open learner model

Introduction

Today's classrooms may comprise a range of tools [1] producing much data that can be tapped to support formative assessment. There is a need for methods to capture and present the data so teachers can interpret and transform it to a meaningful form for students and themselves. We are developing such tools and methods for English as a Second Language. We describe moving from classroom data, through an automated assessment method, to an open learner model (OLM) for use by teachers to support their formative assessment work.

The Common European Framework of Reference for languages (CEFR) offers competence-based common reference levels in language learning [2]. These are based on language use and abilities (what students *can do*). CEFR is not detailed enough to design diagnostic testing items or define task difficulty, but is a useful starting point [3]. A similar focus is at the forefront of many current language courses and applications. In Norway, for example, a specified set of learning goals and competences must be integrated into English teaching in schools [4], and teachers plan activities to address the competences. Our OLM provides students and teachers with an overview of current competence levels, enabling better planning of teaching and student recognition of their learning. The approach also offers a way to facilitate teachers' classroom orchestration [5].

In this paper we introduce the OLM as a teacher and learner feedback tool, describe data available to teachers, how they can transform interaction data to include in a learner model, and outline how such data may be displayed to help raise awareness of competencies.

1. Open Learner Models and Classroom Data

A learner model is a representation of a user's skills and abilities, as inferred during their interactions, and enables a system to adapt to the needs of the individual. Increasingly, learner models are being opened to users as a means to help prompt learner reflection, help teacher planning and decision-making, etc. [6]. There are now also strong arguments for placing OLMs in the centre of contexts where there are multiple sources of data available for the learner model [7],[8],[9] since a variety of tools are in use in classrooms [1]. While an OLM can be likened to technology-based student progress and performance reports, rather than reporting progress, it *models* and externalises competences and skills. The problem in technology-rich classrooms is that data is not always available in a form that matches competence descriptors, and is often not able to pass data to a learner modelling service. We therefore offer teachers a means to transform activity data for an OLM.

Usually activity results are stored with scores or qualitative descriptors in an overview. An illustration of a teacher's spreadsheet recording results is given in Figure 1. This allows the teacher to see at a glance, how an individual is progressing in goal-related competences. As time advances and further items are added, we expect to see a shift towards good and excellent - as is indeed happening in this example. We aim to support teachers with an approach that is similar to their self-generated methods (e.g. Figure 1), or methods with which they are already familiar, but providing a focus on overviews of *current competences*. These can be presented through an OLM, so students may more readily recognise the importance of competences (rather than specific activities), and teachers can gain an overview they can act on in the classroom or in later planning.

	Goal 1	Goal 2	Goal 3	Goal 4	Goal 5
Level test	Starting competence	Starting competence	Starting competence	Good competence	Excellent competence
Test politics and democracy	Starting competence	Low competence	Good competence	Low competence	Good competence
Student comments					
Homework test	Excellent competence		Good competence	Low competence	Good competence
group work	Good competence	Good competence	Good competence	Low competence	Excellent competence
Student comments					
Self assessment	Excellent competence	Good competence	Excellent competence	Good competence	Good competence
Oral presentation		Excellent competence	Excellent competence		Excellent competence
5 hour test	Good competence	Excellent competence	Good competence	Excellent competen	Excellent competence

Figure 1: Example of a teacher's record of competencies that combines colour with text

This is in line with education policy in Europe moving from a focus on knowledge to a focus on competence. For example, in Norway, the learning goals and competences cover three areas: communication; language learning; culture, society and literature – each of which comprises sets of competences [4]. For example, two of the "communication" competences are that after four years of English students should be able to "read and understand the main content of texts on familiar topics" and "understand and use common English words and phrases related to daily life, leisure time and interests, both orally and in written form". Teachers plan how to incorporate appropriate activities into their classrooms to enable students to develop the competencies.

We illustrate with a set of activities aimed at 11-12 year-olds, including an electronic reading and listening test; interactions in a virtual world (Second Life); and an electronic self-assessment (from the European Language ePortfolio). Assessment methods, automatic and manual, are applied to data from these activities to determine achievement level for relevant competencies. The first activity, the online listening and reading test, has a mix of item types: multiple choice, click item, click text, click name, click word, move paragraph. Each item is weighted according to difficulty by professional test developers and these weights, along with student answers and other test item information, is used by ProNIFA (an automatic assessment method – see below), to generate competence levels for students taking the test before data is passed to the OLM. The second data set derives from activity within Second Life, and includes chat logs and video recordings of activity in 3D space. For example, from Second Life we get (i) a simple chat log file (time stamp, chatting person/entity, chat text); (ii) a set of competencies (CEFR skills [2] shown below), specified in a text file (number, id, initial probability that students have that skill, short description); and educator-defined (scripted) rules, which vary from very simple such as checking whether a certain entity writes a certain text; to more complicated, such as computing distances travelled in Second Life. ProNIFA parses the log files, checks whether the rules apply and updates the probabilities of the competencies (and the probability distribution over the competence states).

*(i) [07:21 UTC] <i>Teacher</i>Well done, Svein.
*
(ii) 001 CEFR#094 0,5 Listening A1
(iii) [Rule1] Who*=Teacher What=Well done, <NAME>. ASkills=1;2 AUpdate=0,2 LSkills=3 LUpdate=0,1*
***NB:** If the teacher says "Well done" and a name, the probabilities of skills 1 and 2 for learner <NAME> are increased by 0.2; and for skill 3, decreased by 0.1.*

The third data set is produced by student self-assessments. The European Language ePortfolio self-assessment grid was used to elicit self-assessment of speaking, listening and reading skills. Questions relate to various “can do's”, e.g. “I can understand simple, short greetings and expressions, such as hello, thank you or you are welcome” and students assess themselves between “I can do this a bit / quite well / very well”. The teacher interprets these data sets and the results are manually entered directly into the OLM – i.e. not all data needs to be transformed using ProNIFA.

As explained above, not all data is immediately available in competence form, and needs to be assessed either automatically or manually. ProNIFA (probabilistic non-invasive formative assessment) is a tool to support teachers in the assessment process. It establishes a user interface for data aggregation and analysis services and functions. Conceptually, the functions are based on Competence-based Knowledge Space Theory (CbKST), originally established by Doignon and Falmagne [10], a well-elaborated set-theoretic framework for addressing the relations amongst problems (e.g. test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. While the original idea considered performance (behaviour, e.g. solving a test item), extensions introduced a separation of observable performance and latent, unobservable competencies, which determine the performance [11]. CbKST assumes a finite set of more or less atomic competencies (in the sense of some well-defined, small scale descriptions of some sort of aptitude, ability, knowledge, or skill) and a prerequisite relation between those competencies. A prerequisite relation states that competency a is a prerequisite to acquire another competency b. If a person has competency b, we can assume they also have competency a. Because more than one set of competences can be a prerequisite for another (e.g., competency a or b are a prerequisite for acquiring competency c), prerequisite functions have been introduced, relying on and/or type relations. A person’s competence state is described by a subset of competencies. Due to the prerequisite relations between competencies, not all subsets are admissible competence states. Using interpretation and representation functions, the latent competencies are mapped to a set of tasks (or test items) covering a domain: mastering a task correctly is linked to a set of necessary competencies; not mastering a task is linked to a set of lacking competencies. This assignment induces a performance structure: the collection of all possible performance states. Recent versions of the conceptual framework are based on probabilistic mapping of competencies and performance indicators, accounting for lucky guesses or careless errors. This means, mastering a task correctly provides evidence for certain competencies and competence states, with a certain probability.

ProNIFA retrieves performance data and updates the probabilities of competencies and competence states in a domain. When a task is mastered, all associated competencies are increased in their probability, and failing in a task decreases the probabilities of associated competencies. A distinct feature in formative assessment is the multi-source approach. ProNIFA allows connecting the analysis features to a range of evidence sources (such as the listening and reading test or activity in a virtual world). The interpretation of the sources of evidence depends on a-priori specified and defined conditions, heuristics and rules, which associate sets of available and lacking competencies to achievements exhibited in the evidence. The idea is to define certain conditions or states in a given environment, for example: the direction and speed a learner is moving, following instructions in English in an adventure game, or a combination of correctly and incorrectly ticked multiple choice tasks in a regular online test. The specification of such states can occur in multiple forms, ranging from simply listing test items and the correctness of the items, to complex heuristics such as the degree to which an activity reduced the 'distance' to the solution in a problem solving process (technically this can be achieved by pseudo code scripting). The next step of this kind of planning/authoring is to assign a set of competencies that can be assumed available and also lacking when a certain state occurs. This assumption can be weighted with strength of the probability updates. In essence, this approach equals the conceptual framework of micro adaptivity (e.g. [12]). Figure 2 shows ProNIFA-analysed data from a Second Life activity (see Section 1). The resulting model built around atomic competencies and related probability distribution, is passed to an OLM platform as a next step to support teacher appraisal efforts (Figure 3).

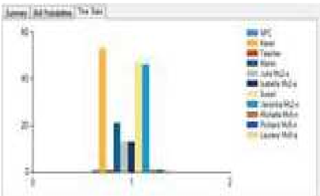


Figure 2: Screenshot of ProNIFA

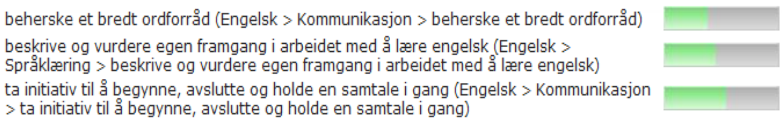


Figure 3: OLM skill meters

2. Competence Visualisation using an Open Learner Model

Using the easy-to-interpret ProNIFA display, teachers can add competency information to the OLM, as shown in Figure 4. They provide a numerical value for the model (by clicking on the stars) and may also include additional (non-modelled) feedback. The example shows competences in English according to the required learning goals and competences [4]. So, for example, if ProNIFA-analysis of recent Second Life logs indicates increased competence in some aspect of a student's learning, the teacher can easily update the OLM accordingly. This can happen alongside other, possibly automated input to the learner model, self-assessments, etc., if other activities are also ongoing. Thus, both teachers and students can flexibly use the OLM for formative assessment support.

Competency	Grade	Strengths (Text)	Impr
bruke teknisk og matematisk informasjon i kommunikasjon (Engelsk > Kommunikasjon)			
produsere tekster med sammensatt innhold i digitale medier (Engelsk > Kommunikasjon)			

Figure 4: Teacher updates to the OLM

As stated previously, information at this broad level of granularity is intended primarily to help gain a quick overview of students' competences which can, for example, be highly useful in classrooms where teachers are trying to manage classroom activities, give formative feedback, or update their teaching plan. In addition to the simple skill meters (Figure 3), student rankings by competence, and a table overview are available. Work is underway on word clouds – providing another way for teachers to quickly identify where to focus their attention [13]; and treemaps, which will allow drill-down to more detail, supporting more reflective formative assessment. These (and possibly other) learner model views will help teachers easily interpret the kind of information they already collect (e.g., Figure 1), but in a more immediately usable format (or, in the case of the planned treemaps, in a way that facilitates access to detail). Student use of the OLM, as well as promoting awareness of their learning [6], will help focus students on thinking in terms of competences (for English [4]), rather than activity-specific results (as in the example in Figure 1).

3. Summary

This paper has introduced a way to help teachers take the range of data now available about students, and transform it into a form that can be used in an OLM. This can help students note the importance of language competences, and help teachers' classroom orchestration.

Acknowledgements

The project is supported by the European Community (EC) under the Information Society Technology priority of the 7th Framework Programme for R&D, contract no 258114 NEXT-TELL. This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

References

[1] Richardson, W. (2009). *Blogs, Wikis, Podcasts, and Other Powerful Web Tools for Classrooms*, California: Corwin Press.

[2] Council of Europe (nd). The *Common European Framework of Reference for Language*s, http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf. Accessed 1 June 2012.

[3] Huhta, A. & Figueras, N. (2004). Using the CEF to Promote Language Learning through Diagnostic Testing, In K. Morrow (ed), *Insights from the Common European Framework*, Oxford: OUP, 65-76.

[4] Utdanningsdirektoratet. Læreplan i Engelsk, http://www.udir.no/Lareplaner/Grep/Modul/?gmid=0&gmi=166919 . Accessed 1 June 2012.

[5] Dillenbourg, P. & Jermann, P. (2010) Technology for Classroom Orchestration, In M. S. Khine & I. M. Saleh (eds). *New Science of Learning: Cognition, Computers and Collaboration in Education,* Berlin: Springer Verlag, 525-552.

[6] Bull, S. & Kay, J. (2007). Student Models that Invite the Learner In: The SMILI Open Learner Modelling Framework, *International Journal of Artificial Intelligence in Education* 17(2), 89-120.

[7] Morales, R., Van Labeke, N., Brna, P. & Chan, M.E. (2009). Open Learner Modelling as the Keystone of the Next Generation of Adaptive Learning Environments. In C. Mourlas & P. Germanakos (eds), *Intelligent User Interfaces*, Information Science Reference, 288-312, London: ICI Global.

[8] Mazzola, L. & Mazza, R. (2010). GVIS: A Facility for Adaptively Mashing Up and Representing Open Learner Models, In M. Wolpers et al (eds), *EC-TEL 2010*, Berlin: Springer Verlag, 554-559.

[9] Reimann, P., Bull, S., Halb, W. & Johnson, M. (2011). Design of a Computer-Assisted Assessment System for Classroom Formative Assessment, *CAF11*, IEEE.

[10] Doignon, J., & Falmagne, J. (1999). Knowledge Spaces. Berlin: Springer Verlag.

[11] Korossy, K. (1999). Modelling knowledge as competence and performance. In D. Albert & J. Lukas (Eds.), *Knowledge spaces: Theories, empirical research, and applications* Mahwah, NJ: LEA, 103–132

[12] Kickmeier-Rust, M. D., & Albert, D. (2011). Micro adaptivity: Protecting immersion in didactically adaptive digital educational games. *Journal of Computer Assisted Learning*, 26, 95-105.

[13] Reimann, P., Bull, S. & Ganesan, P. (in press). Supporting the Development of 21st Century Skills: Student Facilitation of Meetings and Data for Teachers, *TAPTA Workshop Proceedings*, EC-TEL 2012.

Voice-based Computer-mediated Communication (VCMC): An Exploratory Study on EFL Students' Perceptions and the First Language Use for Oral Practice

Mei-Jen Audrey SHIH*, Shih-Ting WANG & Jie Chi YANG
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*audreyshih32@cl.ncu.edu.tw

Abstract: A consideration to gradually boost students' willingness to communicate in the target language (L2) within the EFL instruction is seen to be an important preliminary to encourage EFL students to be able to have more English oral practice and product. While EFL students are in sharing and discussion in pairs or in groups, the intervention of the mother tongue (L1) use, as well as their background information in terms of confidence in English use and challenge in English use, might affect students' perceptions for English oral practice. In order to provide an efficient English speaking setting, a means of integrating synchronous computer-mediated communication in the instructional setting is employed to facilitate pairs and groups sharing. To this end, the aim of the study is to examine the relations of students' perceptions on oral practice within their background and the intervention of L1 use in the integrated EFL learning setting. 195 senior high school students were recruited after engaging in the semester-long English speaking class. Based on a prior study qualitatively conducting students' feedback in this learning process, a questionnaire, Perception for Synchronous Oral Computer-mediated Communication Questionnaire (PSOCMC) is developed. The findings demonstrate six factors revealed in terms of interactiveness, autonomy in English use, clarity of audio processing, intermediary of technology, instant support, and efficiency are analyzed and adopted for investigation. It shows the L1 use in students' sharing would affect them being autonomous in English use, and an indispensable role of confidence in English use in the EFL in-class environment.

Keywords: EFL, perception, mother tongue, synchronous, voice-based computer-mediated communication (VCMC)

Introduction

In English as a Foreign Language (EFL) learning environment, the four language skills are regarded as a means to effectively and distinctly train students to gain sufficient practice in English (Dekeyser, 2001). However, a happening reveals that although most of the EFL students are able to comprehend and produce English (L2) in written forms, it is hardly to see the equal output within oral structure. The difficulty of producing English speaking might further cause an incident of unwillingness for English communication. Then those circumstances might affect their willingness to communicate in English afterward. As a result, a means to promote students to be willing to communicate in English should be the primary consideration in the L2 instruction due to its profound influence of language learning (Léger & Storch, 2009). Also, being in the EFL setting, a condition of mother tongue (L1) use in the L2 affiliation of learning might unwittingly affect students' willingness in L2 speaking since all the interlocutors speak the same language, but are instructed to interact with others in L2 for learning (MacIntyre et al., 1998).

On the other hand, even though EFL students are willing to do English communication, a conventional learning environment setting might illustrates some problems. These problems could interrupt the opportunities and occasions for EFL students to not be able to actively engage in the English speaking setting and insufficient oral practice in class. To this end, an English speaking environment within ICT support in terms of a voice-based synchronous computer-mediated communication (CMC) is employed as an English enclosure to inspirit EFL students for English oral practice. Thus, the study aims to examine the relations of students' perceptions on oral practice within their background, as well as the intervention of L1 use underlying the context.

1. Literature review

1.1. Student perception toward English speaking

English speaking skill is an output process for students while learning the language. In EFL learning setting, students are used to being guided to receive the language knowledge in terms of reading and listening skills (Dekeyser, 1998, 2001), an inactive means to acquire the L2. Also, the outputting progress in terms of speaking skill is arranged as a follow-up L2 learning since without sufficient input knowledge, students rarely to make a proper oral production afterward. Due to the EFL circumstance, students merely are able to gain L2 knowledge in the classroom; and most of the class time is contributed for obtainment of knowledge (Cheon, 2003), yet, the time for appropriately generating output seems to be unwittingly overlooked. Meanwhile, an intervention of L1 or L2 use might alter and affect EFL students' perceptions on speaking performance and willingness (Carless, 2008; Mak, 2011; Storch & Aldosari, 2010). In order to encourage EFL students not only to acquire L2 knowledge but also to be able to yield more oral practice in class, their perceptions toward English speaking should be taken into a premier consideration to understand causes of their willingness to communicate to each other.

1.2. Oral computer-mediated communication

Engaging facilitation for classroom instruction based on pedagogical design tends to have a potential of benefiting efficient L2 learning (Zou, 2011). From dissimilarity on a conventional classroom environment, an oral or voice-based computer-mediated communication (OCMC) (VCMC) is implemented in English speaking classroom setting via a variety of instructional design (Alastuey, 2011; Ko, 2012; Yanguas, 2010). In terms of those recent studies on how synchronous VCMC effect students learning a foreign language in class, learning environments are set as audio or video CMC, and a comparison as the conventional face-to-face (FTF) interaction. The results are various; on the one hand, based on the VCMC support, it not only reveals a significant contribution to the foreign language acquisition but also provides competence of different interaction patterns as well as reducing defects that students might encounter in the conventional FTF speaking class (Alastuey, 2011; Yanguas, 2010). On the other hand, although the conventional FTF setting seems to have a weaker position than the VCMC setting for EFL oral development in class, the study suggested that the affiliated components, such as pedagogical design and strategy use should be generated into consideration as a whole perspective as well (Ko, 2012). The environment setting itself would not display a maximum of effective influence on oral foreign language learning without applicable elements.

2. Methodology

2.1. Participants

195 senior high school (grade 10) students who were male were recruited to participate in and give their feedback following a semester long English speaking class. They did not have any experience of a formal English speaking class at school prior instead of a general English class.

2.2. Instruments

The study is based on an extension of a prior study on three explored subscales of students' perceptions after one-month instruction of the synchronous English speaking class setting (Shih & Yang, 2012). The three subscales of perception were in terms of ability, interaction, and attitude. A questionnaire, Perception for Synchronous Oral Computer-mediated Communication Questionnaire (PSOCMC), was developed by the three subscale criteria, which contained 32 items for examination.

2.3. Procedure

An English speaking class was instructed during a semester long period. Via a theme-based English speaking instruction every other week, students were not only taught relevant vocabulary, phrases, and sentence structures but also asked to think aloud the guided questions towards the subject within the thematic context. Then, students were distributed randomly into pairing or a group for sharing their personal opinions via OCMC supports. Meanwhile, the instructor was able to join each pairing or group for assistance. The questionnaire was completed at the end of the semester to assemble students' perceived feedback to the OCMC-based English speaking class.

2.4. Data analysis

The questionnaire data was analyzed via the statistical analysis computer package SPSS within exploratory factor analysis which was able to illustrate and investigate the description of factors. The questionnaire of the study was based on five-point Likert scale to allow students to rate their agreement of scale on each item, in terms of 1= strongly disagree, 2= disagree, 3= neutrally, 4= agree, 5= strongly agree. The descriptive analysis was also engaged in to explicitly delineate the factors and relevant students' background information.

3. Results and discussion

The findings display a variety of factors revealed students' perceptions within distinct subcategories, and the interrelationship between their background information. Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) = .822 as well as Bartlett's test of sphericity= 2140.964 ($p=.000<.05$) were revealed from PSOCMC which showed an appropriateness for factor analysis. The results extracted six factors via principal axis factoring show the eigenvalues which were F1= 7.301, F2= 2.048, F3= 1.792, F4= 1.329, F5= 1.207, F6= 1.038, and relatively refer to 31.340, 7.832, 6.595, 4.519, 3.465, 2.927 % of variance. It leads to a significant validity of PSOCMC. The six extracted factors were characterized as F1: interactiveness, F2: autonomy in English use, F3: clarity of audio processing, F4: intermediary of technology, F5: instant support, F6: efficiency. The reliability of each factor displayed F1= .854, F2= .783, F3= .829, F4= .851, F5= .611, and F6= .751 which in total presents a significant reliability (.897>.07). Moreover, the mean score and standard deviation are described for the six factors and students' background information. First, it was accumulated as the representative of factor scores among each factor, which is F1 (M=4.10, SD=.724), F2 (M=4.06, SD=.675), F3 (M=4.23, SD=.687), F4 (M=3.55, SD=1.083), F5 (M=3.70, SD=.800), and F6 (M=3.60, SD=.721). Meanwhile, the background information of students, in terms of (1) mainly use Chinese to talk and discuss while sharing (Main_in_C) (M=2.16, SD=.992); (2) be confident to talk and discuss in English (Confidence) (M=3.16, SD=1.076); (3) feel challenged while talking and discussing in English (Challenge) (M=3.94, SD=1.085), were also investigated to refer to the correlation within PSOCMC.

Pearson correlation was utilized as a means to analyze the correlation among the six factors and the students' background information. Table 1 demonstrates the correlations between six factors and students' background information. Among six factors, it showed positive correlations between each factor (p<.01). Along with the students' background, it showed a negative correlation between Main_in_C to F2 (autonomy in English use) (p<.01) and F5 (instant support) (p<.05). Confidence revealed a positive correlation among the six factors; Challenge revealed a positive correlation between F4 (intermediary of technology) (p<.05). Within background information, it presented negative correlations between Confidence to Main_in_C (p<.01) and Challenge (p<.01).

Table 1: The correlations among six factors and students' background information

		F1	F2	F3	F4	F5	F6	Main_in_C	Confidence	Challenge
Pearson Correlation	F1	1								
	F2	.446**	1							
	F3	.466**	.409**	1						
	F4	.273**	.401**	.175*	1					
	F5	.313**	.463**	.377**	.259**	1				
	F6	.571**	.436**	.469**	.382**	.373**	1			
	Main_in_C	-.105	-.415**	-.133	-.106	-.146*	-.091	1		
	Confidence	.386**	.436**	.256**	.148*	.278**	.448**	-.291**	1	
	Challenge	-.019	.096	-.028	.176*	-.071	-.076	.109	-.323**	1

It is interesting to see that all the six factors explaining a positive interaction between either two factors among all. The interplay between the six factors of students' perceptions and their background information shows a number of different considerations. Students who mainly talked in Chinese while in pairs or group discussion seem not to be able to hold the autonomy in English use while discussing with other students. Even more, it also shows that the instant support, such as an immediate oral modification from the instructor or be able to check online dictionary for understanding does not favored much by those students who particularly did the discussion mainly in Chinese (L1) instead of in English (L2) to other students in pairs or group discussion. Due to the reluctant willingness to communicate to each other in L2, as well as the disfavor of instant supports within sharing, it might be worth of further investigating on the reasons of the unwillingness of communicating in L2 for oral sharing in class, as well as a concern on analyzing their discourse while in sharing to distinguish the cause of L1 use within L2 oral practice setting.

Meanwhile, on the one hand, as the students who have more confident in English discussion, it seems that they would hold more autonomous in English use, enjoy the interaction with others, and be more efficient in sharing their ideas in English via the facilitation of technology intermediary. On the other hand, the students who feel challenged to discuss in English in pairs or among groups incline to need the aid or intermediary in terms of headsets in this study to facilitate for a more comfortable English oral practice progress. The intervention of medium seems to lessen some concerns which allow students to be able to have less confrontation while doing oral practice in pairs and groups in L2. Thus, the confidence in English (L2) speaking ability appears to play an essential role while students are doing English oral discussion in class; as long as they are able to gain sufficient confidence in English use, they are more likely to neither use much Chinese (L1) in sharing nor feel it is challenges on using English (L2) to do the speaking practice in class setting.

4. Conclusion

It is a continuous progress for the instructors to consider an optimal EFL learning setting to inspire and boost EFL students' willingness to communicate in L2 for speaking instruction. According to this study, it tends to reveal that the intervention of L1 use in the EFL speaking environment causes students to be less autonomous in English speaking; it seems to be less necessary for students to adopt instant supports while the discussion. Last but not least, a hold of confidence in English oral use might need to be concerned precursory due to its significant influence on all perceived factors from EFL students, as well as the role of intermediary is able to ease off the uncomfortableness of English speaking in the EFL setting. In sum, the study suggests that the instructors should keep an awareness on students' perceptions and their willingness towards the speaking skill learning in EFL instruction since it is able to not only have a thorough understanding on students' perceived learning feedback but also allow the instructors to realize what the students might encounter while having oral practice in L2 and further adapt the speaking instructional design.

References

[1] Alastuey, M. C. B. (2011). Perceived benefits and drawbacks of synchronous voice-based computer-mediated communication in the foreign language classroom. *Computer-Assisted Language Learning, 24*(5), 419-432.

[2] Carless, D. (2008). Student use of the mother tongue in the task-based classroom. *ELT Journal, 62*(4), 331-338.

[3] Cheon, H. S. (2003). The viability of computer mediated communication in the Korean secondary EFL classroom. *Asian EFL Journal, 5*, 1.

[4] DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In *Focus on form in classroom second language acquisition*, 42-63. Ed. by C. J. Doughty & J. Williams. Cambridge: Cambridge University Press.

[5] Dekeyser, R. M. (2001). Automaticity and automatization. In *Cognition and second language instruction*, 125-151. Ed. by P. Robinson, Cambridge: Cambridge University Press.

[6] Ko, C.J. (2012). Can synchronous computer-mediated communication (CMC) help beginning-level foreign language learners speak? *Computer-Assisted Language Learning, 25*(3), 217-236.

[7] Léger, D. S., & Storch, N. (2009). Learners' perception and attitudes: Implications for willingness to communicate in an L2 classroom. *System, 37*, 269-285.

[8] MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation. *Modern Language Journal, 82*, 545-562.

[9] Mak, B. (2011). An exploration of speaking-in-class anxiety with Chinese ESL learners. *System, 39*(2), 202-214.

[10] Shih, M. J. A., & Yang, J. C. (2012). The Perception of Taiwanese EFL High School Students on Synchronous Oral Practice in ICT-based Classroom. *Proceedings of the 15th International CALL Research Conference (CALL 2012)*. Taichung, Taiwan.

[11] Storch, N., & Aldosari, A. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research, 14*(4), 355-375.

[12] Yanguas, Í. (2010). Oral computer-mediated interaction between L2 learners: It's about time! *Language Learning & Technology, 14*(3), 72-93.

[13] Zou, B. (2011). Teachers' support in using computers for developing students' listening and speaking skills in pre-sessional English courses. *Computer-Assisted Language Learning*, 1-17.

Affective States Arising from the Removal of Captioning Support in EFL Multimedia Environments

Aubrey Neil LEVERIDGE*, Jie Chi YANG
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*neill@lst.ncu.edu.tw

Abstract: This study explores affective states resulting from the removal or absence of captioning support in movies for the instruction of English as a foreign language listening comprehension. It aims to provide evidence of different affective states, which arise from the removal of this support, according to levels of proficiency. Prior knowledge of these states may enable EFL instructors and course designers to weigh the costs of exposing learners to this support and prepare learners for possible negative effects, avoiding any negative emotions, which may affect learning.

Keywords: Affective states, captioning EFL, listening comprehension

Introduction

Due to its many affordances, the use of multimedia as an instructional medium has become increasingly prevalent in foreign and second language (L2) classroom settings (Leveridge & Yang 2012; Sun & Dong, 2004). This is particularly true in the realm of L2 listening comprehension instruction. Multimedia integrated classrooms provide L2 instructors with various methods of presenting and supporting authentic L2 listening material (Vandergrift 2007). One such support is captioning, redundant visual text that matches spoken audio signals in the target language. On one hand, research investigating the use of captioning for L2 listening comprehension has indicated various potential benefits: in the facilitation of immediate understanding of L2 content (Robin 2007; Stewart & Pertusa 2004); enhancing vocabulary acquisition (Chai & Erlam 2008); and assisting L2 beginners when the audio is too fast (Robin 2007). On the other hand, research has also pointed out pitfalls associated with captioning support such as: transferability of skills from a learning context to a real-life context (Vandergrift, 2004); the impedance of schema building (Diao, Chandler, & Sweller, 2007); and their lack of compensation for more difficult vocabulary levels and high rates of speech (Danan, 2004). While inconsistencies in research findings exist, the majority of studies agree that at some point, as learners progress captioning support must eventually be eliminated as this support is not generally transferable to authentic L2 listening environments (Leveridge & Yang, 2012; Vandergrift, 2004).

Despite the extensive use of captioning as a tool to support and facilitate L2 listening comprehension, one of the major drawbacks is that learners may become reliant on the support (Vandergrift, 2004). A study by Leveridge & Yang (2012), which found wide variances in learner perceptions of captioning, addressed the issue of reliance by proposing a framework designed for the timely elimination of captioning, thus avoiding learner dependence. However, learner affective states, subsequent the removal of captioning support, is an area that has eluded empirical investigation. Moreover, captioning support, as previously mentioned, is prevalent in L2 listening comprehension instruction, thus, learners already reliant, may become frustrated, irritated, or upset, resulting from the removal of this support. These affective states may negatively impact learning, motivation, and willingness to continue learning. Accordingly, the current study aims to satisfy this research gap by examining learner perceptions concerning the removal of captioning support in L2 listening comprehension and the ensuing affective states. More specifically, this study strives to answer the following two research questions:

1. What affective states arise from the elimination of captioning support?
2. Is L2 listening proficiency a predictor of the types of affective states that arise from the elimination of captioning support?

The findings of this study may illuminate affective states arising from learner dependence on captioning support. Instructors and course designers may then weigh the costs of exposing learners to captioning, and prepare learners for possible negative affects. This in turn may circumvent negative emotions connected to the loss of captioning support, allowing the learner to remain motivated with a willingness to continue learning.

1. Literature Review

1.1 Definition and background of Captioning

Captions may be defined as redundant text that matches spoken audio signals and appears in the same language as the target audio. Captions are not to be confused with subtitles, which are textual versions of dialogue, but may not necessarily be in the same language as the audio. Captioning emerged in foreign language classrooms in the 1980s as a method of supporting listening comprehension that: increased learners' attention, reduced anxiety, increased motivation, and provided students with immediate verification of what was heard (Froehlich, 1988). Moreover, studies indicated that learners held positive attitudes toward captions (Froehlich, 1988). However, what did not appear in these studies was how the removal of captioning support affected the learners.

Then, in the late 1990's Guillory (1999) investigated the effects of different modes (audio only, keyword captions, and full captions), of captioning on learner comprehension, full captioning being the most beneficial. This created a rationale for the use of captions. Utilizing captions were deemed easier than listening alone (see Dio et al. 2007; Smidt & Hegelheimer 2004; Stewart & Pertusa 2004), learners became accustomed and tended to rely on captions, experiencing negative affective states when the captions were unavailable. More recently, studies have focused on how captions benefit learners at various levels of proficiency, the findings inconsistent (Winke, Gass, Sydorenko, 2010).

1.2 Affective States & Captioning

As previously mentioned, learners may become overly dependent on captioning support (Vandergrift, 2007). For example, Leveridge & Yang (2012) examined learner perceptions on reliance of captioning support and its subsequent removal and found variances between proficiency levels, i.e. low-proficiency learners became frustrated by the removal of captioning.

The question of transference to a real-life listening context remains. Thus begging the question: When learners, accustomed to captioning support, enter into real-life listening contexts, will they become frustrated and anxious due to the lack of support to which they have become accustomed?

2. Method

2.1 Participants

Participating in this study were 146 students from a high school in northern Taiwan. The students had an average age of 17 and were from three twelfth grade classes. All participants had been enrolled in English as a Foreign Language (EFL) classes for a minimum of five years. While all students were of a similar age and in the same grade, their individual English aptitudes varied considerably. Furthermore, captioning support was a familiar medium, as all students had prior instructional experience in captioned multimedia environments, as it was a standard instructional medium in their classes. All participants reported normal hearing ability and either normal or corrected to normal eyesight.

2.2 Instruments

The instruments employed in this study were as follows: 1) a multimedia system consisting of a computer, projector, screen, and public address system; 2) the intermediate General English Proficiency Test (GEPT); 3) an open-ended question; and 4) semi-structured interviews.

1. A multimedia system consisting of a computer capable of playing DVDs, projector, screen, and public address system were used to present the movies.
2. GEPT: The GEPT, developed in 1999 in Taiwan, provides individual evaluation of English language proficiency (Roever & Pan, 2008). The GEPT covers the testing of four major aspects of English language learning: listening, speaking, reading, and writing. Furthermore, each category is tested separately, thus making it possible to test only specific areas required, as in this case, listening comprehension. As two aims of the current study were to determine if the participants' L2 listening comprehension level were predictors of: 1) perceptions of captioning support; and 2) emotions arising from the addition and/or elimination of the support, thus the listening comprehension section of the GEPT was employed enabling categorization of participant results into one of three proficiency levels: low, intermediate, or high.
3. Open-ended question: An open-ended question was employed to gain insight into learner emotions arising from the use of captioning support. The open-ended question is as follows: How did you feel when the captions were removed or not presented?
4. Interviews: Semi-structured interviews were given to participants who answered the open-ended question in a stronger-than-average manner. The interviews were structured around the same open-ended question.

2.3 Procedure

First, the GEPT was administered and three categories were created: low (n=54), intermediate (n=46), and high (n=46). Secondly, the participants watched the three movies, 1 per week. The first movie was presented with captions, the second with captions for approximately the first half, followed by audio and picture only, and the third movie without captions, audio and picture only. Immediately following the third movie, participants were asked to recall their experience of viewing of the three movies; the different modes of captioning employed, and answer the open-ended question. Participants whom expressed more extreme affective states were interviewed using the semi-structured interview.

3. Results & Discussion

The following five underlying themes from the interview comments regarding the removal or absence of captioning support were identified, the first two themes emerged irrespective of proficiency, while the last three were particular to a proficiency level:

1. 40% of the participants (n=59) indicated that the removal of captioning support increased their mental workload and/or made them feel tired more quickly.
2. Of the participants, 72% (n=105) indicated a negative affective state, i.e. confusion/frustration (69%), insecurity (7%) or anxiety (24%) arising from the removal and/or absence of captioning support. However, the remaining 28% of the participants (n=41) reported they felt "okay" with the removal or absence of the captions or had no feelings on the matter. Of these participants, 6% (n=9) felt confident they could train listening ability when captions were not presented.
3. Low-level learners feel anxious (28%) when captions are removed.
4. The intermediate-level learners indicated equal amounts of both anxiety (17%) and normalcy (17%) when captioning was removed or not presented.
5. High-level learners indicated that they felt a loss of security (13%) when captions were removed.

The first two themes relate to the first research question, while the last three themes relate to the second research question.

The first theme to develop was related the affective states that may arise from the elimination of captioning support. Participants indicated that the absence of captioning support made them feel more tired as it seemed to increase in their cognitive load, or mental processing. Example: "Without captions, I have to concentrate very hard on the audio to gain comprehension." Example: "I felt sleepy. It is too difficult. Without captions, I must completely rely on the audio. It is tiring to listen to unfamiliar words."

The second theme to develop was confusion and frustration. Example: "I felt confused without captions. With captions, I have time to think. Without captions, I have no time to stop and think."

The third theme was more prevalent in both the low and high proficiency groups of learners: anxiety. Low proficiency learners indicated that they became anxious. Because the listening ability of this group is not as well established, as in the intermediate or high proficiency groups, the amount of cognitive processing needed, outweighed the resources available (i.e. vocabulary). Example: "When captions were removed, I had a difficult time understanding anything. I was disappointed because I cannot understand." High proficiency learners also indicated anxiety, but for different reasons compared to the low proficiency group. Anxiety was reported was because this group lost the mechanism they employed to check for correctness.

The fourth theme was an even distribution of anxiety and normalcy for the intermediate group that indicated a lower degree of anxiety, but an equal amount felt no difference between the non/captioned movies. Suggesting this group may have a stronger vocabulary, and improved listening skills. Even though not as reliant on the captions, some expressed that they would like the support to remain because they could more easily comprehend the movie content. Example: "Without captions, I will be able to train myself how to listen. However, I still want to see the captions in the movie."

The fifth theme was specific to the high-proficiency group: insecurity. Without captioning support, this group could not check listening accuracy, creating a sense of insecurity. Example: "I feel that there is no difference without captions but I cannot check my understanding and I can't get some of the words meanings without captions. [Without captioning] I feel a loss of security." One participant was reading the captions multiple

times as a way of feeling secure in what she had heard. Example: “I can read the captions twice. So I will have a greater sense of security in understanding what is being said.”
The results indicate that the removal or absence of captioning support does induce negative affective states and that proficiency is an influential factor in that various affective states are specific to particular levels.

Conclusion

By presenting learners with three videos, each with different modes of captioning support: the first with full captioning, the second with captioning for the first half only, and the third with no captioning, low-proficiency learners indicated anxious affective states, intermediate-proficiency learners specified either anxious states or contended that there was no difference, while high-proficiency learners indicated a loss of security.

Acknowledgements

We thank all the people who provided assistance for this study.

References

[1] Chai, J. & Erlam, R. (2008). The effect and the influence of the use of video and captions on second language learning. *New Zealand Studies in Applied Linguistics, 14*(2), 25-44.

[2] 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.

[3] Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta, 49*(1), 67–77.

[4] Diao, Y., Chandler, P. & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *American Journal of Psychology, 120*(2), 237-261.

[5] Froehlich, J. (1988). German videos with German subtitles: A new approach to listening comprehension development. *Die Unterrichtspraxis/Teaching German, 21*(2), 199-203.

[6] Guillory, H. G. (1999). The effect of keyword captions to authentic French video on learner comprehension. *CALICO Journal, 15*(1-3), 89-108.

[7] Leveridge, A. N. & Yang, J. C. (2012). Effect of Medium: A Conceptual Framework for the removal of Supporting Captions for EFL Listening Comprehension in Multimedia Instructional Delivery. Proceedings from the: *15th International CALL Research Conference.* Taichung, Taiwan.

[8] Robin, R. (2007). Commentary: learner-based listening and technological authenticity. *Language Learning & Technology, 11* (1), 109–115.

[9] Smidt, E. & Hegelheimer, V. (2004). Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition and strategy use. *Computer Assisted Language Learning, 17*(5), 517–556.

[10] Stewart, M. A. & Pertusa, I. (2004). Gains to foreign language while viewing target language closed-caption films. *Foreign Language Annals, 37* (3), 438–443.

[11] Sun, Y. & Dong, Q. (2004). An experiment on supporting children’s English vocabulary learning in multimedia context. *Computer Assisted Language Learning, 17* (2), 131-147.

[12] Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics, 24*, 3–25.

[13] Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40* (3), 191–210.

Development and Evaluation of ASR-based Speaking System to Support English Proverbs Learning

Yi Hsuan Wang, Shelley Shwu Ching Young, Chi-Ting Wei
Institute of Information Systems and Applications, National Tsing Hua University, Taiwan
annywang12345@hotmail.com

Abstract: In this study, the researchers from different background cooperatively completed the interdisciplinary research. The aim of this study was to explore a CALL system by combining the speech recognition (ASR) technology with English learning in the hope that the state-of-the-art technology could provide learners with more opportunities of bi-directional language learning in both of formal and informal English learning. The ASR-based CALL system was constructed which provided learners with the opportunities to practice English speaking with immediate diagnosis of their utterance and three levels of pedagogical feedback was embodied to assist learners enhancing English speaking. The quasi-experimental design was adopted in this study and a total of 32 Taiwanese students participated in the experiment. The results revealed that using the ASR-based CALL system in learning had positive effects on learners' speaking performance especially and students had positive attitude toward using the ASR-based system for language learning. Moreover, the three-level corrective feedback of the ASR-based CALL system could help students to improve language faults.

Keywords: CALL, English Learning, Automatic Speech Recognition, Feedback Design

Introduction

The advancement of computer assisted language learning (CALL) facilitates learning and teaching. The automatic speech recognition (ASR) based CALL system benefits learners by providing them with integrated learning stimulation and opportunities of enhancing learners' English speaking (Chiu, Liou, & Yeh, 2007; Chen, 2011). The ASR technology provides a flexible learning environment where learners obtain immediate evaluation of their English speaking and they can practice at any time that suits him. However, many issues in the feedback design of ASR-based CALL system requires further research (Chen, 2011). Related research has shown that in the speaking-practiced CALL system, learners tend to produce more accurate utterances when they are provided with corrective feedback, instead of the opportunity of to speak (Lyster & Ranta, 1997). Nevertheless, little study evaluated the feedback design of the ASR-based CALL system because the requirement to develop the ASR technology is technically demanding and challenging which needs researchers from different domain knowledge, both from technical expertise and technology learning theory background, to work together. Understanding the above research background, we attempted to construct and evaluate a CALL in which the ASR web-service and pedagogical corrective feedback were integrated to provide learners with a flexible learning environment for English speaking. The researchers from different backgrounds of computer science and learning technology worked together with the English educators to bring out the interdisciplinary study. The results will inform our following research project on how to well design and implement an ASR-based CALL system to address the actual needs of learners in Taiwan.

1. Literature Review

1.1 Computer assisted language learning

Over the last few decades years, there has been a dramatic increase in the variety of research on CALL including the mobile assisted language learning (Shield & Kukulska, 2008), tangible companions for learning conversation (Young, Wang & Jang, 2010) and the web-based voice recognition system for acquiring second language learning (Chiu, Liou, & Yeh, 2007). An intelligent Computer Assisted Language Learning (ICALL) involved the application of state-of-the-art computing technology such as automatic speech recognition technology (ASR). The timely evaluation based on learners' speaking performance forms an individual learning environment which provides each learner with step by step learning opportunities The ASR-based pronunciation learning system attracts more and more interests from researchers and English instructors and the future of applying ASR into CALL system in language learning is promising (Lu & Jaw, 2010).

1.2 The importance of feasibility feedback in language learning

The importance of providing learners with corrective feedback while using CALL system has been recognized from several studies (Neri et al., 2006; Chen, 2011). Learning feedback or immediate reward of learners' performance is necessary for learners to improve their ability (Hawkins, 1987). It could help learners learn effectively while providing them with corrective feedback rather than only giving them learning input (Lyster & Ranta, 1997). Corrective feedback could be presented in the implicit form or explicit form. According to the study results from Lyster and Ranta (1997), it revealed that teachers used implicit feedback, recast, in the language class the most to correct learners' language errors. Students pointed out that the explicit and segmentation feedback was more efficient than implicit ones for learners to acquire their language faults (Bigelow et al., 2006). While considering students' responses to various types of feedback, it is also important to consider the pedagogical purpose at the same time. An integrated feedback should be designed for the target learners and an explicit feedback may come after the implicit one. After reviewing the related literatures, we approached the research goal from the view point of associating ASR techniques with web-based learning concept and aimed to develop a CALL system which integrating multiple levels of corrective feedback to facilitate learning.

2. The ASR-based CALL System

The ASR-based CALL consisted of four modules, Expert module, Instruction module, Student module and ASR module, and each module had its specific role in language learning. The Expert module and Instruction module played the role of an English tutor that provided students with appropriate learning materials and guideline depending on the learner's speaking proficiency. The ASR module acted as an on-line learning partner and speaking evaluator that listens to each learner's English utterance and provided language feedback. The Student module acted as a recorder which recorded the learner's operation of the system and generated a language portfolio.

2.1 Corrective feedback design and presentation

Under this framework, both implicit and explicit feedbacks were provided and were organized into three levels. At the first level, which focused on providing implicit feedback, shows the learner's pronunciation score and the audio waveform. At the second level, which aimed to provide explicit feedback, contained a comment, an emoticon (a smiley face or crying face depending on the score), list of words that were pronounced accurately and inaccurately and an audio toolbar for replay of the learner's utterance. At the third level, demonstration of the accurate utterance, in both full sentence and single-word form were available. The learners could also play the sentence in normal and slow speed.

3. Methodology

To understand the learning effectiveness of using this system, we adopted the quasi-experiment. The control group practiced English speaking using single-level-feedback system. The design of single-level-feedback system was referred to the previous CALL study discussed in the literatures in which only the waveform diagram were presented as feedback to evaluate learners' speaking. On the contrary, the experimental group was given the three-level-feedback system which integrating the implicit and explicit elements into feedback presentation. A total of 32 seventh graders in the middle Taiwan participated in this study. The experimental group contained 16 students and the control group contained the other 16 students. The study used comparative test data and empirical experiments to report on the performance of learning English in the ASR-based CALL system with different levels of learning feedback. The learning topic chosen for this study was based on our previous research results (Wang & Young, 2012) that indicated English proverbs were one of the most desired learning contents for Taiwanese learners. Learners were required to take a pre-test and a post-test each contained an oral evaluation and several multiple-choice questions.

3.1 Research questions

The research questions of this study are as follows:
1. Can students achieve better English speaking after using the ASR-based CALL system?
2. Can the three-level feedback ASR-based CALL system promote learning effectiveness and motivation more effectively than the one-level feedback system?

4. Data Analysis

Following the research questions, the researchers analyzed and reported the study results based on quantitative data collected from the pre-test and post-test and the qualitative results from questionnaires and system login records. The Independent Samples T-test and Paired Samples T-test were conducted using the SPSS 12software package.

4.1 Learning effectiveness of experiment group (E.G.) and control group (C.G.)

Learning performance of the E.G. and C.G. are reported below. The independent sample t-test showed that there was no significant difference in the scores of pre-tests and post-tests for E.G and C.G. However, there were significant differences between the pre-test and post-test speaking scores for the E.G. and C.G. Base on the audio recordings of the participants, it was found students tended to pronounce the vocabulary correctly after practicing speaking with the ASR-based CALL system. Besides, their speaking fluency improved toward the end of our data collection period. For example, the proverb “Don’t put off till tomorrow what should be done today”, during the first few weeks, students tended to experience difficulty of pronouncing full sentences. They only could pronounce the first three words accurately but pronounce the last part inaccurately. However, after several weeks of practice, their speaking was more complete and they spoke with much more confidence.

4.2 Evaluation of the corrective Feedback

The ASR-based CALL system provided both explicit and implicit feedback to learners. Three levels of corrective feedbacks were provided for students in the E.G. and only one level of corrective feedback for the C.G. From the pre-test and post-tests data, the result indicated that both the E.G and CG’s speaking fluency was enhanced after eight weeks of practice. Considering the possible influence of students’ learning achievement, we further analyzed their speaking fluency according to their achievement level. It was found that the means of the speaking post-test scores were higher than the speaking pretest scores in E.G.. Furthermore, there were significant differences between pre-test and post-test scores from the results of Paired Sample T-test for the low-achievement learners in E.G. ($t= -4.79$, $P= 0.01$). It showed that the low-achievement learners in E.G. had improved significantly with the English speaking but not for the learners in C.G.. On the other hand, from the results of questionnaire, it showed there were only 40% of the students reflected they could read the information from the first-level of feedback. More than half of the students were confused and had problems understanding the audio waveform diagram. Furthermore, 53% of the students suggested that they did not know how to adjust their pronunciation just by using the first-level feedback information. Approximately 60% of students in the C.G. suggested the need for more details from the system.

5. Discussion and Conclusion

The study results showed that the ASR-based system efficiently improved Taiwanese students’’ English speaking. Research analyses indicated that students’ English speaking fluency and pronunciation were enhanced significantly. It was observed that some students experienced difficulty pronouncing English proverbs at the beginning but they were able to produce full sentences after several weeks. From the recorded system login data, it is found that learners became much more active in speaking English. Besides, the corrective language feedback in the proposed system was organized into three levels, containing both implicit and explicit forms. The research results in this study accorded with the earlier ones (Lyster & Ranta, 1997; Giuliani, Mich & Nardon, 2003; Chiu, Liou & Yeh, 2007). Students in the C.G. pointed out they had difficulty detecting their pronunciation errors from the level-one implicit feedback. The audio waveform was not helpful in providing useful information to the learners and in-depth analysis was needed and expected. On the other hand, the system provided students in E.G. with all three levels of learning support, feedback in implicit format firstly and then in explicit format. It indicated that corrective feedback in explicit format of immediate audio replay (recast) with the textual description and model pronunciation benefited learners the most especially for the low-achievement ones. The pre-test and post-test speaking scores for the low-achievement learners in the experiment group were improved through the integrated information of feedback. The textual feedback interpreted the waveform of students’ utterance more specific and the audio replay recast students' speaking utterance immediately so that they could read and hear the evaluation of the system accordingly to achieve better language learning.

5.1 Future work

It can be concluded from the current study that the ASR-based system is promising and the cooperation of the interdisciplinary study could provide tremendous value to language learning. In the future, the researchers will continue to improve the ASR-based CALL system and develop more interactive activity types such as interactive storytelling or speaking cloze test. Besides, we tend to extend the use of the system and evaluate the system for learners at different ages for addressing the actual needs of learners in Taiwan. Other research findings yielded from this series of ASR technology in language learning will be shared in the near future.

Acknowledgements

The research project is funded by the National Science Council in Taiwan, NSC 98-2511-S-007-003-MY3 and NSC99-2519-S007-003-MY3.

References

[1] Bigelow, M., Delmas, R., Hansen, K., and Tarone, E. (2006). Literacy and the Processing of oral recasts in SLA. TESOL Quarterly, 40 (4), 665-685.

[2] Chen, H., N. (2011). Developing and evaluating an oral skills training site supported by automatic speech recognition technology.ReCALL Journal, 23(1), 59-78.

[3] Chiu, T. L., Liou, H. C., & Yeh, Y. (2007). A study of Web-based oral activities enhanced by automatic speech recognition for EFL college learning. Computer Assisted Language Learning, 20, 209-233.

[4] Hawkins, E. (1987). Modern Languages in the Curriculum, 2nd Edition. Cambridge, Cambridge University Press.

[5] Lu, J.h. & Jaw, G.Y. (2010). Applying MyET Oral Practice Database to Upgrade College Students' English Speaking Abilit. Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, USA, 1111-1115.

[6] Lyster, R. & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. Studies in Second Language Acquisition, 19, 37-66.

[7] Neri, A., Cucchiarini, C. & Strik, H. (2002). Feedback in computer-assisted pronunciation training: Technology push or demand pull? Proceedings of ICSLP 2002, Denver, Center for Spoken Language Research, UK, 1209-1212.

[8] Young, S. S.-C., Wang, Y. S., & Jang, J. S. R. (2010). Exploring Perceptions of Integrating Tangible Learning Companions in Learning English Conversation. British Journal of Educational Technology, 41(5), 78–83

[9] Wang, Y. S., Jang & Young, S. S.-C. (2012). Exploring young and adult learners' perceptions of the corrective feedback in the ASR-based CALL system. British Journal of Educational Technology, 43(3), 77-80.

The metacognitive development of adult language learners in self-directed settings: A look at a distance learning course

Claudia Guidi
Assistant Professor of Spanish and Assistant Academic Director of Foreign Languages, University of Maryland University College, USA
claudia.guidi@gmail.com

Abstract: This paper presents a research design for an empirical study on the metacognitive development of adult language learners in a distance learning context. The paper describes a mixed methods design that will be used during the implementation phase of the project to elicit the experiences, thoughts, and feelings of learners during the process of learning, and to examine how the verbalization of experience, thought, and feeling drives metacognitive development.

Keywords: metacognition, metacognitive development, learner autonomy, distance learning, self-directed learning

Introduction

This paper presents a research design for a study on the metacognitive development of adult language learners learning Spanish online. The issue of how metacognition develops has been addressed in studies of language learners in both *enhanced settings* and *non-enhanced settings*. *Enhanced-settings* are learning contexts that include pedagogical interventions in the form of advising and strategy training, while *non-enhanced settings* do not include such interventions (e.g., learning outside the classroom, immersion, distance learning). Does the enhanced setting make a difference for learners' metacognitive development? And can learners in non-enhanced settings still be successful? These important questions are still open. This design of an exploratory study aims at better understanding the metacognitive development of learners in non-enhanced settings.

1. Literature Review

Prior studies on metacognitive development in self-directed settings generally support a constructivist perspective on learning. For example, studies that found positive metacognitive development [2] [6] [11] describe the process as disruptive, not linear; a process in which uncertainty (cognitive, emotional, or both) plays a critical role in moving the learning forward. Prior studies suggest that some sort of intervention, that is, some 'cueing' or 'prompting' of behavior through advising and strategy training is necessary to jumpstart metacognitive development. This finding emerges from the observation that studies of enhanced settings showed positive metacognitive development [2] [6], while some studies of non-enhanced settings [1] [3] reported lack of development. This finding would lead one to conclude that little metacognitive development occurs in non-enhanced settings. However, this finding should be taken with caution because (1) non-enhanced studies tend not to collect longitudinal data, causing internal validity issues, and (2) one study of a non-enhanced setting where these data were collected [11] did show development. At first glance, studies seem to show that, with the exception of 'expert' learners, some explicit intervention in the form of learning strategy training and advising is necessary for metacognitive development, as learners tended not to develop much metacognition in non-enhanced settings [1] [3]. However, this conclusion should be taken with caution due to the one-shot, single instrument, methodologies used in non-enhanced studies. Since the study on distance learning [11] did show improvement of metacognition in a non-enhanced setting, the issue of the metacognitive development of adult language learners in settings where metacognitive interventions are not available (i.e., non-enhanced settings) should be revisited.

2. Research questions

In light of the rationale discussed above, the study will address the following questions: (1) What are the uses, thoughts, and emotions of adult language learners in non-enhanced self-directed settings, (2) What is the role of emotion in the learning process of adult language learners in non-enhanced self-directed settings? Are there emotional stages, (3) Would adult language learners in non-enhanced self-directed settings show metacognitive development? If so, what factors would contribute to this development?

3. Methodology

3.1. Participants

A non-enhanced setting is operationalized in this study as an 8-week distance language learning course with no metacognitive enhancements such as advising or strategy training. The participants in this study are adult language learners enrolled in such a course. They are female and male and from a variety of majors. They are speakers of English learning Spanish as a foreign language. Demographic and background data will be collected via a questionnaire one week prior to the course. Upon obtaining permission from the course instructor, the researcher will visit the courses to solicit students to participate on a volunteer basis.

The study explores the behaviors of adult language learners enrolled in an 8-week non-enhanced self-directed university online course. The design consists of mixed-methods including a pre-and post- beliefs questionnaire, weekly collected reflective journals, and exit interviews.

3.2 Pre- and post-course beliefs questionnaire

Upon receiving written consent to participate from each participant, the researcher will administer a 10-item Likert scale beliefs questionnaire adapted from a prior study [2]. This pre-and post-beliefs questionnaire aims at eliciting learners' own perceptions of their metacognitive knowledge. This tool is used to establish a knowledge baseline at the onset of the study, and capture any increases in perceived metacognitive development at the end of the 8 weeks experience.

3.3 Journals

The researcher will ask the participants to keep a reflective journal to record their behaviors (e.g., any activities, strategies they do in relation to language learning), as well as their thoughts and feelings as they go through the learning experience. There will be no restrictions on format or length of the entries in the journals, but learners will be asked to start each new entry with the date. Journals will be submitted to the researcher at the end of each week during the 8 weeks of the course. Weekly submissions are a way of reminding students to keep working on their journals on a regular basis.

3.4 Interviews

The critical incident technique [8] will be used to elicit how individuals make sense of events in their environment in the process of learning. Critical incidents will be collected by the researcher through face-to-face video-taped interviews with a convenient sample of participants. Content analysis will be applied to the interview data and participants will be added until the data reaches the point of saturation [7] [10].

4. Significance and potential impact of the study

Language learning studies on metacognition tend to overlook the role that emotional states play in the learning process. This research design looks at learning holistically, as an interplay of thought, feeling, and action, and attempts to gain a better understanding of learning as a process. Furthermore, this design improves the research methodology of studies in metacognition by collecting critical incidents to gain an in-depth understanding of what works and what does not work in the process of learning how to learn in a distance learning course.

References

[1] Bordonaro, K. (2006). Language learning in the library: An exploratory study of ESL students, Journal of Academic Librarianship, 32(5) 518-526.

[2] Cotterall, S. & Murray, G. (2009). Enhancing metacognitive knowledge: Structure, affordances and self, System, 37(1), 34-45

[3] Inozu, J., Sahinkarakas, S. & Yumru, H. (2010). The nature of language learning experiences beyond the classroom and its learning outcomes. US-China Foreign Language, 8(1), 14-21.

[4] Kuhlthau, C. C. (2004). Chapter 3: The information search process. In Seeking meaning (2nd ed., pp. 29-52). Westport, CT: Libraries Unlimited, Inc.

[5] Lewis, T. & Vialleton, E. (2011). The notions of control and consciousness in learner autonomy and self-regulated learning: a comparison and critique. Innovation in Language Learning & Teaching, 5(20), 205-219.

[6] Navarro, D. & Thornton, K. (2011). Investigating the relationship between belief and action in self-directed language learning, System, 39(3), 290-301.

[7] Powell, R. R. and L. Silipigni Connaway (2004), Qualitative research methods, in Basic research methods for librarians (4th ed.), Westport, CT: Libraries Unlimited.

[8] Radford, M. L. (2006). The critical incident technique and the qualitative evaluation of the Connecting Libraries and Schools project, Library Trends, 55(1), 46-64.

[9] Rivers, W. P. (2001). Autonomy at All Costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners, Modern Language Journal, 85(2), 279-290.

[10] Westbrook, L. (2010), Qualitative research methods, in L. Connaway and R. R. Powell, Basic Research Methods for librarians, Westport, CT: ABC-CLIO (5th ed.). Retrieved from http://site.ebrary.com/lib/umd/Doc?id=10437271&ppg=220

[11] White, C. J. (1999), The metacognitive knowledge of distance learners, Open Learning, Nov (37-46).

TPACK-in-Action: An Innovative Model to Help English Teachers Integrate CALL

Shu-Ju Diana TAI[a*] & Hsueh-Hua CHUANG
[a] *Iowa State University, USA*
[b] *National Sun Yat-Sen University, Taiwan*
[*]shujutai@gmail.com

Abstract: In this paper, we propose a TPACK-in-Action model to guide the design of the CALL Workshop in helping inservice English teachers develop their TPACK proficiency and integrate CALL in their classrooms. Following the five steps design: (1) Modeling; (2) Analysis; (3) Demonstration; (4) Application; and (5) Reflection, the workshop centers at helping teachers learn to integrate CALL by doing CALL within the TPACK framework [22]. In other words, the workshop aims for teachers to walk away knowing how to teach with technology with pedagogical decisions as well as to transfer what they have learned in the workshop to their teaching in classrooms

Keywords: CALL, teacher education, professional development, TPACK-in-Action

Introduction

It has become increasingly clear that the future of CALL is closely tied to language teacher education because teachers are key to the realization of its educational potentials [13]. They are the gatekeepers, determining whether or what technologies enter the classroom and how they are used [6]. Moreover, rapidly changing CALL technology and the widening scope of technology-enhanced environment place more weight on the significance of teachers' perceptions and actions in order to successfully implement technology in the L2 classroom [8]. In other words, teachers "need to know why they do what they do" (p. 11) [15] in the technology enhanced L2 teaching and learning environment, which lends itself to the importance of teacher education in CALL.

1. Literature Review

Given the significance of CALL in language teacher education, one important factor related to CALL teacher education is the content and approach employed to deliver the training course(s). Extending the notion about the approach adopted to conduct the CALL education, Levy [19] proposed that a CALL course should be looked at from a more holistic view rather than whether or not teachers are trained to be computer experts. In addition, Chapelle and Hegelheimer [3] stressed the need to clarify the key competences of language teachers in the 21^{st} century to "effectively and critically engage in technology-related teaching issues . . . within a world that is decisively supported and interconnected by technology" (p. 300). In responding to this notion, Peters [23] specifically identified that there is a need to help English teachers learn to integrate technology effectively in the classroom rather than be technical or technology experts. Moreover, many CALL researchers have made suggestions that language teachers should develop a variety of competences and knowledge in order to integrate technology effectively within their classroom settings [5, 9, 14, 20]. The TPACK framework [22] advocates the incorporation of the three fundamental knowledge types among teachers, content knowledge (CK), pedagogy knowledge (PK), and technology knowledge (TK) and emphasizes the importance of the interactions and the complexities among all three basic knowledge domains. In other words, the TPACK framework goes beyond looking at these three knowledge domains in isolation but examines the new kinds of knowledge that gather at the intersections between and among the three domains, pedagogical content knowledge (PCK), technological content knowledge (TCK), technological pedagogical knowledge (TPK), and TPACK [18]. In sum, a well-organized and well-prepared training course should help prepare teachers to obtain and develop enough knowledge relating to language teaching technology solutions [11, 17] i.e., TPACK competency. Moreover, teachers need experiences with consistent modeling of effective use of technology in order to become familiar enough with the use of different forms of technology and see the wider range of affordances available [1, 2, 12]. Therefore, this study proposes that the CALL intervention be based on an innovative approach, TPACK-in-Action, to help English teachers develop their TPACK competency and integrate technology in their teaching.

2. The CALL Workshop: TPACK-in-Action

Adapted from the TPACK-in-Practice model [7], the TPACK-in-Action model proposes that a CALL teacher training follows the five steps: [*Modeling*] The CALL workshop starts with modeling an activity to situate teachers in context as Chapelle [4] noted, "The way that students will learn to do applied linguistics with technology is by learning applied linguistics through technology" (p. 31). During this step, teachers have the opportunities to see how a CALL activity/task can be implemented in classrooms [12, 16, 17]. [*Analysis*] Acknowledging the notion that teachers need to know not only how to use technology but also understand why they are doing so [3], an analysis of the modeled activity within the TPACK framework will be implemented to help teachers understand the rationale behind the choice of tools and pedagogy incorporated into the content in the modeled activity. [*Demonstration*] Through demonstration, teachers will learn about features of tools incorporated in the modeled activity. Moreover, alternative tools will also be introduced to allow teachers more options to meet learners' needs. [*Application*] Teachers will apply what they have learned, i.e., creating a lesson plan based on their curriculum and teaching it. [*Reflection*] Teachers will take the opportunity to reflect on their learning and development. As research noted, one of the most important factors that fosters teachers' professional development is reflective practice because critical reflection raises teachers' awareness about teaching, enables deeper understanding, and triggers positive changes [10, 21].

3. Conclusion

Different from the traditional techno-centric technology approach in which teachers developed technology knowledge, the TPACK-in-Action CALL Workshop aims to situate technologies in contexts, in which teachers learn how to incorporate technology to suit the instructional and learning needs. In addition, the ample opportunities to engage in hands-on activities, i.e., to integrate CALL by doing CALL as Chapelle [4] noted, contribute to teachers' greater confidence in their instructional ability and lead to more successful teaching experiences [24]. The goal of the TPACK-in-Action workshop is for teachers to walk away with the ability to teach with technology with pedagogical decisions and to transfer what they have learned in the workshop to their teaching in classrooms.

References

[1] Bird, T., & Rosaen, C. L. (2005). Providing authentic contexts for learning information technology in teacher preparation. Journal of Teaching and Teacher Education (JTATE), 13(2), 211-23.

[2] Brook, C. & Oliver R . (2005). Online learning communities and community development. In P. Kommers & G. Richards (Eds). Proceedings of Ed-Media 2005, World Conference on Educational Multimedia, Hypermedia & Telecommunications (pp. 2518-2525). Norfolk, USA: Association for the Advancement of Computing in Education.

[3] Chapelle, C. & Hegelheimer, V. (2004). The English language teacher in the 21st century. In S. Fotos & C. Browne (Eds.), New Perspectives on CALL for Second Language Classrooms (pp. 299-316). Mahwah, NJ: Lawrence Erlbaum.

[4] Chapelle, C. A. (2003). English Language Learning and Technology. Philadelphia, PA: John Benjamins Publishing Company.

[5] Chishome, I. M. & Beckett, E. C. (2003). Teacher preparation for equitable access through the integration of TESOL standards, multiple intelligences and technology. Technology, Pedagogy and Education, 12(2), 249-275.

[6] Cuban, L. (1986). Teachers and Machines: The Classroom Use of Technology Since 1920. New York, NY: Teacher College Press.

[7] Figg, C. & Jaipal, K. (2012). TPACK-in-Practice: Developing 21st Century Teacher Knowledge. In P. Resta (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2012* (pp. 4683-4689). Chesapeake, VA: AACE.

[8] Godwin-Jones, R. (2002). Technology for prospective language teachers. Language Learning & Technology, 6(3), 10-14.

[9] Hegelheimer, V. (2006). When the technology course is required. In P. Hubbard & M. Levy (Eds.), Teacher Education in CALL (pp. 117-133). Philadelphia, PA: John Benjamins Publishing Company.

[10] Ho, B. & Richards, J. C., (1993). Reflective thinking through teacher journal writing myths and realities. Prospect, 8, pp. 7-24.

[11] Hong, K. H. (2010). CALL teacher education as an impetus for L2 teachers in integrating technology. ReCALL, 22, 53-69.

[12] Hoven, D. (2007). The affordances of technology for student teachers to shape their teacher education experience. In M. A. Kassen, R. Z. Lavine, K. Murphy-Judy & M. Peters (Eds.), Preparing and Developing Technology-Proficient L2 Teachers (pp. 133-163). San Marcos, TX: CALICO.

[13] Hubbard, P. (2008). CALL and the future of language teacher education. Calico Journal, 25(2),175-188.

[14] Hubbard, P. (2004). Learner training for effective of CALL. In S. Fotos & C. Browne (Eds.), New Perspectives on CALL for Second Language Classrooms (pp. 45-67). Mahwah, NJ: Lawrence Erlbaum.

[15] Hubbard, P. & Levy, M. (2006). The scope of CALL education. In P. Hubbard & M. Levy (Eds.), Teacher Education in CALL (pp. 2-20). Philadelphia, PA: John Benjamins Publishing Company.

[16] Hughes, J. (2005). The role of teacher knowledge and learning experiences in forming technology-integrated pedagogy. Journal of Technology and Teacher Education, 13(2), 277-302.

[17] Kessler, G. & Plakans, L. (2008). Does teachers' confidence with CALL equal innovative and integrated use? Computer Assisted Language Learning, 21(3), 269-282.

[18] Koehler, M. and Mishra, P. (2008). Introducing TPCK. In: AACTE Committee on Innovation Technology (Eds.). Handbook of Technological Pedagogical Content Knowledge (TPCK) for Educators (pp. 3-29). London: Routledge Taylor and Francis Group.

[19] Levy, M. (1997). A rationale for teacher education and CALL: The holistic view and its implications. Computer and the Humanities, 30, 293-302.

[20] Levy, M., & Stockwell, G. (2006). CALL Dimensions: Options and Issues in Computer-Assisted Language Learning. New Jersey: Lawrence Erlbaum Associates.

[21] Liou, H. (2001). Reflective practice in a pre-service teacher education process for high school English teachers in Taiwan. System, 29(2), 197-208.

[22] Mishra, P. & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for integrating technology in teacher knowledge. Teachers College Record, 108(6), 1017-1054.

[23] Peters, M. (2006). Developing computer competencies for pre-service language teachers: Is one course enough? In P. Hubbard & M. Levy (Eds.), Teacher Education in CALL (pp. 153-165). Philadelphia, PA: John Benjamins Publishing Company.

[24] Ross, J. A., Hogaboam-Gray, A. & Hannay, L. (1999) Predictors of teachers' confidence in their ability to implement computer-based instruction. Journal of Educational Computing Research, 21, (1), 75-97.

Chances and Challenges of Using WebQuest in Academic Reading

Min-Tun CHUANG* & Kuan-Yin CHEN
Department of Foreign Languages, National Chiayi University, Taiwan
*mtchuang@mail.ncyu.edu.tw

Abstract: WebQuest, embedded with task-based learning is a scaffold learning structure in which the instructors provide various online resources to help learners complete designed tasks and study the focus language at the same time. However, there are few empirical researches exploring the students' perceptions toward WebQuest, especially in higher education. Therefore, to fill in the gap, this qualitative research attempted to examine how English major college students feel about learning academic reading through a researcher-designed WebQuest. A total 86 sophomore English major students were involved in this study. The findings indicated that through the process of step- by- step guidance provided in the WebQuest, critical thinking could be elicited while some of the challenges were also found in the study. The major contribution of this study is to provide some insights on the application of integrating WebQuest into academic reading, and to understand students' voice toward the use of WebQuest.

Keywords: WebQuest, Academic Reading, Task-Based Learning

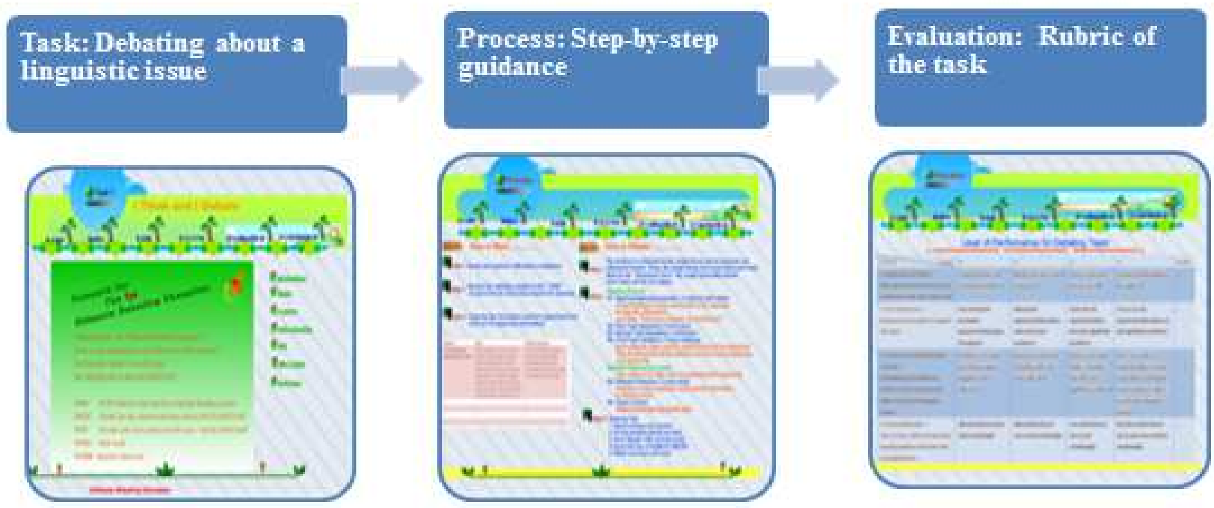
Introduction

With the advantages and easy accesses to the Internet, the EFL or ESL teachers use computers and the Internet more and more frequently in the classroom to make the teaching-learning process more effective. Oliver (2000) indicated that exposure to information through Web sites can provide students with environments that support inquiry-based and constructivist learning. However, a large-scale study result indicated that there was no effectiveness of access to the Internet in school on students' achievement (Trotter, 2002). The finding of this study revealed the fact that merely expose to the Internet is not enough to improve students' learning. Therefore, WebQuest, an inquiry-oriented instructional template embedded with learning tasks is necessary for improving students' learning. WebQuest, through which students interact with resources on the Internet, can elicit students' higher-order thinking by analyzing the data for completing the task embedded in the WebQuest (Dodge, 1995). Bearing with these in mind, a researcher-designed WebQuest was planned and implemented in a college-level academic reading instruction. It is expected that through the step by step instruction of WebQuest, students would learn how to synthesize the data and to complete a task related to a linguistic issue.

1. Methodology and Research Procedure

Eighty-six sophomore English major students who were attending academic reading at National Chiayi University were invited to participate in this research-designed WebQuest reading course. Students' novice status to WebQuest is useful to explore their voice toward Internet-based learning. Students completed the WebQuest task over a four week period of time. Firs, students were introduced to the WebQuest (see Figure 1). Addresses for topic-related Web sites were provided, but learners were not restricted to these sites. Then, after going through these data, students needed to organize the findings and then debated over the topic "Whether the earlier the better in language learning", and the rubric would be provided in the evaluation section of WebQuest. Third, learners ought to write down their reflection toward the whole process. Last, the researcher analyzed the written forms of reflection, and used it to describe the findings.

Figure1. The WebQuest Design (http://student.ncyu.edu.tw/~s0991015)



2. Results and Discussions

The spirit of WebQuest and task-based learning were the rubric used to examine students' voice and learning process. In terms of information providing in the WebQuest, most students found that WebQuest provided enough information to them. Student were not passively received from the teacher, they took over the learning process such as searching for extending information (Shiuan). Nowadays, the educators put more efforts on cultivating and reinforcing student's higher levels of cognition. In this study, students needed the ability to distinguish what kind of information (Yi) they needed and synthesize the data they found to debate. WebQuest is also embedded with the task-based learning. It met students' interest (Ling). Furthermore, students needed to use the target language to complete the task-debate. Most of all, students even evaluated their own learning process and result (Shiuan).

However, there were some challenges for the researchers to reflect. The WebQuest designer might create some space for learners to exchange their information and discuss. More explanation about the elements and function of WebQuest should be provided. Furthermore, due to the limitation of the task and time, not each one of the class had the chance to speak out his arguments (Chen-Hua).

To conclude, the results could response to previous researches. WebQuests have four constructs: Critical thinking, knowledge application, social skills, and scaffolded learning (Zheng, 2005). The results also received positive feedback from students toward the use of WebQuest in teaching reading (Tuan, 2011).

3. Conclusion

This study examines the nature of English major students' experiences with the process of learning through WebQuest. On the practical level, the researcher carefully designed a task-based WebQuest and combined it into academic reading. On the theoretical level, the spirit of WebQuest, namely, using learners' time well, supporting learners' thinking and levels of analysis, synthesis and evaluation, and task-based learning add more knowledge to this study.

References

[1] Dodge, B. (1995). Some thoughts about WebQuests. The Distances Educator, 1(3), 12-15

[2] Oliver, K. (2000). Methods for developing constructivist learning on the Web. *Educational Technology,* 40(G), 5-16.

[3] Trotter, A. (2002) Internet access has no impact. Education Week. Retrieved March 11th, 2012, from http://edweek.com;ew;ewstory.cfm?slug=01internet.h22.

[4] Tuan, L. (2011). Teaching Reading through WebQuest. *Journal of Language Teaching and Research*, 2(3), 664-673.

[5] Zheng, R., Stucky, B., McAlack, M., Menchana, M. & Stoddart. (2005). WebQuest learning as perceived by higher education learners. *TechTrends,* 49, 41–49.

Exploring the effects of using mixed-modality vocabulary learning strategy on vocabulary retention

Fang-Chuan OU YANG[a*] **& Kuo-Yu LIU**[b]
[a]*Department of Information Management, Chien Hsin University of Science and Technology, Chung-Li, Taiwan*
[b]*Department of Computer Science and Communication Engineering, Providence University, Taichung, Taiwan*
[*]ouyang@uch.edu.tw

Abstract: This study investigated the effects of mixed-modality vocabulary learning strategy use on vocabulary retention. To fulfill the purposes of this study, a CALL system called MyEVA was employed. A within-subjects experiment was conducted to examine the effects of three vocabulary learning tools: MyEVA, internet dictionary, and traditional paper-based dictionary. The findings indicate that mixed-modality with preference strategy setting stimulates the greatest vocabulary acquisition and the best retention for L2 students.

Keywords: mixed-modality, vocabulary learning strategy, vocabulary retention

Introduction

College students should have at least 4000 words to understand and read English textbooks (Hu and Nation, 2001). However, many college students in Taiwan only have less than 2000 words, and therefore have bottlenecks to comprehend their English textbooks (Huang, 2004). Oxford (1990) claimed that vocabulary learning strategy (VLS) can help English learners to recollect plenty of words more effectively. Learners can perform good learning outcomes if they use adaptive learning strategies (Oxford & Crookall, 1990). The results of Tight's study (2010) indicated that instruction through multiple modalities stimulates the better learning and retention than individual preferences. Style matching also promoted significantly greater retention than mismatching. Surprisingly, little work discusses the use of vocabulary learning strategies on e-learning systems. This research aims to fill this gap and focuses on exploring the effects of using mixed-modality vocabulary learning strategy on e-learning systems. A vocabulary learning system called MyEVA (My English Vocabulary Assistant) is developed for this research. In this study, it was observed that L2 students employed three vocabulary tools to memorize the target words, as revealed by the students' pretest/posttest scores of VKS (Vocabulary Knowledge Scale). A sample of 93 L2 undergraduates signed up to participate in the experiments. They were indicated to use the basic mode of MyEVA, the preference mode of MyEVA, internet dictionary (Yahoo Dictionary), and traditional paper-based dictionary. Finally, the researchers analyzed if the mixed-modality vocabulary learning strategy is more beneficial than individual vocabulary learning strategies on vocabulary retention.

1. The Design of MyEVA

MyEVA is a mixed-modality vocabulary learning system including 3,569 TOEIC (Test of English for International Communication) words and 8 VLSs which are designed for L2 students to improve their English vocabulary capability. Students can navigate any one of those VLSs when they study the words to expand their vocabulary size and enhance the retention. The 8 VLSs designed in MyEVA were initially selected from the memory strategy classified by Schmitt & McCarthy (Schmitt & McCarthy, 1997) and then adjusted to be suitable for the L2 students in Taiwan. The 8 VLSs in MyEVA are: *word card strategy, flash card strategy, synonym strategy, antonym strategy, assonance strategy, clipping strategy, grouping strategy, and imagery strategy*. The screenshot of MyEVA is shown in Fig 1.

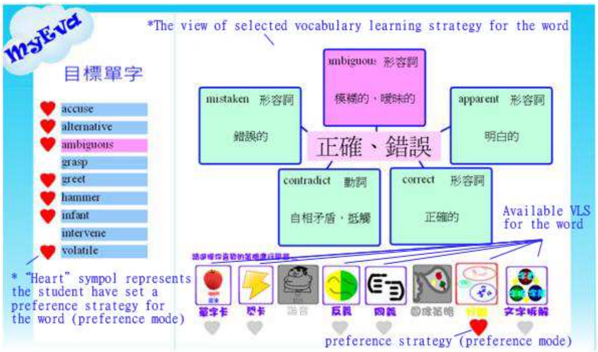


Figure 1. The screenshot of MyEVA

	All Learners (N=93)		Poor Learners (N=54)		Good Learners (N=39)	
	Mean	p-value	Mean	p-value	Mean	p-value
Basic Mode v.s. Preference Mode	4.43/4.83	0.108	3.94/4.39	0.127	4.69/6.08	0.000***
Basic Mode v.s.Internet Dictionary	4.43/4.10	0.280	3.94/4.33	0.407	4.69/4.38	0.352
Basic Mode v.s. Traditional Dictionary	4.43/3.23	0.000***	3.94/3.00	0.028*	4.69/2.77	0.000***
Preference Mode v.s. Internet Dictionary	4.83/4.10	0.005**	4.39/4.33	0.862	6.08/4.38	0.000***
Preference Mode v.s. Traditional Dictionary	4.83/3.23	0.000***	4.39/3.00	0.000***	6.08/2.77	0.000***
Internet Dictionary v.s. Traditional Dictionary	4.10/3.23	0.005**	4.33/3.00	0.000***	4.38/2.77	0.000***

* p<0.05   **p<0.01   ***p<0.001

Figure 2. The summary of experiment results

On the other hand, two navigational modes are designed in MyEVA to examine the effects of preference strategy setting: 1) Basic mode: system is freely opened for students to learn the words and use diverse VLSs. 2) Preference mode: students can set the most favorite learning strategy for the word and the system will show the preference strategy by default when they navigate the word every time.

2. Experiments and Data Analysis

The focus of the study is to determine whether mixed-modality VLS applied on e-learning system has a significant effect on L2 vocabulary retention. Four types of learning activities were conducted: basic mode of MyEVA, preference mode of MyEVA, internet dictionary, and traditional dictionary. In this within-subjects design, all 93 participants practiced the same 24 target words. However, the words were divided into four equivalent groups (A, B, C, D) of six words each, and subjects practiced each group of words under one of the four learning activities. The learning activity was randomly assigned to each group and each subject practiced all groups in the experiments.

This study used the selecting policy of target words with reference to the Folse's research (2006). The main concern in selecting the target words for the experiments was that they be unknown to the subjects. Totally 36 words in the within-subjects design were used and 24 of the words are actual target words. Three-level VKS was used for both pretest and posttest to detect even partial gains in degrees of knowledge. Each word could receive a score of 0, 1, or 2.

To analyze the data collected from the tests, an independent two-sample t-test was performed to determine whether the pretest and posttest exhibit significant differences in learning achievement (at the 95% confidence level).The t-test results showed significant differences between the pretest and posttest (t=18.27, p<0.001). The mean scores of pretest and posttest were 1.06 and 16.61. An advanced observation divided the subjects into good learners (top 40% subjects getting higher scores in pretest, N=39) and poor learners (N=54), and compared the learning outcomes by different learning activities. The results are shown in Fig 2. 1) Overall: preference mode performed the best learning outcomes for L2 students (M=4.83). Both basic mode and preference mode had significant learning effects than internet dictionary and traditional dictionary. However, there were no significant differences between basic mode and preference mode (p=0.108). 2) Poor learners: preference mode performed the best learning outcomes (M=4.39). Internet dictionary (M=4.33) and basic mode (M=3.94) also performed good learning outcomes. However, there were no significant differences between preference mode, basic mode, and internet dictionary. 3) Good learners: preference mode performed the best learning outcomes (M=6.08) and had significant differences compared to the basic mode, internet dictionary, and traditional dictionary (p<0.001); nevertheless, there were no significant differences between basic mode and internet dictionary (p=0.352).

3. Conclusion

This study analyzed if four vocabulary learning tools have different effects on vocabulary retention for L2 students. The results showed that mixed-modality VLS with preference strategy setting (preference mode) stimulates the best vocabulary retention. Findings also indicated that subjects with different prior knowledge performed distinct learning outcomes. Mixed-modality VLS without preference strategy setting (basic mode) had the similar effect with internet dictionary on poor learners. It seems that good learners are aware of using preference strategy and achieve more effective vocabulary retention. The researchers believe the experimental results will have insight into language teachers, curriculum designers, and, in particular, system developers of English e-learning systems.

Acknowledgements

The author would like to thank the NSC of the Republic of China, Taiwan, for financially supporting this research under contract No. 100-2511-S-231 -001.

References

[1] Folse, K. (2006). The Effect of Type of Written Exerciseon L2 Vocabulary Retention. *Teachers of English to Speakers of Other Languages*, 40(2).

[2] Hu, M. & Nation, I.S.P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13 (1), 403-430.

[3] Huang, C. C. (2004). High school and university students' vocabulary knowledge, content knowledge and differential reading comprehension. Taipei , Taiwan : Crane.

[4] Oxford, R. & Crookall, D. (1990). Vocabulary learning: a critical analysis of techniques. *TESL Canada Journal*, 7(2), 9-30.

[5] Oxford, R. (1990). Language learning strategies: What every teacher should know. Boston: Heinle & Heinle Publishers.

[6] Rubinj, J. (1981). Study of cognitive processes in second language learning, *Applied Linguistics*, 2, 117-131.

[7] Schmitt, N. & McCarthy, M. (1997) . Vocabulary Description, Acquisition and Pedagogy. Cambridge Language Teaching Library, 207-209.

[8] Tight, D. G. (2010). Perceptual Learning Style Matching and L2 Vocabulary Acquisition. *Language Learning*, 60(4), 792-833.

Towards a system architecture for ICALL

Elena VOLODINA[a*], Hrafn LOFTSSON[b], Birna ARNBJÖRNSDÓTTIR[c], Lars BORIN[a] & Guðmundur ÖRN LEIFSSON[d]

[a] *Språkbanken (Swedish Language Bank), University of Gothenburg, Sweden*

[b] *School of Computer Science, Reykjavik University, Iceland*

[c] *School of Humanities, University of Iceland, Iceland*

[d] *School of Engineering and Natural Sciences, University of Iceland, Iceland*

*elena.volodina@svenska.gu.se

Abstract: In this paper, we present an on-going project whose overall aim is to develop open-source system architecture for supporting ICALL systems that will facilitate re-use of existing NLP tools and resources on a plug-and-play basis. We introduce the project, describe the approaches adopted by the two language teams, and present two applications being developed using the proposed architecture.

Keywords: Intelligent Computer-Assisted Language Learning, Natural Language Processing, System Architecture, Writing Feedback, Exercise Generation

1. Introduction

It is a remarkable fact that, despite the existence of various accurate Natural Language Processing (NLP) tools and resources that can potentially benefit language learning, very few projects are devoted to development of Intelligent Computer-Assisted Language Learning (ICALL)[1] applications.

This situation calls for a change. We have therefore joined our forces in order to design and develop open-source system architecture for supporting ICALL systems. We make the architecture open-source in order to encourage participation from other researchers and developers, and to facilitate re-usability of existing NLP tools and resources. To test the architecture, we are currently developing two specific applications for Icelandic and Swedish, described in the sections below.

2. An emerging ICALL platform for Icelandic

In the Icelandic part of the project, we are developing a platform which chains together various NLP tools. Internally, the platform uses the *Text Corpus Format* (TCF; an XML format), proposed in the *WebLicht* project [1], for communication of information between the various components. All annotations for a particular text are stored in a single XML file, where each annotation, e.g. at the level of tokens, part-of-speech (PoS) tags, or constituents, is stored in a separate layer. In addition to using the layers proposed in WebLicht, we have added a layer for information about grammatical errors. An important part of the design of the platform is language-independence, i.e. it should be simple to add

1 Intelligence in CALL systems can be understood differently by different researchers. In this paper, we define ICALL as NLP-based CALL, i.e. intelligence in CALL is ensured through the use of NLP tools and resources like parsers, part-of-speech (PoS) taggers, corpora, lexicons, etc.

NLP tools for supporting various languages, the only requirement being that they must communicate with other tools in the platform using TCF.

Currently, no ICALL application exists for the Icelandic language. However, a free and open CALL application named *Icelandic Online (IOL;* http://icelandiconline.is) was launched in 2004. IOL is a pedagogically driven web course which has evolved over time [2]. It has, currently, almost 90,000 registered users, and has received universally positive feedback. In IOL, second-language learners of Icelandic receive feedback from a teacher regarding short written texts. Currently, teachers use special codes for hand-marking specific types of errors, i.e. spelling errors, feature agreement errors, case errors in objects of verbs, etc.

In order to automate part of the error-marking and to test our platform, we are currently in the process of developing a web service which allows students of IOL to send texts to the service for the purpose of detecting particular types of grammatical errors. This will allow students to correct potential errors, re-submit the texts for error detection again, and so forth, before finally submitting the text to the teacher. The web service merely identifies error candidates, but does not attempt to correct errors. Writing feedback is immensely labour intensive and this service will allow students to identify certain types of errors of form, allowing the teacher to focus on content feedback.

The web service submits Icelandic text (input by a student) to the platform, which, in turn, uses tools from the IceNLP toolkit [3], i.e. a tokeniser, a PoS tagger and a finite-state parser, to detect the following types of grammatical errors: (1) feature agreement errors in noun phrases, i.e. errors in gender, number and case; (2) feature agreement errors between subjects and verb complements; (3) feature agreement errors between subject and verbs, i.e. errors in person and number; and (4) incorrect case selection of verb objects. In addition, spelling errors are flagged.

IceNLP outputs TCF containing information from the analysis, i.e. about the individual tokens and their PoS tags, individual constituents and error candidates. The platform forwards the TCF to the web server, which converts it to a human readable HTML format and displays a resulting page to the student, containing the original text submitted and the error candidates highlighted.

3. Lärka – an emerging ICALL platform for Swedish

Language technology research has a long history in Sweden, going back to the 1960s. Most of the basic NLP components exist for Swedish in quite stable and mature forms, e.g. PoS taggers and parsers, annotated reference corpora, and large lexical databases with morphological analysers. Swedish ICALL, however, has a shorter history. Only two projects – ITG [4] and Grim [5] – have resulted in concrete ICALL applications that are used in real-life teaching settings. However, they both use a technology which was state of the art at the time, but which practical experience shows is not the optimal solution today.

The application developed by the Swedish partner is web-based, and has the working title *Lärka* (http://spraakbanken.gu.se/larka/). Its principle is to have all functionalities web-service-based to ensure flexibility and reuse. The main components of Lärka's architecture are the following.

(1) Lärka's *frontend* is the graphical user interface that handles user interaction, sends requests to the backend and assigns behaviour to buttons and fields.

(2) Lärka's *backend* consists of a number of web services for generating language training exercises, selecting distractors, interacting with databases and generating syntactic trees. The backend depends heavily on Korp and Karp described below.

(3) *Korp* (http://spraakbanken.gu.se/korp/) is Språkbanken's web-service based infrastructure for maintaining and searching a constantly growing corpus collection, at the moment amounting to about one billion words of Swedish text [6]. The corpora available through Korp contain multiple annotations, e.g. lemmatisation, compound analysis, PoS tagging, and syntactic dependency trees, which can form the basis for versatile exercises.

(4) *Karp* (http://spraakbanken.gu.se/karp/) is the corresponding web-service based infrastructure for maintaining and retrieving information from Språkbanken's collection of computational lexical resources [7].

Korp and Karp together provide the necessary information for Lärka's learning activities. Once the sentence or lexical information is retrieved, the relevant algorithm is applied to generate an output for the exercise. The output from Lärka's backend can be used by any program, not only Lärka's frontend, for example in apps for mobile phones.

Each exercise (or any other learner activity) is added as a separate module consisting of backend and frontend parts. Exercises can thus be developed separately and be "plugged in" with minimal efforts into the architecture.

At the moment, Lärka offers three (multiple-choice) exercise types, all based on authentic corpus data: (1) identifying parts of speech; (2) identifying syntactic relations; and (3) vocabulary exercises. Each of the exercise types can be deployed as a test or as a self-study exercise. Next on our "to-do" list is to extend Lärka's exercise scope, add syntactic trees to every sentence, and add an "encyclopaedia" of basic linguistic terms.

4. Concluding remarks

The main idea of our project is to propagate the re-use of existing accurate NLP tools and resources in language learning by designing and implementing a system architecture for ICALL, at the moment on a more abstract level – where our two subprojects share the general philosophy of making NLP components available via web services – and in the next phase of the project on the concrete level of having a common data exchange format (e.g. TCF). It is thus clear that ICALL researchers and developers can be affected by our project. In addition, language learners will also be affected because the system architecture and the two test applications will benefit language learners in the form of a more versatile and open-ended CALL experience, thanks to the NLP components.

References

[1] Hinrichs, M, Zastrow, T., & Hinrichs, E. WebLicht: Web-based LRT services in a distributed eScience infrastructure. *Proceedings of LREC 2010*. Valletta, Malta: ELRA.

[2] Arnbjörnsdóttir, B. 2004. Teaching morphologically complex languages online: Theoretical questions and practical answers. In P. J. Hendrichsen (Ed.*) CALL for the Nordic Languages*. (Copenhagen Studies in Language 30.) Copenhagen: Samfundslitteratur.

[3] Loftsson, H., & Rögnvaldsson, E. 2007. IceNLP: A Natural Language Processing Toolkit for Icelandic. *Proceedings of InterSpeech 2007, Special Session: 'Speech and language technology for less-resourced languages'*. Antwerp, Belgium.

[4] Borin, L., & Saxena, A. 2004. Grammar, incorporated. In P. J. Hendrichsen (Ed.), *CALL for the Nordic languages* (pp. 125–145). (Copenhagen Studies in Language 30.) Copenhagen: Samfundslitteratur.

[5] Knutsson, O. 2005. *Developing and evaluating language tools for writers and learners of Swedish.* Doctoral thesis in human-computer interaction. KTH, Stockholm.

[6] Borin, L., Forsberg, M., & Roxendal, J. 2012. Korp – the corpus infrastructure of Språkbanken. *Proceedings of LREC 2012.* Istanbul, *Turkey: ELRA.*

[7] Borin, L., Forsberg, M, Olsson, L.-J., & Uppström, J. 2012. The open lexical infrastructure of Språkbanken. *Proceedings of LREC 2012.* Istanbul, *Turkey: ELRA.*

Preliminary Study on Factors Affecting Aptitude Level for Social Learning Focusing on EFL Online Discussion

Yoshikko GODA[a], Masanori YAMADA[b], Hideya MATSUKAWA[c], Kojiro HATA[d], & Seisuke YASUNAMI[a]
[a]*Research Center for Higher Education, Kumamoto University, Japan*
[b]*Research Center for Higher Education, Kanazawa University, Japan*
[c]*Center for Education in Liberal Arts and Sciences, Osaka University, Japan*
[d]*Faculty of Social and Management Studies, Otemae University, Japan*
*ygoda@kumamoto-u.ac.jp

Abstract: The purposes of this preliminary study are (1) to investigate possible factors for English as a foreign language (EFL) learners' behavior and attitudes towards computer-supported collaborative learning focusing on online discussion. Comment numbers, satisfaction and perceived group contributions on two online discussions of 58 EFL students were analyzed with multiple regression correlation, in relation with items of six inventories; (a) social skill (Kikuchi, 2007), (b) Self-efficacy for English learning with CSCL based on Matsunuma (2006), (c) English learning strategy (Kubo, 1999), (d) Social presence and cognitive presence with text-chat developed based on Gunawardena & Zittle (1997), (e) Felder-Soloman Index of Learning Style (ILS) (Felder & Silverman, 1988), and (f) researcher-developed questionnaire for a discussion task. The results show that 11 items collectively account for 76.2% of the comment number variance ($F(11,46)=13.39$, $p<.01$), 15 items and "sequencel" type score of learning style significantly explain 89.2% of the learners' satisfaction ($F(16, 41)=21.12$, $p<.01$), and 2 items and "sequential" type score of learning style describe 32.8% of the learners' perceived contribution ($F(3, 54)=8.80$, $p<.01$). There are three overlapped items and one learning style type for all prediction equations. It indicates that 25 items and 11 items for sequential/global type in the learning style types may be utilized to predict EFL learners' behavior and attitudes towards online discussion.

Keywords: online discussion, CSCL, EFL, social skills, learning style, social learning

Introduction

The purposes of this preliminary study are to investigate possible factors for English as a foreign language (EFL) learners' behavior and attitudes towards computer-supported collaborative learning (CSCL) focusing on online discussion. As social media has become popular, trials of applying it to educational setting is increasing. One of major problems of such applications is that learning activities with social media seems beneficial for learners with high aptitude towards online social activities. In other words, learners with low aptitude seem not to obtain much benefit from social learning. This research is placed as a basic research for a central algorithm to determine learners' aptitude level towards social learning in order to provide individualized appropriate supports.

1. Research Methods

There were 109 Japanese university students of CALL class as initial research participants. The data of the 58 students with all research tasks completed were analyzed. Three discussion topics were prepared. First one was used to provide learners chances to practice using the bulletin board system of the learning management system (LMS), and the discussion 1 was not included for the data analysis. Students were required to discuss on each topic in a group of four to six members in their class. A span for each discussion was one week between two classroom instructions. Teachers provided some feedback only in the face-to-face instructions, not during the discussion spans.

In order to collect data related to learners' attributions and characteristics, six inventories were employed; (a) social skill (Kikuchi, 2007), (b) Self-efficacy for English learning with CSCL based on Matsunuma (2006), (c) English learning strategy (Kubo, 1999), (d) Social presence and cognitive presence with text-chat developed based on Gunawardena & Zittle (1997), (e) Felder-Soloman Index of Learning Style(c) (ILS) (Felder & Silverman, 1988), and (f) researcher-developed questionnaire for a discussion task. Brief description of each inventory is as followed. All inventories except for (e) and (f) are 5-point-Likert scales (1: Not agree at all - 5: very much agree). (a) Social skill inventory includes 21 items in four categories. (b) Self-efficacy for English learning with CSCL consists of 17 items. (c) English learning style inventory contains 13 items. (d) Social presence and cognitive presence with online discussion provides 25 items and six items are related to online discussion. (e) ILS consists of 44 dichotomous items for four dimensions, with the internal consistency reliability of ranging .55 to .77 (Litzinger, Lee, Wise & Felder, 2007). (f) Researcher-developed questionnaire consists of four questions; a 4-point Likert item for satisfaction of the discussion, a perceived contribution to the discussion group with a response of percentage, and two open-end questions about pros and cons of the learning activity. It's conducted on each discussion topic before the instructors' feedback.

For the data analysis, the comment numbers recorded on LMS, the satisfaction points with the inventory (f), and the perceived contributions (%) with the inventory (f) were utilized as dependent variables (DVs). The average numbers for two discussion topics were employed as DVs. Fifty one items of the inventories (a) to (d) and the raw scores with the full score of 11 for each dimension of learning style were used as independent variables (IVs). Each DV was performed with multiple regression correlation (MRC) analysis in the presence of all IVs. MRC would be repeated three times to create three prediction models for each DV and its experimental-wise *alpha* is .03 with the priori-set alpha level .01.

2. Results and Conclusion

According to the results of stepwise MRC analyses (Table 1), three models for the DVs were suggested with statistical significance. Eleven items collectively account for 76.2% of the comment number variance ($F(11,46)=13.39$, $p<.01$), 15 items and "sequence" type score of learning style explain 89.2% of the learners' satisfaction ($F(16, 41)=21.12$, $p<.01$), and two items and "sequence" type score of learning style describe 32.8% of the learners' perceived contribution ($F(3, 54)=8.80$, $p<.01$). There are some overlaps for three items and one learning style type among the all prediction equations. It indicates that 25 items and 11 items for sequential/global type in the learning style types may be utilized to predict EFL learners' behavior and attitudes towards online discussion. Furthermore, there could be a possibility to apply them to classification of learner aptitude types for social learning. The limitations of this research should be considered in the future research; grouping and group memberships, kinds of discussion topics, and learners' English proficiency.

Table 1. Inventory Items in Each Prediction Model and Its Entry Order

DV	Entry Order	Inventory*	Item #	*B*	Item
Number of Comments	a**			1.26	
	1	d	6	0.39	I felt comfortable participating in the course discussions.
	2	c	10	-0.74	I guess meanings of the contents based on a whole story.
	3	c	8	0.67	I guess meanings of unknown words when I read English materials.
	4	d	23	-0.46	I can describe ways to test and apply the knowledge created in this course.
	5	b	2	0.19	I can get high marks and scores in English classes.
	6	d	17	0.14	I felt motivated to explore content related questions.
	7	c	7	0.28	I make tables and charts of relationships when I learn English grammar.
	8	c	5	-0.27	I determine what I don't understand when I read English materials.
	9	a	6	-0.20	I can handle criticisms from others well.
	10	d	16	0.37	Course activities piqued my curiosity.
	11	d	25	-0.34	I felt motivated to explore content related questions.
Satisfaction	a**			2.90	
	1	a	11	-0.27	I can come to a composition with uncomfortable peoples.
	2	d	12	0.16	I was able to form distinct individual impressions of some participants via a text-based medium.
	3	b	8	-0.35	I think that I know how to study English.
	4	d	17	0.17	I felt motivated to explore content related questions.
	5	b	5	0.27	I know a lot about learning contents of English.
	6	b	10	-0.25	I can read English newspaper.
	7	c	7	-0.22	I make tables and charts of relationships when I learn English grammar.
	8	d	24	0.20	I have developed solutions to course problems that can be applied in practice.
	9	a	19	0.11	I like discussions and debates.
	10	e		0.07	Learning Style "Sequence" score
	11	a	14	0.15	I can help others well.
	12	a	10	-0.15	When I feel fear or anxiety, I can handle them well.
	13	d	7	0.13	The members created a feeling of an online community.
	14	c	13	-0.12	When I read English materials, I first check the meanings of unknown words and write them down on the text.
	15	a	6	-0.10	I can handle criticisms from others well.
	16	c	9	-0.09	I try to comprehend the overall meaning without caring details that I don't understand well when I read English materials.
Contribution	a**			-28.65	
	1	d	8	7.00	The moderators facilitated discussions on the online discussion.
	2	e		2.48	Learning Style "Sequence" score
	3	b	4	5.22	I can comprehend What to be taught in English classes.

Note. *Alphabets are match to the inventories in 1. Research Method. **Constant Number.

References

[1] Felder, R.M., & Silverman, L.K. (1988). Learning and teaching styles in engineering education. *Engineering Education*, *78*(7), 674-681.

[2] Gunawardena, C.N. & Zittle, J. F. (1997) Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. The American Journal of Distance Education, 11(3), 8-26

[3] Kikuchi, A. (2007). 社会的スキルを測る:KiSS‐18 ハンドブック[Measure Social Skills: KiSS-18 Handbook]. 川島書店

[4] Kubo, N. (1999). 大学生の英語学習における動機づけモデルの検討 : 学習動機,認知的評価. [Motivational model of undergraduates' English learning: Learning motivation and cognitive evaluation.]. 学習行動およびパフォーマンスの関連 教育心理学研究 47(4), 511-520

[5] Matsunuma, M. (2006). 英語自己効力感(ESE)尺度の作成 [Development of English Self-Efficacy (ESE) Inventry]. 早稲田大学大学院教育学研究科紀要別冊, 89-97.

[6] Litzinger, T.A., Lee, S.H., Wise, J.C., & Felder, R.M. (2007). A psychometric study of the index of learning styles. *Journal of Engineering Education*, *96*(4), 309-319

Podcasting for Language Learning Using an Integrated Approach

Hsiu-Ting HUNG
Department of English, Kaohsiung First University of Science and Technology, Taiwan
hhung@nkfust.edu.tw

Abstract: Podcasts are a rich source of language input and an alternative to ready-made pedagogical materials that can be freely incorporated into content-based instruction or skill-based classes. This paper argues that when developing language lessons with the use of podcasts, it is important for language teachers to consider an integrated way to provide effective conditions for language acquisition. To explore the potential of podcasts for language learning, this paper discusses their instructional benefits and recommends that podcasts should be integrated into task-based learning lessons via a WebQuest approach, with consideration of second language acquisition principles pertaining to extensive input and comprehensible input.

Keywords: English as a foreign language, podcasts, WebQuest, task-based learning

1. Educational benefits of podcasts

English as a foreign language (EFL) teachers often find it challenging to expose students to authentic speech by native speakers and provide meaningful opportunities for communicating in the target language in and out of the classroom. However, the emergence of podcasting appears to offer rich source of material for use in an EFL context and to assist foreign language instruction in general. Lazzari [5] defined podcasting as "a method for distributing digital video and audio contents over the Internet" (p.27). Podcasts enable students to choose when and where to learn and to do so at their desired learning pace and path. Moreover, podcasts can be easily downloaded, shared, and even edited, granting learners more control and ownership of their learning. Given the flexibility to adopt and adapt podcasts according to individual needs, students are generally more engaged in the learning process and learning outcomes are thus believed to be maximized.

In developing a taxonomy of podcasting for language learning, Rosell-Aguilar [6] identified two major uses: creating podcasts and using already available ones. This paper focuses on the latter, the use of existing podcasts as language learning resources. In the foreign language classroom, podcasts are a rich source of language input that can serve as an alternative to ready-made pedagogical materials and can be freely incorporated into content-based instruction or skill-based classes. EFL teachers can access or subscribe to authoritative and authentic podcasting services (e.g., BBC Learning English) and then use the downloaded audio or video podcasts for in-class practice. EFL teachers can also ask their students to subscribe to podcasts in any content area to enable self-study after class. While these are common practices in and out of school, such uses of podcasts may not lead to sustained and quality learning experiences, due to a lack of systematic integration and learning guidance. Accordingly, this paper argues that when developing language lessons with existing podcasts, it is important for EFL teachers to consider an integrated way to provide effective conditions for language acquisition.

2. WebQuest as an integrated approach to using podcasts for language learning

One integrated approach to using podcasts for language learning is through WebQuests. According to the definition provided on the official website of the WebQuest Research Consortium, a WebQuest is "an inquiry-oriented lesson format in which most or all the information that learners work with comes from the web." As research on uses of WebQuests have grown, the idea has been expanded to refer to an instructional approach that makes use of online resources structured in this specific format [3].

Since the launch of WebQuest by Bernie Dodge and Tom March in 1995, the structure of this format has been refined to five major segments in the form of web pages: 1) Introduction, 2) Task, 3) Process, 4) Evaluation, and 5) Conclusion, with one optional segment, 6) References. The Introduction page orients learners to the task by providing some background information or a scenario to explain that they will be "learning by doing" during the WebQuest. The Task page provides an overview and explanation of the learning task. The Process page details a set of clearly described steps, interwoven with teacher selected Internet resources (often in hyperlinks), which direct the learners to examine the information to complete the task. The Evaluation page includes an assessment to evaluate the learners' performance in relation to the learning objectives of the task, or offers a rubric for the learners' self-evaluation of their learning development. The Conclusion page summarizes what the learners should have learned in the WebQuest and invites them to reflect on their learning. As an optional component, the References page is often labeled differently, such as Teacher Page or Resources. This usually includes: additional learning resources pertaining to the core task of the WebQuest, information sources embedded in all the pages of the WebQuest, or any other relevant information that the WebQuest creator deems helpful (e.g., credits, acknowledgements, or teacher guides). Altogether, these segments are designed in a structured format to guide learners through WebQuests in a sequential presentation of web pages.

Conventional WebQuests are often created using HTML editors or stand-alone web design applications (e.g., Notepad and Dreamweaver), which requires educators to have knowledge of HTML and web design. This prerequisite of technological literacy may discourage educators who have limited time and computer skills from becoming contributors and users of WebQuests. Fortunately, the advent of Web 2.0 technology has made the creation of WebQuests much easier. Web 2.0 authoring tools, such as Google Sties, allow teachers to easily create and edit web pages with a few clicks of a mouse button. Given the ease of creation, this paper proposes that language-learning podcasts can be seamlessly integrated into WebQuests using Web 2.0 authoring tools, to provide focused and contextualized materials that allow for skills development in an input-rich, task-based learning environment.

3. Designing effective WebQuests with integrated podcasts for English learners

When designing WebQuests with podcasts as a source of content integration, language teachers should attend to Second Language Acquisition (SLA) principles, particularly extensive input [2] and comprehensible input [4] as follows.

3.1 Podcast integration in WebQuests should be designed to offer extensive input.

Dodge [1] suggested the following five guiding principles for designing a WebQuest: 1) Find great sites, 2) Orchestrate your learners and resources, 3) Challenge your learners to think, 4) Use the medium, and 5) Scaffold high expectations. As applied to the case of integrating podcasts into WebQuests, the first and foremost guiding principle of "find great sites" is translated into a commitment to locate rich and relevant podcasts of adequate length in support the desired learning task. Given that exposing learners to rich and meaningful input in the target language is a prerequisite for language acquisition, teachers should "(a) make resources available and (b) provide learner-training in how to make effective use of the resources" [2, p.218]. In addition, teachers may provide pre-task exercises to make the key linguistic points more salient and prime students with the necessary linguistic skills (e.g., vocabulary and grammar) required for the learning task in the WebQuest.

3.2 Podcast integration in WebQuests should be modified to offer comprehensible input.

Many researchers and educators have embraced the potential of WebQuests to enhance learning motivation, critical thinking skills, and collaboration skills [3]. While numerous WebQuests have been developed by content area teachers, "most lacked the linguistic supports that [English language learners] are likely to need" [9, p.41]. To make existing WebQuests appropriate for use with English language learners, the content and design of the WebQuests have to be carefully modified to ensure quality learning and match students' proficiency levels. Accordingly, Sox and Rubinstein-Avila [7] proposed a rubric centering on the linguistic, multimedia and organizational features of WebQuests for teachers to cater to the specific needs of English language learners and make input more comprehensible. Such adaptations are considered as an essential step in the utilization of podcasts for language learning through the WebQuest approach.

In sum, this paper advocates for the integrated use of podcasts in WebQuests to aid language learning. The proposed approach allows learners to access podcasts as extensive and comprehensible input in a task-based and contexualized manner under the WebQuest infrastructure. A research agenda examining the learning effects of podcasts based on this integrated approach is thus needed to provide empirical guidelines for future applications of podcast pedagogy.

Acknowledgement

This paper was partially supported by the National Science Council of Taiwan (NSC 101-2511-S-327-002).

References

[1] Dodge, B. (2001). FOCUS: Five rules for writing a great WebQuest. *Learning and Leading with Technology, 28(8),* 6-10.
[2] Ellis, R. (2005). Principles of instructed language learning. *System, 33(2),* 209-224.
[3] Halat, E. (2008). A good teaching technique: WebQuests. *The Cleaning House, 81(3),* 109-111.
[4] Krashen, S. (1985). The input hypothesis: Issues and implications. London: Longman.
[5] Lazzari, M. (2009). Creative use of podcasting in higher education and its effect on competitive agency. *Computers and Education, 52(1),* 27-34.
[6] Rosell-Aguilar, F. (2007). Top of the pods – In search of a podcasting "podagogy" for language learning. *Computer Assisted Language Learning, 20(5),* 471-492.
[7] Sox, A., & Rubinstein-Avila, E. (2009). WebQuests for English-language learners: Essential elements for design. *Journal of Adolescent and Adult Literacy, 53(1),* 38-48.

The Annotation Functionality to Support Language Teaching and Learning

Shih-Hsien YANG[a*], Sheng-Shiang TSENG[b] & Hui-Chin YEH[c]
[a]*Department of Applied Foreign Languages, National Formosa University, Taiwan*
[b]*Department of Applied Foreign Languages, National Yunlin University of Science and Technology, Taiwan*
[c]*Graduate School of Applied Foreign Languages, National Yunlin University of Science and Technology, Taiwan*
shiyang@nfu.edu.tw

Abstract: This project intends to develop a supportive English website management system for both teachers and students to manage the available internet resources to make authentic learning more inviting and effective. The annotation functionality was designed to scaffold, support, and evaluate students' learning processes and outcomes. With the annotation functionality, English as a Foreign Language (EFL) students can take initiative in their own learning with the virtual assistance. This website learning management system intends to make it possible to turn teachers as authoritative figures into virtual language learning facilitators through annotations.

Keywords: Annotation, the virtual facilitator, language learning, scaffolding

1. Introduction

Due to the rapid growth of technology, language learning websites are widely employed in different fields. Learners at all ages have increasingly access to many websites in English, which have the potential to provide seemingly limitless amount of authentic learning materials for EFL learners. In terms of language teaching, teachers have been seeking for different possibilities to assist students in their learning through the learning websites. English websites emerge to support English learning and teaching in the past few decades. It is evident that a wide variety of teaching websites are designed and developed based on various teaching principles ranging from behaviorism, cognitive theories to constructivism in order to assist students in improving their English proficiency. However, the rich resources on the internet did not provide all the answers to language learning and teaching unless there are enough scaffolding functions designed to support learning. Particularly, EFL learners from all different levels with limited language proficiency found it almost impossible to get access to the authentic language learning websites. The purposes of the study are to develop a supportive language website management system to assist both EFL teachers and students for teaching and learning.

It has been argued that it is effective for language learners to get access to the authentic materials for meaningful learning. With available resources online, it is imperative for the teachers to provide scaffolding to assist learners to obtain the adequate authentic website resources and to make full advantage of them. For EFL contexts, two compelling concerns were identified to meet the EFL learners' needs. First, how exactly can the EFL learners benefit from online learning through authentic English websites? How is it possible for EFL learners to get access to those learning websites so as to sustain their ongoing learning? Through the annotations, a virtual instructor could possibly be expected to facilitate students' learning through the annotations functionality.

The design of this supportive language learning management system for web content management was grounded in the scaffolding theory. Scaffolding theory has been advocated to provide learners with assistance who could not accomplish a task beyond their current level on their own and such support can be gradually decreased while learners gain autonomy over their learning (Azevedo; Cromley; Thomas, Seibert & Tron, 2003). According to Jonassen (1994), in order to arm students with the ability to accomplish a complex and challenging task, it is imperative to provide learners with individualized scaffolding and support; otherwise, learning will be ineffective. With the scaffolding, learners can go beyond the zone of proximal development; that is, extend what they can do now toward the goal of accomplishing more difficult tasks with support (Dunlap & Grabinger, 1996). Furthermore, learning takes place and persists when students are motivated to learn by meeting their individual interests and needs. This compelling need for the individualized learning speaks to the ZPD principle, which stresses the importance of providing assistance to learners to reach their potential learning level.

2. The website management system

2.1 Annotations in dialogue boxes in a website management system

The system consists of the mechanisms, including the guiding annotations in dialogue boxes to provide students instruction, and assessment functions to evaluate to what extent the students have learnt. The annotation functionality could play the role of a virtual facilitator to scaffold, support, and evaluate students' learning outcomes. Students have multiple accesses to acquire a wide range of information presented in different forms, such as texts, graphics, or animations to foster their proficiency level and particularly enhance their language ability through the websites. However, not every language learner is able to improve his/her own language skills from the web-based learning environment. They indicated that for those beginners and lower achievers, owing to their limited language proficiency, they would be overwhelmed with too much information in the learning websites in non-linear fashion, and become less motivated to learn. In light of this, the annotations in dialogue boxes are developed for teachers to write the specific instruction on the website and provide specific and adaptive instruction to assist learners to learn how they can undertake the tasks step by step. The content of dialogue box can be filled out by teachers. Other learners can also write down instructions and add messages, information either in simple English or Chinese to facilitate other learners' comprehension in the website management system. The annotations in the dialogue boxes allow teachers to move around bubbles into the sessions that need more explanations and these individualized and adaptive functions make it possible for teachers to select the appropriate learning websites to accompany with teachers' explanations and written instructions. After the learners have undertaken the tasks, they can also write their own annotations for other learners to get access to the available websites for further learning. The collective effort in writing annotations creates different opportunities for learning.

2.2 Assessment

Through the language assessment, teachers can come into realize the effectiveness of teaching, appropriateness of the materials, and the effectiveness of the syllabus (Hardley, 2000). Furthermore, students can have multiple accesses to self-evaluate their own learning so as to discover what needs to be improved to reach respective goals of learning (Sadler, 1989). The formative assessment can further lower students' affective filter and foster learning motivation, because the feedback of the formative assessment allows lower achievers to realize their own drawbacks to overcome the difficulty possibly stemming from their lack of innate ability through efforts (Ames, 1992). The language learning system with the annotation functionality intends to incorporate formative assessment and summative assessment to design the following test formats to evaluate what students have learned from the particular selected websites. In order to realize how much learners have learnt from the particular websites, two types of tests are adopted: journal writing and vocabulary test. To start with, according to Long & Richards (1987), journal writing is regarded as a type of language assessment in investigating what students have acquired by having students reflect upon their own learning. As a consequence, it is hoped that through the journal writing, learners can constantly examine their own learning by reflecting on what has learned, what is expected to learn, and what interests them.

Conclusion

Even though the available resources on the websites provide learners with opportunities to get access to multiple resources to foster language proficiency, there are still problems remaining unresolved for EFL learners to comprehend the English learning websites with their limited language proficiency. The problems mainly result from the fact that EFL learners often do not receive enough support in their learning process. To this end, a website management system provides the annotation functionality to scaffold language learning and teaching. The system aims to serve as a virtual language facilitator in helping learners to lower their affective anxiety and increase the accessibility for them to continuously learn at their own pace. The system mainly consists of annotations in dialogue boxes and assessment. The guiding instruction in dialogue boxes is designed to increase learners' understanding of how to access to English learning websites. The open-ended space intends to go with teachers' pedagogical needs and scaffolding instruction to prevent learners from overwhelmingly discharged from the targeted language learning. The assessment formats are designed to measure what students have learned for both summative and formative assessment.

References

[1] Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.

[2] Azevedo, R., Cromley, J. G.., Thomas, L., Seibert, D., & Tron, M., (2003). *Online process scaffolding and students' self-regulated learning with hypermedia*. Paper presented at the 2003 Annual Meeting of the American Educational Research Association Chicago, IL.

[3] Dunlap, J. C., & Grabinger, R. S. (1996). Rich environments for active learning in the higher education classroom. In B. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design* (pp. 65-82). Englewood Cliffs, NJ: Educational Technology.

[4] Hardley, A. O. (2000). *Teaching Language in Context*. Boston: Heinle & Heinle.

[5] Rubens, W., Emans, B., Leinonen, T., Skarmeta, A.G., & Simons, R.J. (2005). Design of web-based collaborative learning environments. Translating the pedagogical learning principles to human computer interface. *Computers & Education, 45*(3), 276-294.

[6] Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

[7] Long M.H & Richards, J.C. (1987). *Methodology in TESOL*. USA: Heinle&Heinle.

[8] Sung, Y. T., Chang, K. E., Chiou, S. K., & Hou, H. T. (2005). The design and application of a web-based self- and peer-assessment system. Computers and Education, *45*(2), 187-202.

[9] Wiburg,K.M.(2003).*Technology and teaching English language learners*. New York:Personal Education.

Using Peer Review to Enhance EFL Writing via a Web-based System under Cross-Institutional Setting

Wen-Chi Vivian Wu [a], **Ching-Huei Chen**[b] **& Yu-Chuan Joni Chao** [a]
[a] *Department of English Language, Literature and Linguistics, Providence University, Taiwan*
[b]*Digital Content Technology and Management, National Changhua University of Education, Taiwan*
wcwu@pu.edu.tw

Abstract: The researchers of this study self-designed an online writing system, entitled "Collaborative Online Peer Editing and Reviewing (COPER)" in order to investigate EFL student satisfaction and self-efficacy regarding their writing class employing anonymous cross-institutional peer review. The use of COPER allowed the researchers to create a more active environment in which the students received constructive feedback from sources from their cross-school peers. In a cross-campus mode via COPER, a semester long writing project was carried out by 66 English-majored students at the two schools in which two rounds of anonymous peer review tasks (one movie reflection and one news reflection) were accomplished including reviewing and commenting on the reflective writing of cross-school peers. The researchers employed a mixed research methodology, including a questionnaire and open-ended questions to investigate the students' perceptions regarding their satisfaction , self-efficacy, and critical thinking skills resulting from this on-line task.

Keywords: EFL, satisfaction, self-efficacy, COPER, peer familiarity, peer familiarity

Introduction

One of the most commonly used strategies in ESL/EFL writing classrooms is peer review, used in order to alleviate the instructors' heavy burden as a result of grading an enormous amount of their students' writing work, as well as to ensure that their students to receive timely, quality feedback about their writing. The recent literature review on the application of peer review in college-level writing [1, 2] supports peer review and feedback as a way to enhance identification and production of writing produced by L2 (second language) learners in local and global contexts as well as to facilitate critical thinking. Anonymous peer review activities not only can serve as a solution to the problems caused by peer familiarity, as mentioned above, but also have proven to be useful in enhancing the quality and amount of critical feedback [3]. In order to maximize the effectiveness of peer review (to avoid peer familiarity while at the same time to guarantee the genuine anonymity of the peer review), a double-level-anonymous (cross-institutional) peer review activity was utilized in this study. In this study the researchers paid special attention to the students' satisfaction and self-efficacy. It was believed by the researchers that an investigation into student satisfaction could indirectly reflect student achievement in the peer review task, as well as sustainability, i.e. whether the students would be willing to continue their participation. The researchers were interested in investigating whether the students' self-efficacy (self-confidence) would be elevated as the result of the double-level anonymity featured in the cross-institutional task.

Data collection and experimental procedures

A total of 66 undergraduates participated in this study. They were all English majors registered in required writing classes when the study was conducted. The participants were recruited from two distinct groups—34 from a comprehensive university in central Taiwan, and 32 from a technical college in southern Taiwan. They anonymously reviewed two rounds of reflective writing by the other school's students using the online platform, "COPER." For data collection, a 7-point Likert-scale questionnaire with 43 question items and two open-ended questions modified from papers [4, 5] was used to investigate the participants' perceptions with regard to their satisfaction and self-confidence. Each student in the peer review task completed two rounds of reflective writing, with each round including both drafting and revising processes. A movie entitled "Cry Freedom" was used in the first round of reflective writing; on the other hand, in the second round, the instructors assigned three current news issues from which the students could choose. The students first researched the three news articles on the Internet and subsequently decided which to write about for their reflection.

The Design of the Peer-review System (COPER)

The design idea: The COPER System allowed users to engage into online group discussion by posting their reflective papers for their assigned cross-institutional peers to view while at the same time they review the papers of the other members in their group by providing constructive feedback and comments. An additional function of the system was Facebook instant message notification. The writer was notified of the receipt of the feedback and comments on their writing paper from their anonymous peers when they logged into the System Website. In addition, the writer was also able to discover who had or had not provided feedback. At the same time, those who had given feedback on the assigned paper could also find out who else had or had not provided feedback on the same paper. All of the original reflective papers and revised papers, along with feedback and comments, were stored in the database of the system for ease data retrieval.

The software used to create the system included Html and PHP+MySql. As for the interface of the system, there were two types of user authority, that is the administer and user. At the front page of the system, the status of users was identified first. The log-in pages were separated by the two different types of users. Administrators were able to make use of four features: (1) to view all the group members, their posted papers, and received feedback, (2) to announce updates on the interfaces controlled by the administer, which allowed all the users to view the announced news on their front pages, (3) to add new users manually on the administrator's interface, including student identification numbers, user names, passwords, and groups, and (4) to configure deadlines and starting points for posting papers and providing feedback. Student users were able to view the updates and the articles assigned to them for peer review. During the review time period, they were able to evaluate their assigned papers and, furthermore, provide the writers of the papers with feedback.

Results and Discussion

The results showed that students were moderately satisfied with the peer feedback they received, with a mean score of 4.35 (M=4.35) measured on the modified 7-point Likert-scale questionnaire. The moderate level could be associated with a perception of incomplete training and insufficient practice on how to generate constructive and critical feedback, even though the students received a series of coaching sessions at the outset of this experiment. Although the quantitative data demonstrated that the students were not extremely contented with the feedback and comments provided by their peers, the findings from the open-ended questions revealed that the participants did appreciate and value the comments from their peers.

The results of this study also showed that the students' self-efficacy sustained a satisfactory level (M=4.89). This finding implied that the students developed self-confidence in executing the peer-review tests because of the anonymous nature of the review. They did not know whose writing drafts they were reviewing and thus were relieved of the concern that negative feedback on their peers' drafts could damage the group harmony of their class. Paper [3] discovered that incorporating anonymity into e-peer review activities increased the amount of critical comments generated by peers as well as the students' self-confidence in reviewing and commenting their peers' writing works.

In addition to quantitative findings, the findings of the open-ended questions also show that the students in the study perceived that their critical thinking was facilitated because they were engaged in several rounds of reviewing and feedback. This echoed the study of the paper [6]. Their discoveries suggested that the givers/student reviewers benefited more than receivers/student writers. When givers of feedback were reviewing and commenting on the writing drafts of others, they kept connecting the critical feedback they gave to others with their own writing in order not to make the same errors.

Acknowledgment

This research was partially supported by the National Science Council in Taiwan through Grant NSC 100-2511-S-126-004.

References

[1] H. F. Shang, "An exploratory study of E-mail application on FL writing performance." *Computer Assisted Language Learning*, vol. 20, no. 1, pp. 79-96, 2007.

[2] M. Guardado and L. Shi, "ESL students' experiences of online peer feedback." *Computers and Composition*, vol. 24, pp. 443-461, 2007.

[3] Lu, R., & Bol, L. "A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback." *Journal of Interactive Online Learning*, vol.6, no.2, pp. 100-115, 2007.

[4] S. C. Tseng and C. C. Tsai, "Taiwan college students' self-efficacy and motivation of learning in online peer assessment environments." *Internet and Higher Education*, vol. 13, pp. 164-169, 2010.

[5] Y. Xiao and R. Lucking, "The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment." *Internet and Higher Education*, vol. 11, pp. 186-193, 2008.

[6] K. Lundstrom and W. Baker, "To give is better than to receive: The benefits of peer review to the reviewer's own writing." *Journal of Second Language Writing*, vol. 18, pp. 30-43, 2009.

“Facebook and Pedagogy” Heating Up Everyday English Communication

Hsing-chin LEE*
National Taipei College of Business University, Taiwan
*hsingchinlee@gmail.com

Abstract: Computer Mediated Communication (CMC), when conducted under time and space constraints, has been closely linked to effective language learning. This paper first examines and evaluates the pedagogical effects of teacher-student interaction on Facebook Group Page and then analyzes sampled data concerning online collaborative discussions based on children’s literature. The level of teacher participation affects student motivation and increases a more positive teacher- student relationship. In response to teacher-generated questions based on a children’s book, participations are asked to express themselves in English. The messages/reflections that the teacher self discloses encourage participants frequently to access the Facebook Group Page. Discussions of student feedback concerning their involvement in writing and participating in English and the effects of teacher self-disclosure via Facebook are stated. A survey on how often students post entries in English to the Facebook Group Page is conducted. Finally, potential drawbacks of relying too heavily on the use of technology are cited. Methods used in this study include Computer-Mediated Communication and literature-based on-line Collaborative Learning.

Keywords: Teacher-student interaction, Computer Mediated Communication, teacher-student relationship, on-line collaborative learning; English children’s literature

Introduction

This paper describes the elements incorporated within Computer Mediated Communication and literature-based Collaborative Online Learning (COL) approach. Stevick (1976) emphasizes that language learning methods should take account of both learners’ and teachers’ attitudes toward language teaching to achieve the goal of effective learning. This paper proposes the idea of integrating the Collaborative Online Learning (COL) approach with the teaching of English children’s literature. The emphasis here is to enhance students’ reading and writing skills and engage their interest by an online learning environment.

The paper advances and encourages the creation of “digital natives” (Prensky 2001), and seeks both to foster authentic conditions for language learning and heat up classroom discussion. According to Lashinsky (2005), approximately 8 million students from over 2,000 colleges and 22,000 high schools use Facebook.

1. Research Purpose

The traditional Taiwanese classroom organizationally permits teachers to spend more time talking than their students. In so doing, teachers may self-disclose information about themselves at the same time that they lecture on course content. Wheeless & Grotz (1976: 47) define self-disclosure as “any message about the self that a person communicates to another.” How frequently a teacher self-discloses affects the rate of student participation on the Facebook group page. This paper proposes a “literature-based” Facebook environment to facilitate collaborative learning. Specifically, the proposed learning approach advances the following objectives in teaching literature online to

1) develop students’ problem-solving skills
2) provide online learning environment to encourage independent learning
3) recognize everyday expressions through the reading of English children’s literature
4) encourage teachers of “digital natives” (Prensky 2001) to create authentic conditions for language learning that heat up classroom discussions.

2. Computer Mediated Communication via Utility Social Networks

Computer mediated communication (CMC) has been closely linked to effective language learning (e.g., on-line social networks via mobile learning [Prensky 2005]) due to time and space constraints. Facebook has shown itself to be an effective online social network. Recent evidence as of 2012 accounts for 800 million Facebook subscribers worldwide. Stelter (2008) stresses that ubiquitous learning using Facebook and mobile devices has become normative and acceptable to large numbers of educators.

Scholars have also found that students who use CMC collaboratively with other students adopt more direct uncertainty reduction strategies (e.g., more intimate questions and self- disclosures) than those in face-to-face conversations (Tidwell & Walther, 2002).

3. Using Facebook Group Page in Teaching English Children’s Literature

This paper employs the elements incorporated in the Collaborative Online Learning environment. A situation is collaborative if students are at the same level, perform the same activities, have a common goal, and interact in a collaborative way (Lee 2009: 87). Yeh (2007: 830); in effect, these conditions, when met, propose a synchronous scaffolding environment on collaborative technical writing. Yeh (2007:830) states that “collaborative learning helps students use their own prerequisite knowledge to go beyond what they currently think.”

The paper is interested in discovering how teacher self-disclosure affects the students’ level of participation in discussions over open-ended literary questions. The research question being posed is: *How much does the teacher’s consistent self-disclosure on Facebook affect the level of student participation and ultimately teacher-student relationships?*

4. Participants

The participants were 32 undergraduates (23 first-year students, 6 second-year students, 3 third-year students) enrolled in a selective course on English Children’s Literature at a vocational College. The sample consisted of 4 males and 28 females, (age ranged from 18 to 21).

5. Qualitative data

Participants responded to open-ended questions. Levels of participation reflected participants’ perceptions of teacher self-disclosure on Facebook. The following aspects were observed: (1) the frequency of teacher-posted messages; (2) the frequency of participants who responded to open-ended questions; (3) the levels of English of the participants; (4) a student satisfaction questionnaire that was completed to evaluate student feedback regarding use of Facebook.

6. Discussion

32 aspiring EFL college students were observed over one semester for the survey. The results of the evaluation indicate that after reading continuously at least half of the novel and answer open-ended questions about it. The teacher's self disclosures/confessions of her view points aroused students' interest and generated more discussion.

A Student Satisfaction questionnaire was carried out at the end of this study. The survey procedure revealed approval of the Facebook activity, with mean scores on a 5 point Likert scale ranging from 4.71 to 4.48.

7. Conclusion

This paper discusses the use of Facebook for language learning and evaluates the effect of teacher self-disclosure on Facebook Group Page. If employed effectively, the integration of the children's literature and online collaborative learning in EFL learning has great pedagogical value. It can stimulate learner autonomy, create authentic conditions for language learning, and heat up classroom discussions that hold everyone's attention.

Acknowledgements

We thank all the people who wrote previous versions of this document.

References

[1] Lashinsky, A. (2005). Facebook stares down success. Fortune, 152, 40.

[2] Lee, C.Y. (2009). A Case Study of Using Synchronous Computer-Mediated Communication System for "Spoken English Teaching and Learning Based on Sociocultural Theory and Communicative Language Teaching Approach Curriculum. A dissertation presented to the faculty of the College of Education of Ohio University.

[3] Prensky, M. (2001). Digital Natives, Digital Immigrants. From On the Horizon (MCB University Press, Vol. 9 No. 5, October 2001) http://www.marcprensky.com/writing/prensky%20-%20digital%20natives,%20digital%20immigrants%20-%20part1.pdf

[4] Prensky M. (2005) What can you learn from a cell phone? Almost anything! The Innovate Gateway 1 (June/July). Available at: http://www.innovateonline.info/index.php? view=article&id=83 (last accessed 7 June 2006).

[5] Stevick. E. W. (1976). Memory, Meaning and Method: some psychological perspectives on Language Learning. Newbury House.

[6] Stelter, B. (2008). AOL buying No. 3 social networking site. The New York Times, Retrieved

[7] Tidwell, L. C., & Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. Human Communication Research , 28, 317-348.

[8] Yeh, S.W., Lo, J.J., Huang, J.J., & Fan, Z.Y. (2007). A Synchronous Scaffolding Environment for Collaborative Technical Writing. Lecture Notes in Computer Science, 4469, p.829840. (NSC 94-2411-H-033-010).

English for Academic Purposes in an Online Writing System

Hui-Chin YEH* **& Yi-Fang CHANG**
Graduate School of Applied Foreign Languages, National Yunlin University of Science and Technology, Taiwan
hyeh@yuntech.edu.tw

Abstract: Novice and non-native postgraduate students always face a number of difficulties in the complexity in academic English writing. However, they are seldom taught the effective strategies to regulate their writing processes. This research intends to develop an academic writing course incorporating an online writing system for postgraduate students to develop their metacognitive genre awareness. In the four metacognitive stages, the postgraduate students are assigned different kinds of writing tasks. All of their writing products, such as drafts of thesis proposal, final version of thesis proposal, self-reflections, are recorded and observed in the online writing system.

Key words: Online writing system, EAP, online learning, academic writing

1. Introduction

English proficiency has become a key factor for learners to succeed in pursuing their higher education. It is reported that more than 65% of international academic journals are written in English (Johns, 1993). In response to the needs of English learning, English for Academic Purposes (EAP) courses are dramatically increasing. A genre-based approach is an effective way to improve writers' genre-awareness in EAP instruction (Fortanet, Posteguillo). It explicitly and systematically explains to writers how languages work in a social context, how texts are structured, and why they are written in the ways they are. Hyon (2001) concluded that ‘‘genre-knowledge gained through explicit instruction can be remembered by EAP students over an extended period of time and facilitate aspects of L2 reading and writing" (p. 434). Swales' (1990, 2004) CARS (Create a Research Space) model is regarded as the most influential EAP genre analysis framework. It sheds light on the structural organizations and the academic writing conventions. According to Swales (1990), any introduction to research articles (RAs) consists of three obligatory moves, which include move 1—establishing a territory (establishing the topic), move 2—establishing a niche (justifying the present study), and move 3—occupying a niche (describing the present study). Each move can be further divided into a number of steps, which are hierarchically organized in distinct sections.

Writing Master's thesis is a demanding task that requires sophisticated linguistic knowledge and is a tough process for novice writers. Therefore, students should be taught the effective metacognitive strategies to regulate their writing processes. Wenden (1991) claimed that metacognition plays an important role in effective learning because learners are equipped with the abilities to *plan, monitor, evaluate*, *and revise* their cognitive and mental activities. Xiao (2005) indicated that the differences between mature and immature writers lie in the awareness of metacognition. EFL unskilled writers know little about writing knowledge and strategies, while skilled writers extensively undergo the processes of self-regulation and metacognitive control (Flower, 1990; Kellogg, 1996). Moreover, writing is not merely considered as an individual work; instead, it is supported by feedback from peers who effectively review texts, see logical gaps, identify problems in the organization, and most importantly, improve the writing products (Vass, Littleton, Miell & Jones, 2008). The aim of this study is to identify how L2 students develop their genre awareness in an online writing system.

2. Design of the Academic Writing Course and the Online Writing System

The design of both the academic writing course and the online writing system is based on the theory of metacognition, which includes four processes—planning, monitoring, evaluating, and revising. Each student's metacognitive processes and writing products can be preserved and observed in the online writing system. In the planning stage, the postgraduate students are required to analyze three moves in research paper introductions according to Swales' CARS Model. They identify the moves and steps and denote them numerically at the end of each sentence of the research paper in the function *Post a New Essay*. In the monitoring stage, they have to apply the generic features that they have discovered and learned from analyzing research paper introductions to compose their own thesis proposals. Then, they post their drafts and final thesis proposal in the function *Post a New Essay*. In the third stage of evaluating, the students are asked to edit peers' proposals in terms of the generic features and the text organization. In the function *Suggestion to Global Revision*, a list of suggestion types in the text organization is provided, and the definitions and examples for each suggestion type are also provided for students to look up. In the revising stage, the students have to decide which advice provided by their peers should be accepted or rejected in order to revise their proposals. They can look up peers' comments in the function *View Different Versions* and decide whether or not to take them. Both the local revisions (e.g. grammatical mistakes) and the global revisions (e.g. logic and organization) are shown in the system.

Conclusion

In the context of EAP, the genre-based approach provides L2 learners with an explicit chance to undertake writing tasks and to cultivate genre awareness. Genre awareness is closely connected with metacognitive strategies, which can help learners *plan, monitor, evaluate, and revise* their learning in a way that directly improves writing performance. Moreover, peer review in CSCL provides students with the opportunity to meaningfully use language to write to real audiences. In order to investigate how L2 students develop their genre awareness, the academic writing course and the online writing system are designed based on four stages of metacognition.

References

[1] Flower, L. (1990). The role of task representation in reading to write. In Flower, L., Stein, V., Ackerman, J., & Kantz, M. (eds.) *Reading to write: exploring a cognitive and social process.* Oxford: Oxford University Press. 35-75.

[2] Fortanet, I., Posteguillo, S., Palmer, J. C., & Coll, J. F. (Eds.). (1998). *Genre studies in English for academic purposes.* Castello, Spain: Servei de Publicacions-Universitat Jaume I, Castello, Spain.

[3] Hyland, K., & Hamp-Lyons, L. (2002). EAP: issues and directions. *Journal of English for Academic Purposes, 1*, 1-12.

[4] Hyon, S. (2001). Long-term effects of genre-based instruction: a follow-up study of an EAP reading courses. *English for Specific Purposes, 20*, 417–438.
[5] Johns, A. (1993). English for Specific Purposes: International in scope, specific in purpose. In S. Silberstein (Ed.), *State of art TESOL essays: Celebrating 25years of the discipline* (pp. 115-153). Alexandria, VA: TESOL.
[6] Kellogg, R. (1996). A model of working memory in writing. In Levy, C. & Ransdell, S. (eds.) *The science of writing: theories, methods, individual differences, and applications.* Hillsdale, NJ: Lawrence Erlbaum Associates.57-72.
[7] Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior, 19*, 335-353.
[8] Shaw, P., & Liu, E. (1998). What develops in the development of second-language writing? *Applied Linguistics*, *19*(2), 225-254.
[9] Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.
[10] Swales, J. (2004). *Research genres: explorations and applications.* Cambridge, UK: Cambridge University Press.
[11] Vass, E., Littleton, K., Miell, D. & Jones, A. (2008). The discourse of collaborative creative writing: peer collaboration as a context for mutual inspiration. *Thinking Skills and Creativity, 3*, 192-202.
[12] Wenden, A.L. (1991). *Learner strategies for learner autonomy.* London: Prentice-Hall International.
[13] Xiao, Y. (2005). *The EFL beginning writers' perception and metacognitive knowledge of English writing: A study on the freshmen at Sun Yat-Sen University.* Unpublished paper. Guangzhou: Sun Yat-Sen University.
[14] Yang, Y. F. (2011) A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology*. *42*(4), 687-700.

Automatic Exercise Generation in an Essay Scoring System

Itziar ALDABE[a], Montse MARITXALAR[a*], Olatz PEREZ DE VIÑASPRE[a], Larraitz URIA[b]
[a]*IXA NLP Group, University of the Basque Country UPV/EHU, Spain*
[b]*IKER (UMR5478), IKERBASQUE*
*montse.maritxalar@ehu.es

Abstract: In this paper, we present an automated essay scoring system for language learners to improve their writing skills. The system combines statistical and Natural Language Processing (NLP) techniques to compute the evaluation measures of the essays. It also incorporates a module to automatically generate multiple-choice questions (MCQ) to test the language level in Basque. The automatic generation of the options that constitute the MCQs is based on the linguistic information compiled in the rules written for a Basque grammar checker, which detects, among others, determiner errors.

Keywords: generation of MCQs, evaluation of essays, NLP in ICALL

Introduction

In the last two decades several research has been developed on computer-based automated essay scoring (AES) systems [4] which provide students with feedback to improve their writing. Although some research communities have criticized this type of system, recent studies [2] demonstrate that AES systems help improving the school performance. One of the advantages is that they measure all essays using the same scoring model. This way, AES systems offer “objective” data to improve on the controversial task of manual essay evaluation.

The essay scoring system presented in this paper combines statistical and Natural Language Processing (NLP) techniques to compute the evaluation measures of the essays. The application, named IAS, uses a client-server architecture [1]. On the one hand, the client interprets the results calculated by the server in order to decide the evaluation measures that the application will give to the user as the result of the evaluation. On the other hand, the server includes a request manager which triggers its language modules depending on the language requested by the client. Those language modules are responsible for the linguistic process of the essay. By means of this linguistic process the texts are analysed and students' errors are detected. Both the interface and the application itself are multilingual. So far, we have integrated NLP tools for two very different languages, such as Basque and Spanish.

In this article, we present an improvement of the system, where a new module to automatically generate MCQ questions for Basque has been incorporated. The automatic generation of the options that constitute the MCQs is based on the linguistic information compiled in the rules written for a Basque grammar checker, which is also integrated in IAS.

1. Automatic Generation of MCQs to Improve on the Use of Determiners

The experience acquired by teachers when teaching a language is a valuable source when establishing the criteria for creating tests. In contrast, a learner corpus containing a collection of errors is an alternative way of defining the heuristics required to generate learning material. In our opinion, the use of learner corpora for creating exercises offers interesting advantages: (i) error analysis studies on learner corpora show us a realistic picture of the written output of learners; (ii) they provide us with objective data about the typical error types; and (iii) the development of grammar checkers based on corpora to detect and diagnose errors provides us with information on error diagnosis that can be used for the automatic generation of learning material.

Determiner errors are relatively common among Basque language learners. The Basque learner corpora collected in the last years have provided us with enough data to perform a deep study of this phenomenon. Based on that information, we have created a grammar to automatically detect some types of determiner errors using the Constraint Grammar (CG) formalism [3]. Based on the linguistic information compiled in the rules of the grammar checker, IAS generates MCQ tests to check the use of determiners in Basque.

The next example presents an automatically generated MCQ to deal with the R_DET error type (repetition of the determiner). The source sentence, which is a text written by a learner of Basque, contains a R_DET error (option d) and the system generates the rest of the options, i.e. the correct answer (option a), and the two distractors (options b and c).

Nire bizitzaren .... orain dela 5 urte gertatu zen (*... in my life took place 5 years ago).*

a) egun zoriontsu **bat** (**one** happy day)

b) egun**a** zoriontsu**a bat** (**a one** happy the day)

c) egun**a** zoriontsu **bat** (**one** happy the day)

d) egun zoriontsu**a bat** (**the one** happy day)

In order to create the above presented MCQ example, the system first detects two determiners in the same DP in the source sentence (option d). Then, it deletes the determiner from the adjective (zoriontsua - zoriontsu) to generate a correct DP (option a). The new DP is set as the correct answer of the MCQ. In order to generate new errors, i.e., the distractors, different strategies are applied: (i) to add determiners to various words in the same DP; (ii) to delete the determiner to all the elements in the DP; (iii) to change the declension mark of the phrase; (iv) to change the order of the words inside the phrase, etc. In the example above, both distractors are created by adding the definite article -a to different words inside the DP (options b and c).

2. Experiment

This experiment deals with the correct use of determiners in Basque by means of the use of IAS. Here we present the results related to the R_DET error type, since it is one of the most common error. A linguist first supervised the automatically detected errors in a text written by a Basque language learner. Then, she manually added new error instances in order to create different types of determiner errors. After that, the system automatically generated the MCQs embedded in the text.

The linguist was responsible for creating the different types of R_DET errors which were previously observed and detected in the Basque learner corpus. She had to establish different examples of R_DET errors, obtaining the highest level of casuistry. In total, she marked 15 Determiner Phrases (DPs) out of a total of 68 within the text, that is to say, 22.05% of the DPs. Once erroneous DPs were defined, the system generated the corresponding correct answer for each of them as well as two more distractors. The test generated for the experiment was answered by thirty low-language-level learners from Argentina, who took part in this experiment on-line. Those students were not therefore the writers of the Basque learner corpus. The experiment was carried out with low language level learners since the most R_DET errors have been detected in the mentioned corpus (proportionally and with a statistically significant difference) at this level.

Based on the fact that 60% of the students (18) correctly answered more than 11 MCQs, we could say that the generated test was not difficult for most of them. There was a significant difference in percentages when comparing the number of errors made by the Argentinian students when filling the MCQs and those errors detected in the Basque language corpus. Based on the results of the students who answered incorrectly at least 4 MCQs, the percentage of R_DET errors per DP of Argentinian students was 5.88% or higher. In contrast, the percentage in the low-level sample of the Basque language corpus was 0.77%. These results can lead us to consider (i) that offering this type of test could cause the emergence of new errors that students would not produce in an open task; or (ii) that perhaps foreign language learners of Basque make different errors than those made by second language learners of Basque.

3. Conclusions

IAS is an automated essay scoring system for students of Basque to improve their writing. The system is the core of the first bilingual web application developed to handle Basque and Spanish essays. IAS incorporates a module which automatically generates MCQ to test the Basque language level. The automatic generation of the options that compose the MCQs is based on the linguistic information compiled in the rules written for a Basque grammar checker, which is also integrated in the system.

The results of the experiment leave open a new research challenge related to the type of exercise that should be proposed to improve the writing of language learners. In addition, we should study the type of learner corpus to be used as the basis for the generation of language exercises.

Acknowledgements

This work has been supported by Olatz Perez de Viñaspre PhD grant for the Basque Government and the OpenMT-2 project (TIN2009-14675-C03-01).

References

[1] Castro-Castro, D., Lannes-Losada, R., Maritxalar, M., Niebla, I., Pérez-Marqués, C., Alamo-Suarez, N.C., & Pons-Porrata, A. (2008) A Multilingual Application for Automated Essay Scoring. *Lecture Notes in Advances in Artificial Intelligence - LNAI 5290* - Springer New York 243-251.

[2] Grimes, D. y Warschauer, M., (2010). Utility in a Fallible Tool: A Multi-Site Case Study of Automated Writing Evaluation. *Journal of Technology, Learning, and Assessment,* 8 (6).

[3] Karlsson F., Voutilainen A., Heikkila J. & Anttila A. (1995) Constraint Grammar: a Language Independent System for Parsing Unrestricted Text. Mouton de Gruyter.

[4] Landauer, T. K., Laham, D., Foltz, P. W.: Automated Essay Scoring and Annotation of Essays with the Intelligent Essay Assessor. In: Mark D. Shermis and Jill C. Burstein (Eds.). Automated Essay Scoring: A Cross Disciplinary Perspective, pp. 87-112. Mahwah, NJ: Lawrence Erlbaum Associates (2003)

Designing for Model Progression to Facilitate Students' Science Learning

Daner SUN [a*], Chee-Kit LOOI [b]
[a,b.] *Learning Sciences Laboratory, National Institute of Education, Singapore*
*daner.sun@nie.edu.sg

Abstract: WiMVT (**W**eb-based **i**nquirer with **M**odeling and **V**isualization **T**echnology) is designed as a learning system combining guided inquiry, modeling and visualization with the social interaction. In the paper, we first present the design rationale of the system, briefly describe the main functions, then discuss the features supporting model progression in science learning. Following it, we describe a pilot study of WiMVT implementation in the secondary science class. The data analysis demonstrates the pedagogical value of WiMVT on students' conceptual understanding, and indicates that appropriate peer feedback can promote students' model elaboration in the modeling activities.

Keywords: WiMVT, model progression, collaboration, peer feedback

Introduction

Due to the learning effectiveness of the models and modeling tools for science education, in particular, the computer-supported Model-based Science Learning Environment (MbSLE), a number of researchers have invested great effort in building and implementing MbSLEs in science learning (e.g. Model-it, ModelingSpace, NetLogo, WISE). Besides modeling tools, these applications may have design elements such as curriculum materials, proposed pedagogy (e.g. inquiry, CSCL, model progression) and communicative tools. However, few of them have been designed with the integration of all these design elements. For example, most of them are unable to support online collaborative modeling; some of them do not allow for importing of multiple visual representations; and some of them do not facilitate model progression in pupils. With the intention of creating an innovative application for secondary school students to acquire sophisticated understanding of scientific conceptions, develop critical learning skills, we have developed a web-based science learning environment named WiMVT (**W**eb-based **i**nquirer with **M**odeling and **V**isualization **T**echnology, http://www.sstlsl-wimvt.sg/wimvt). It is designed as a system in which guided inquiry, modeling and visualization, and social interaction, are integrated - which is unique among existing science learning environments (Sun & Looi, 2012).

The work reported in this paper focuses on one of core features of the system: model progression. In this paper, we first introduce the design rationale and the basic functions of the system. Then we emphasize the feature of model progression in the system. Finally, we present some results from a pilot study to illustrate the educational value of the model progression with WiMVT system. The study is used to answer the research questions below:
1) How to integrate a science learning environment featuring model progression into a real learning context?
2) What is the learning efficacy of the model progression for students' conceptual understanding in science?
3) Does the peer feedback promote students' modeling progression in the WiMVT lessons?

1. The Design Rationale of WiMVT System

The learning efficacy of the model-based inquiry in science education has been demonstrated in many studies (Schwarz & Gwekwerere, 2007). Its generic learning pattern can be summarized as question, hypothesis, plan, investigation, model, and conclusion (Bell, et al., 2009). The pattern serves as the guide for the form of WiMVT inquiry cycle. Additionally, informed by the principle of POE (Predict-Observe-Explain) adopted in science class in some Singapore schools (White & Gunstone, 1992), we propose a phase called Pre-model (to explicate students' initial ideas) with the corresponding phase: Model (to explicate students' post ideas) in the inquiry. The design guides students to present their predictions of the science phenomena before investigation, and then to verify their predictions through investigation, thus ultimately improving their understanding of the science phenomena. Hence, the main purpose of embedding model progression by including Pre-model and Model into guided inquiry is to help students elicit and expose their prior knowledge through pre-models and to elaborate their models in Model phase after a series of activities. Afterward students refine their understanding and seek validation them in the Reflect and Apply phrases. Finally, a revised model-based inquiry cycle incorporating eight phases is created: Contextualize, Question & Hypothesize (Q&H), Pre-model, Plan, Investigate, Model, Reflect, and Apply (WiMVT inquiry cycle).

WiMVT is a complicated system, so we employ a standard approach: the Rapid Application Development (RAD) for the system development in consideration of the research condition. The development process consists of five short development cycles which mostly involve: design → discussion → adopted features development→ discussion & usability test → redesign. At each stage, consultants and collaborators from different research including science education, computer technology, and educational technology areas are invited to give feedback and comments on the design of the system. Subsequent revisions are made based on the feedback. During the whole process, to verify the validity of the system at each stage, usability testing is conducted to collect data for revising and improving the design and development. Up to date, we have finished several usability tests and two pilot studies. The existing WiMVT system has been revised and improved based on the usability report and feedback from trial implementation in the pilot school.

2. The Introduction of the WiMVT System

The WiMVT system operates via the Internet and is accessible through a general web browser. Figure 1shows the work flow of the system as a pedagogical scenario. The number tags (A-D) corresponds to the four operational procedures of the system, see discussion below.

A. Establishing the project: the authoring tool in the teacher module supporting the design of the project[1]: 1) Present brief project description, learning objectives, and tasks for Home; 2) Edit content accompanying various types of information (e.g. videos, images, simulations) for Contextualize; 3) Design questions for Q&H, and assign tasks for Plan, Pre-model and Model, Reflect, and Apply. 4) Insert simulations together with guided questions into the Investigate tab where students are required to do certain virtual experiments. Finally, the teacher configures the students' groups in the Group Management section and assigns the projects to the students.

[1] The teacher module consists of six sections: My Profile, My Subjects, Project Management, Solutions Review, Simulation Library, My Mailbox. The section of Project Management provides an authoring tool for teachers to design and edit the instructional content.

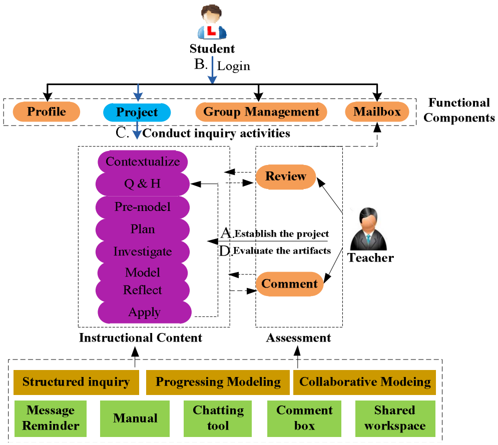


Figure 1. The work flow of WiMVT system

B. Logging onto account: the student module comprises four functional components: My profile, My project, Group Management and Mailbox. The general information of the assigned project can be retrieved in My Project section after the students log into the system with their accounts and passwords.

C. Conducting inquiry activities: The typical work surface in student module is illustrated in Figure 2. It is split into four panes: shared workspace which holds the textual information or tools associated with each phase, status of group members (online students' username is visible at all times), name list of group members, and a chat box.

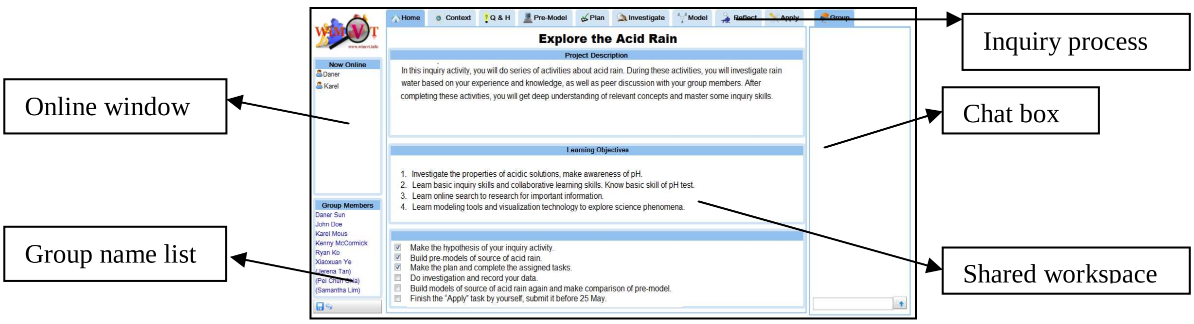


Figure 2. The interface of students' work session

The process of WiMVT inquiry can be briefly described as below: After accessing the "Home" tab, the students are guided to engage in a series of learning activities: students formalize their hypothesis of the questions in Q&H; create pre-models of scientific concepts they will learn based on their prior knowledge when in Pre-Model; design the plan in the Plan and then collect and discuss the data in Investigate. Additionally, they can also engage in the manipulation of several simulations to do virtual experiments, as well as answer the guided questions. They are then asked to revise their pre-models through peer review and discussion in Model, and to reflect upon artefacts being built when getting into the Reflect phase.

D. Reviewing and commenting on artefacts: The teacher thereby can access the artefacts and interactions generated by students while navigating in Solutions Review, and comment on each student or each group's hypotheses, plan, investigation report, pre-models and models, reflective content, as well as their responses in Apply if any. Thus, the system supports both formative and summative assessments.

3. Model Progression in the WiMVT System

The model progression is a way to present models in increasing complexity gradually through expanding the number of components or the levels of relations among variables of models (Swaak, et al., 1998; Mulder,et al., 2011). As indicated earlier, the design elements of the system support model progression. Specifically, a sketch tool serves as a drawing-based modeling tool is designed to assist low-ability students' creation of the models both Pre-model and Model (Lerner, 2007). Compared to the drawing-based tool, the qualitative modeling tool and quantitative modeling tools in Model provide more opportunities for students to construct high level scientific models. In the system, when students define objects and establish relations between variables of a qualitative model, the modeling functions were mainly executed as an invisible simulation engine for processing relationships which are specified in the form: 'If A increases, B increases' (Avouris, et al., 2005). In the quantitative modeling scenario, the relations are established via precise mathematical forms involving variables. In this way, modeling thereby can be progressive because the students can start from simple (novice) models to complex (expert) models using the sketch tool. Otherwise, they can work from more qualitative modeling without defining formulas and then get into the stage of more quantitative modeling when figuring out the formulas finally. Moreover, synchronous collaboration in Pre-model and Model is facilitated via the shared workspace and a chatting tool. It means that students can co-construct a model in real-time, and then modify and elaborate it through online peer discussion.

4. Research Design and Methods

4.1 Participants

In this study, 46 students from two classes were randomly selected from a junior secondary school in Singapore. A female physics teacher with 9 - year teaching experience conducted the class. She had participated in a series of teacher-researcher working sessions of WiMVT project, and thus had some good understanding of the system. The computer facilities in the school were excellent, and each student owned and used a Macbook for daily lessons in the various subjects.

4.2 The design of WiMVT lessons

The WiMVT lessons were co-designed by WiMVT team, science teachers and collaborators. The classes studied "Current Electricity and D.C. Circuit". The topic was divided into 8 50-minute lessons, in which four lessons were incorporated by the system. Table 1 summarizes the lesson flow and relevant information.

The students drew a model of a circuit needed to run a quiz show for 3 teams of participants, and to point out the direction of the current flow as well in lesson 1. The teacher reviewed the pre-models and identified the major misconceptions amongst the students. Students' initial ideas of simple circuits then further explored and elicited in lesson 2 through doing hands-on experiments of connecting possible circuits. During lesson 3 and 4, the students interacted with three levels of PhET simulations, as well as answered guided questions individually. After obtaining new understanding through investigation, they were guided to Model phase to elaborate initial models drawn in lesson 1, and to reflect on their conceptual changes in Reflect. In this pilot study, the qualitative and quantitative modeling tools are not incorporated into the system. The students mainly worked with the sketch tool to create models. They were encouraged to provide online peer feedback for models creation and elaboration within the group members. Before class time, the students were asked to log onto WiMVT at home to sufficiently familiarize themselves with the system. The teachers integrated the instructional content in the system and managed the grouping of the students. As 23 students in each class were divided into 8 groups with heterogeneous, they mainly worked in triads (one group worked in dyads each class).

4.3 Data source and data analysis

The study aimed to examine the pedagogical value of model progression of the system in students' science learning, as well as the impact of peer feedback on students' modeling performance. We used software to capture the screen activities with the intention of validating data analysis. Videotaped recordings of the teacher and students' interactions were used to identify patterns of change for triangulation purposes. One audio recorder was directed at each of 8 groups in both of classes. After all sessions, we interviewed the teacher and 16 students using a semi-structured interview protocol for approximate 20 minutes. In data analysis, all videotapes and audios were transcribed to examine students' performance in WiMVT activities. The students' peer feedback generated in chatting log during the modeling process was also saved, identified and transcribed. It was used to investigate the relation between students' model quality and their peer feedback. The results were subsequently verified by cross-referencing collected data.

5. Results

5.1 The progression of the models' quality

We used the quality of models as the indicator for evaluating students' modeling performance. Based on the literature review, we classified the quality of models into three levels: 1) High Quality Models (HQMs) are the model representations reflecting appropriate descriptions of science conceptions that involve components with basic properties, and depicting interactions between variables of components. 2) Medium Quality Models (MQMs) are the model representing partially accurate descriptions of scientific conceptions, in which some of appropriate components of models are included in the models. 3) Low Quality Models (LQMs) refer to the model representations which contain inaccurate descriptions of all models components, they are usually at the level of the scribble drawing (Grosslight, et al., 1991). In this study, we collected a total of 11 models in Pre-model phase and 14 models in Model phase. The outcomes of the evaluation of students' models in Pre-model and Model phases are depicted in Table 1[2].

Table 1. The number of different models in the Pre-model and Model phases

Quality of models	Pre-model stage	Model stage
LQM	1	0
MQM	5(without current direction) 3(inaccurate current direction) 1(short circuit)	6 (without current direction) 1 (inaccurate current direction) 3 (broken circuit with current directions)
HQM	1	4

Overall, the quality of students' pre-models and models were at the level of MQMs, around 82% and 71% respectively. Specifically, 45% of models in Pre-model phase presented right representations with components of bulbs, switches and batteries, but failed

[2] As the models were the products of groups' work, the number of models equals the number of groups.

to define the current flow direction, while the rate decreased to 42% in Model phase. In comparison with the 3 groups who drew the incorrect current flow direction, only one group exhibited the same mistake in Model phase. In Model phase, 3 groups defined the right components of models although they drew the current in the broken circuit, while the models of 4 groups achieved the level of HQMs. These findings indicated students' better performance as reflected by the better quality of models generated in the Model phase. The increase in HQMs further demonstrated students' progress in the understanding of core concepts and modeling skills. The students' responses on the process of understanding the circuit in the Reflect phase demonstrated their progressions as well:

- Students A: I used to think that short circuits are very complicated, but they are not. In addition, I thought that parallel circuits have different current for each bulb. But now, I think that for parallel circuits, the bulbs have the same brightness as the same amount of current is being flowed through it. Only, when the switch is closed, then the electrons can starts flowing. Bulbs in series circuits have lower brightness than bulbs in parallel connection.
- Student B: I feel that our design is correct as it is in parallel connection of the bulbs with a switch connected to it. Closing one switch will cause its corresponding light bulb to light up.

5.2 The correlation of models quality and peer feedback

Online peer feedback is particularly advantageous, due to the possibility of a less stressful and intimidating working environment from the lack of face-to-face interaction, which may promote students to be adventurous and be more involved (Guardado & Shi, 2007). In the study, students were encouraged to build, revise and elaborate their models through receiving peer feedback from their team members both in the Pre-Model and Model phases. The peer feedback coding instrument was developed based on the principles of good feedback theory and practice (Nicol & Macfarlane-Dick, 2006). It consisted of five categories: A. task-oriented (clarifies the task specificities), B. knowledge-oriented (provides necessary information on how to solve a problem), C. strategy-oriented (provides strategic plans to derive answers in the best way), D. assessment-oriented (provides constructive comments on the work produced) feedback, and E. affection-oriented (provides comments with intentions to improve motivations). Here are some examples of the peer feedback from the transcription of the discourse of a group doing the modeling, with their coding:

- *Category A +Category B: you press undo and draw the bulb.*
- *Category B: Just put more batteries to make it (electromagnet) stronger.*
- *Category C: Let us first draw and then think it.*
- *Category D: Actually it is possible. But maybe need more batteries.*
- *Category E + Category D: Nice drawing. I will draw the line.*

In the data analysis, the Pearson coefficient was computed to assess the relationship between model scores and the students' peer feedback. Thus, we scored 25 models from the range of 0 -100 according to the components of models and its relations. The LQM score was less than 60, the MQM score was between 60 and 80, and the HQM score was between 80 and 100. We calculated the quantity of peer feedback that happened at each group, as well as the number of each type of peer feedback. The results indicated that there appeared to be an upward trend, namely, as the amount of peer feedback increases, the higher the scores of the models drawn. The statistical analysis with the Pearson's r=0.972, p=0 (at the level of 0.01) reflected a strong positive correlations between students' model scores and the quantity of feedback. Table 2 presents the respective correlation for the five categories of peer feedback and the model scores.

Table 2. The correlations between peer feedback and models' score

Measure	A: Task	B: Knowledge	C:Strategy	D:Assessment	E:Affection
Pearson Correlations	0.839**	0.280	0.574	0.941**	-0.739
Sig.(2-tailed)	.000	.158	.253	.000	.261
N (Model)	25	25	25	25	25

** Correlation is significant at the 0.01 level (2-tailed)

The findings suggested that category A, B, C, D were positively correlated with the model scores. Significant correlation existed between category D and model scores, r=0.941, p =.000. Thus, the higher the quantity of category D, the higher the models' scores, or vice versa. There was also significant correlation between category A and the model scores (r=0.839, p=000). As for knowledge and strategy-oriented peer feedbacks, the correlations (r=0.28, p = .158; r=0.574, p=*.253*) for both measures with the model scores are not significant and they are weakly correlated. This could imply that knowledge and strategy-oriented feedbacks may be less related to the scoring of the model. Also, the correlation between affection-oriented peer feedback and scores is negative, $r(2) = -.739$, p = .261. We would like to explore these further in future empirical studies. Thus, when students are working with the modeling tool in the system, it is suggested assessment-oriented peer feedback and task-oriented feedback be provided by team members or otherwise, using words like:

- *Please don't forget that the switch is one of the objects.*
- *There appears to be a problem in that part, do you mind if you check it again?*
- *I think the two objects that you linked up may be incorrect.*

5.3 Voices of the teacher and students

The teacher and students expressed an overall positive attitude toward the WiMVT implementation in the science class. The teacher had a better understanding of the lessons which could be designed to leverage on the affordances for WiMVT inquiry. She concluded that 1) lesson plans should be adopted for best fit with WiMVT inquiry and the instruction should optimize the core features of the system; 2) The explicit inquiry mode was a good scaffold to guide students' learning activities; 3) Students were suggested to do individual modeling in the Pre-model phase, because they had different initial ideas; 4) In the Model phase, students were encouraged to converge to a solid understanding whereby they could present in one consensus model through the co-constructive way. The students agreed on that the WiMVT learning activities were more interesting and engaging compared what they had used previously. They pointed out that the small group's collaboration provided more opportunities to do tasks in the system, and the synchronized collaboration could help them finish the task faster. The modeling process directly within the system could make drawing more convenient and less time consuming. They thought that they enhanced their understanding of electrical circuits bring taught in the lesson through the comparison of pre-models and models, as well as a reflection phase to concretize the thinking process.

Conclusions

This pilot study on design and implementation of WiMVT lessons addresses the research question on how to design the WiMVT lessons, and the results demonstrated that WiMVT exhibited some value in aiding students' conceptual understanding. Specifically, the students' model progression can be achieved through the design elements of the combination of Pre-model and Model phrases in the inquiry. The preliminary finding that the quantity of peer feedback varies with the quality of the models can help inform the design of collaboration into such an environment. The students are particularly encouraged to heed assessment-oriented feedback in the collaborative activities. The teacher and students' voices suggested that more engagement needs more appropriate instructional support such as guiding students' collaboration, scaffolding students' modeling process. In summary, we provide an illustration of the WiMVT system that supports flexible collaborative students' model-based inquiry. We believed that the inquiry with WiMVT will create unique educational opportunities for students' science learning. In the future work, the investigation of students' conceptual understanding, collaborative skills, inquiry skills and reflective thinking skills will be the main avenues we will pursue with longer-term and larger scale use of the WiMVT system.

Acknowledgements

This research is supported by the project: Fostering Collaborative Inquiry Modeling and Visualization Practices in Secondary Science Learning (WiMVT) funded by National research Foundation in Singapore (Project #: NRF2009-IDM001-MOE-019 project).We thank Baohui Zhang for his initial conception of and contributions to this work. We would also like to thank WiMVT team members and our collaborators: Gao Shan, Charles Chen, Fu Weikai, Evelyn Teo, Karel Mous, Pey Tee Oon, Sze Ghee Tan, and Kin Chuah Chan for working with us on the project.

References

[1]Avouris, N., Margaritis, M., Komis, V., Saez, A., & Meléndez, R. (2005). Modelingspace: interaction design and architecture of a collaborative modeling enviroment. Presentation for Comodeseminar, Patras, December 2005.

[2]Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2009). Collaborative inquiry learning: Models, tools, and challenges. International Journal of Science Education, 32(3), 349-377.

[3]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.

[4]Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. Computers and Composition, 24, 443-461. doi:10.1016/j.compcom. 2007.03.002

[5Lerner, N. (2007). Drawing to learn science: Legacise of agassiz. Journal of Technical Writing and Communication, 37(4), 379-394.

[6]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(5), 614-624.

[7]Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199-218.

[8]Schwarz, C. V., & Gwekwerere, Y. N. (2007). Using a guided inquiry and modeling instructional framework (EIMA) to support preservice K-8 science teaching. Science Education, 91(1), 158-186.

[9]Swaak, J., Joolingen, W. R. V., & Jong, T. D. (1998). Supporting simulation-based learning:the effects of model progression and assignments on definational and intutive Knowledge. Learning and Instruction, 8(3), 235-252.

[10]Sun, D.E, & Looi, C.K. (2012). Designing a web-based science learning environment for model-based collaborative inquiry. Journal of Science Education and Technology, doi: 10.1007/s10956-012-9377-9.

[11]White, R., & Gunstone, R. (Eds.). (1992). Probing Understanding. London and New York: The Falmer Press.

Evolution of Literacy in Software Functions by Creation of Storytelling

Isao MIYAJI[a*]
[a]*Faculty of Informatics, Okayama University of Science, Japan*
*miyaji@mis.ous.ac.jp

Abstract: "Creation of storytelling" using PowerPoint was conducted so as to raise computer literacy and to foster the students' self-understanding. They were required to draw a figure using Excel and stick it on a report. They were expected to write their reports using Word. The pre and post literacy in functions of the software of three kinds was investigated to know the computer literacy which students were able to raise by "creating the storytelling". As a result, the pre literacy in functions of PowerPoint was lower than those of Word and Excel. After this practice, the post literacy in the functions of PowerPoint significantly became higher. This paper will report that students became able to utilize the functions of PowerPoint as same as those of Word and Excel.

Keywords: Storytelling; slide-show storyboards; literacy in software; peer assessment; self-understanding

Introduction

Creative activities that produce works that inform and entertain people by describing real and imaginary events, using graphics, narration, and music are called storytelling [1]. In digital storytelling, still pictures such as photographs, figures, and drawn pictures are displayed sequentially to create a storytelling and narrated. Digital stories can be easily reconstructed, and producers can distribute a story to many people through the Internet.

A practice method or a purpose is reported a lot in until now about storytelling [2], [3], [4], [5], [6]. Sadik calls digital storytelling a meaningful technology-integrated approach [2]. Effects of creating digital storytelling are reported as follows. Still pictures are easy to handle for the producers of such assignments, and students can reflect upon memories or what they have learned through reviewing still pictures. It has been reported that the practical class improved the quality of their text narratives through the practice of digital storytelling [7]. It has been reported that storytelling task showed effect of similar patterns to the other tests in written language comprehension [8].

When persons used WBT as a supplementary material, Hirose et al. have reported that the person who has weak point awareness toward PC operation used the WBT for the number of fewer times as a study on the software literacy [9]. In addition, Yamagishi et al. reported that achievement degree of operation for the applied software is low for persons who have uneasiness about PC utilization [10].

The literacy of word processor, spreadsheet and presentation becomes required as a member of society. The literacy of presentation is inferior to the literacy of word processor and spreadsheet about students in our department. We made a study on designing and devising a class so that these three kinds of literacy became the same degree. We adopted evaluation and modified activities through creating storytelling using PowerPoint. We told meaning to revise it and method to convey oneself idea to the other persons. As a summary of the classes, students were required to write the report which included consideration for the contents of the work, self-evaluation, peer-evaluation after filling in them. It was reported that the blended class which utilized e-learning inside and outside the class was effective [11]. We conducted the blended class which utilized e-learning under such contents. We investigated the literacy in the software function in pre and post time to know the literacy in a computer as an effect of this class. In this paper, we will analyze them and report knowledge obtained from them.

1. Instructional Design and Method

The target subject in this study is called an information science experiment, consisting of three hours per week as one of the compulsory subjects in the second semester at the department of information science in a university. The content which the author was in charge of was “creating digital storytelling.” Students are separated into three groups. Each group creates the storytelling for four weeks. Each class session was 180 minutes long, and the class proceeded according to the plan shown in Table 1. The themes of each group are different as shown in Table 1. Each group carried out the class by four weeks. The first, second and third groups created each theme of storytelling from 1 to 4, from 5 to 8, and from 9 to 12 weeks respectively. The class was conducted by a teacher and a teaching assistant. After explaining the activity contents of the day at the beginning of each class, the teachers walked around the classroom and responded to questions as needed. The number of students attending a lecture of three groups was 63 persons in total, that is, 21, 22 and 20 persons respectively.

Table 1 Lesson plans

<table>
<tr><th rowspan="2">Week</th><th colspan="3">Theme</th><th rowspan="2">Experiment contents</th><th rowspan="2">Distributed survey sheet and deliverable in the session</th></tr>
<tr><th>first group</th><th>second group</th><th>third group</th></tr>
<tr><td>1</td><td rowspan="4">Self-understanding</td><td rowspan="4">Children's story</td><td rowspan="4">Future course</td><td>Submitting a story, Creating story slides</td><td>Attitude related to abilities (pre)</td></tr>
<tr><td>2</td><td>Creating story slide, Narating a story, Self-evaluation</td><td>Complete story slide, Self-evaluation 1</td></tr>
<tr><td>3</td><td>Viewing 1, Peer evaluation 1, Modifying slides, Self-evaluation 2</td><td>Peer evaluation 1, Modified slides, Self-evaluation 2, Report 1</td></tr>
<tr><td>4</td><td>Viewing 2, Peer evaluation 2, Self-evaluation 3</td><td>Peer evaluation 2, Self-evaluation 3, Report 2, Attitude related to abilities (post)</td></tr>
</table>

1.1 Purpose of Classes

One of the purposes of this class is to heighten students' literacy in computer and self-expression through creative activities. The literacy in computer means that the PC can be smoothly operated using functions of Word, Excel, Power Point and so on. Self-understanding and self-analysis are important when searching for jobs. Another purpose is to deepen self-understanding and a meaning to work through creating a storytelling on the assigned theme.

1.2 Lesson Plans

The teacher recruited three groups for a class some days ago. Then he distributed an experiment description document (A4 paper, 31 pages) to the students and explained the outline of class method. Based on the experiment description document, the teacher explained the experiment's purpose, contents, plan, slide creation method, experimental method, and related details. He distributed one A4 sheet on which six pictures and the narrative stories could be entered. Students were assigned to write a story on the right-hand page and to draw a related picture within the square frame on the left-hand page before the first session. The illustration and character for an animation are drawn outside of the square frame. The students were instructed to paint handwritten picture with colored pencils.

The students were instructed to download the story slide of the "My course in the future" as an example for their assignment. The teacher explained how to animate slides in the first session. They subsequently scanned the images of their sheet with an image scanner. Then they imported the images to Paint software, which were installed as part of Windows XP Microsoft? suite of files. The students cut the images on Paint software and pasted them on slides with PowerPoint. The teacher instructed a student who completes the work and has room at time so as to attach an animation using a function of PowerPoint.

At the beginning of the second session, the teacher explained how to write reports. They were required to enter the final image works into PowerPoint to create their slides. They attached an animation to illustrations and characters for deeper understanding for their stories. Students then recorded their storytelling narrations using a microphone while viewing their slide show after their works were completed. At the end of the second session, the students required to submit the file of the storytelling slide.

Students themselves rated their works on an evaluation sheet after completing the works. At the start of the third session, the teacher printed and distributed a peer evaluation sheet and a handout in which all works by the students were printed. The slide shows of all members in the class were sequentially projected on the screen and viewed. Then they were required to evaluate them for one minute. After the students evaluated each story, the evaluations were entered into the peer evaluation sheet. After viewing all the story slide shows, the students entered the rating value and comment to spreadsheet in Excel, and the files were submitted using the Internet. The teacher gathered and summarized the evaluations in each student's file, then gave each student access to an e-learning portal so that they could download the peer evaluations. The teacher also pointed out the points that should be revised in the printed work and distributed instructor feedback to each student. The students were then required to modify their digital stories and slides by referring to the peer evaluations and the instructor feedback provided. Students performed the second self-assessments after modification in the third session. After the correction, they submitted the file of the story slide. Students themselves rated their works again.

The modified story slides were viewed again in the same manner as during the third session, and then evaluated once again in the fourth session. The second peer evaluation was entered into an assessment spreadsheet, and the files were also submitted.

The teacher gathered and summarized the evaluations in each student's file, then gave each student access to an e-learning portal in the same manner as in the third session. Students then pasted the second peer assessment for themselves on an assessment spreadsheet. By comparing the first and second evaluations, students could learn from and interpret the appropriateness of the corrected elements. Students performed the third self-assessments after the final evaluation in the fourth session.

1.3 Theme of Storytelling

The storytelling of the theme directed was created in this class. Students draw six pieces of pictures matched with the scene of the story of the theme and create a slide of PowerPoint in accordance with it. We require students to consider about oneself, to view a work to another persons, to know the reaction, and to revise a work through creating digital storytelling. A student talks about oneself creating a slide work, after writing its scenario so that a person can understand it.

As expression unlike the sentence, a student was made to consider how to draw and express a picture to supplement the sentence in the story. A student was made to tell one's thought to a person utilizing a PC, using characteristic of a picture and the narration, and being made a story plain. A digital picture book completes when a story was narrated using a microphone so as to promote the understanding of the story contents attaching animation and to reach the feeling of the contents of the story.

We made the first group create a work according to theme of "self-understanding", the second group create a work according to theme of "a children's story", and the third group create a work according to theme of "a future course" as a theme of storytelling. Each group creates it in four weeks respectively.

A story related to autobiographical topics was made to deepen self-understanding as the first theme. “My hobby,” “my memories,” “a childhood dream” and so on are shown as an example. A student was required to create a story which reflects and tells oneself about contents related to oneself.

A student was required to create the story for children so as to make a student understand to act for a person as the second theme. "Fantasyland," "a dog’s adventure," "an insect’s life" and so on are shown as an example. Students were told to create a story that a child would interest in or that is educational and useful for a child.

A student is made to create a story about one's future to make a student have consciousness to work after having done a self-analysis enough as the third theme. “A job that I would like to get,” “workplace where I want to work,” “my dream job” and so on are shown as an example. The students were requested to think about a future course and to create the work with an attitude towards work based upon this theme.

2. Analysis Results and Discussion

Literacy in software functions was investigated before and after the course. Based on the assessed changes in literacy, the degree of achievement of the objectives of this course was estimated. In this study, the term *literacy* represents the skill and ability of use of software functions shown in Table 5. Hereinafter the existence of significant difference is inferred using a criterion of a level of significance of 5%.

2.1 Description of Computer Functions in which Literacy is Acquired

Students were advised to describe in a report what they have understood in the course about the use of a personal computer. They stated that they have learned to use either "narration, PowerPoint, personal computers, Word, Paint, animation, image scanners, or Excel", as presented in Table 2. In all, 102 descriptions and an average of 1.65 per person were obtained. All students described that they gained literacy in at least one function.

2.2 Computer Literacy Investigation Method

Literacy in software functions was investigated as shown in Table 5, to measure computer literacy more quantitatively than the descriptions written in the report and to ascertain the details of its improvement. The survey was administered twice, before (at the beginning of the first lesson) and after (at the closing of the fourth lesson) the course. Students were advised to fill out a questionnaire that surveys literacy level in the functions of three application software: PowerPoint, Word, and Excel. Nakamura et al. proposed measurement of basic knowledge about information technology using questionnaire on computer technical terms and reported its application to evaluation of information education that we could know enough the understanding degree easily and in a short time [12].

The questionnaire presented a total of 60 items, respectively including 25, 17, and 18 items on the functions of PowerPoint, Word, and Excel. Responses of 62 students who answered both before and after the course were used. Literacy levels were classified as "1. do not know, 2. know the name but cannot use, or 3. can use." Students were asked to assign an appropriate number from 1–3 to each item. This rating suggests that the present survey is based on students' personal assessments.

2.3 Comparison of Literacy for Three Software Applications

Table 3 presents the variance analysis results of average literacy in all items for each of three software applications, where *m*, SD, *F*, *p*, Cond, Error, and df respectively signify the average, standard deviation, *F* value, significance probability, between-groups, within-groups, and degrees of freedom. The factor of conditions was significant, as evident in Table 3 (F (5,114) = 34.3, $p < 0.001$).

Multiple comparisons were conducted according to Tukey's method. The result is presented in Table 4, where n. s. represents no significant difference, which indicates that literacy in PowerPoint was significantly lower before the course than that in Word or Excel. However, literacy in PowerPoint was significantly improved after the course, up to a level at which no significant difference was found from either Word or Excel. No literacy difference was observed in Word, although literacy tended to be improved significantly in Excel after the course compared with before. Consequently, results suggest that this practice brought about a good effect on literacy not only in PowerPoint but also in Excel.

Table 2 Personal computer functions for which literacy is acquired

Literacy contents	No. of students
Narration	29
PowerPoint	26
PC	17
Word	8
Paint	8
Animation	6
Image scanner	6
Excel	2
Sum	102

Table 3 Result of variance analysis of literacy in three application software

PowerPoint				Word				Excel			
pre		post		pre		post		pre		post	
m	SD	m	SD	m	SD	m	SD	m	SD	m	SD
2.33	0.67	2.92	0.23	2.82	0.45	2.96	0.15	2.58	0.30	2.93	0.10

Sum of square		df		Mean square		F value	p
Cond	Error	Cond	Error	Cond	Error		
6.37	4.23	5	114	1.27	0.04	34.3	***

*** p<.001

Table 4 Result of multiple comparison of literacy for three software applications.

Time		Pre		Post		
Time	Software	Word	Excel	PowerPoint	Word	Excel
Pre	PowerPoint	***	***	***	***	***
	Word	–	n.s.	n.s.	n.s.	n.s.
	Excel	n.s.	–	**	**	+
Post	PowerPoint	n.s.	**	–	n.s.	n.s.

*** p<.001, ** p<.01, + p<.1

2.4 Applications Comparison of Literacy in Each Software Function

Wilcoxon's signed-rank test was performed to rank values obtained before and after the course for each function of the three software applications. The result is shown in Table 5, where z denotes the Wilcoxon test statistic. A significant difference was observed in 23 of 25 items for PowerPoint. Many students were surprised that a story was presentable in a slide show with voice narration. Presumably, this is the reason why "11 Record narrations" is greatly strengthened. Moreover, literacy in items 15–25 related to animation was remarkably improved compared with other items. Students were instructed to assign a motion to characters, personae, etc. in all six pages. This is regarded as having brought about a good effect.

No significant difference was found in items "1 Launch PowerPoint" and "8 Execute a slide show" for PowerPoint. Their respective rankings had been 2.92 and 2.90 before the course. Improvement was not found after the course because there was little room for improvement.

Significant difference was found in 10 of 17 items for Word for the reasons given below. Items "26 Role of windows in Word" and "28 Page layout" were improved because students were asked to use Word to write a report. Because they were instructed to insert a table in their reports, items "34 Draw a ruled line" and "35 Erase a ruled line" were enhanced. Moreover, because they put the graphs for self-evaluation and peer-evaluation or slides for storytelling in their reports, item "37 Insert an image" was improved. Item "42 Count text characters" cannot be used fully yet, although significant difference was found. Students were advised to include the number of characters and graphs into a table in their reports when they were asked to summarize the story contents. Therefore character counting was explained by necessity. Nevertheless, the story was so short that characters were easily countable without using the function "Count text characters." Therefore, not all the students used the function. Significantly different tendencies were found in each of items 32 and 38.

No significant difference was found in five items for Word: items "30 Change font", "31 Character style", "36 Input characters", "39 Print previews", and "40 Print." They had already been well understood (ranked as much as almost 3 before the course), so there was little room for improvement. No significant difference was recognized.

Significant difference was observed in 14 of 18 items for Excel. The reasons are considered below. Items "48 Arithmetic operation", "49 Copy a formula", and "60 Function" were improved because students were ordered to execute arithmetic operations and to copy the results to some cells when calculating the average of self-evaluation or evaluation of others. Items "45 Make a table", "46 Save a table", and "50 Edit a table" were improved by creating a table with Excel. Item "53 Change the number of digits" was enhanced because students were ordered to arrange the place after the decimal point when writing average ranks in a table in the report. When making a table, students were advised to narrow or extend column width and to center numbers in the case of a numerical column, if needed, for better appearance. These operations strengthened items "51 Width change of rows and columns", "52 Centering", and "54 Draw a ruled line." Students used Excel in various scenarios and tasks, such as placing data into an evaluation sheet, creating a table, and drawing a graph. For this reason, item "44 Role of window in Excel" was enhanced.

Although there were few occasions to print something using Excel before this practice, one required part was exclusively printed with a print range specified in this course. This improved item "55 Designate print range" and "57 Print." When making a report and an evaluation sheet, students were instructed to perform operations so that a table might be well understood by everyone. The items strengthened as above suggest that this instruction enhanced effectiveness. Marginally significant difference was found in two items 47 and 59 for Excel. No significant difference was found in two items for Excel. "43 Launch Excel" and "58 Make a graph" had already been well understood, ranked as much as almost 3 before the course. Therefore, there was little room for improvement. No significant difference was found.

Table 5 Survey results of literacy in software functions

Items of evaluation			Pre		Post		Test	
			m	SD	m	SD	z	p
PowerPoint	1	Launch PowerPoint	2.9	0.4	3.0	0.0	1.6	
	2	Role of windows in PowerPoint	2.5	0.7	2.9	0.3	3.4	***
	3	Layout of slide	2.7	0.6	3.0	0.2	3.5	***
	4	Size of place holder and change of position	2.1	0.8	2.5	0.7	3.1	**
	5	Input of title and change of font size	2.7	0.5	3.0	0.1	3.3	***
	6	Input a text	2.8	0.5	3.0	0.1	3.1	**
	7	Save a slide	2.9	0.4	3.0	0.1	2.0	*
	8	Practice of slide show	2.9	0.4	3.0	0.2	1.2	
	9	Interrupt slide show	2.8	0.5	2.9	0.2	2.1	*
	10	Set background	2.6	0.6	3.0	0.2	4.4	***
	11	Record narrations	1.6	0.7	3.0	0.0	6.7	***
	12	Print distributed document	2.4	0.7	2.9	0.2	4.6	***
	13	Set slide show	2.5	0.6	3.0	0.2	4.9	***
	14	Set effect to change a screen	2.2	0.7	2.9	0.3	5.4	***
	15	Set animation	2.3	0.7	3.0	0.2	5.2	***
	16	Method to play back repeatedly	2.0	0.7	2.8	0.5	5.4	***
	17	Set special effects in a text	2.1	0.8	2.9	0.4	4.9	***
	18	Set special effects to an object	2.0	0.8	2.8	0.5	4.9	***
	19	Change animation	2.2	0.8	3.0	0.1	5.5	***
	20	Practice animation	2.3	0.8	3.0	0.1	5.1	***
	21	Change a kind of animation	2.2	0.8	3.0	0.1	5.5	***
	22	Change the order of animations	2.2	0.8	3.0	0.1	5.4	***
	23	Change to distribute time of animation	2.0	0.8	2.9	0.3	5.4	***
	24	Delete animation	2.3	0.8	3.0	0.0	4.9	***
	25	Set the trace of animation	2.0	0.8	3.0	0.3	5.6	***
Word	26	Role of windows in Word	2.6	0.7	2.9	0.4	3.1	**
	27	Save a document	2.9	0.5	3.0	0.0	2.1	*
	28	Page layout	2.7	0.5	3.0	0.2	3.2	**
	29	Set letter format	2.9	0.4	3.0	0.1	2.4	*
	30	Change font	2.9	0.4	3.0	0.0	1.6	
	31	Character style, a size, underline	2.9	0.3	3.0	0.2	1.2	
	32	Make a table	2.8	0.5	2.9	0.3	1.8	+
	33	Insert a table	2.7	0.5	2.9	0.2	2.9	**
	34	Draw a ruled line	2.8	0.5	3.0	0.2	2.6	*
	35	Erase a ruled line	2.8	0.5	2.9	0.2	2.5	*
	36	Input characters	2.9	0.4	3.0	0.0	1.6	
	37	Insert an image	2.8	0.5	3.0	0.1	2.6	**
	38	Print document	2.9	0.4	3.0	0.0	1.9	+
	39	Print previews	2.9	0.4	3.0	0.0	1.6	
	40	Print	3.0	0.3	3.0	0.0	1.0	
	41	Past up of letter	2.8	0.5	3.0	0.0	2.6	**
	42	Count text characters	2.6	0.6	2.8	0.5	2.6	*
Excel	43	Launch Excel	3.0	0.3	3.0	0.0	1.0	
	44	Role of windows in Excel	2.7	0.6	2.9	0.3	2.8	**
	45	Make a table	2.9	0.4	3.0	0.2	2.0	*
	46	Save a table	2.9	0.4	3.0	0.2	2.2	*
	47	Open a saved table	2.8	0.5	2.9	0.2	1.8	+
	48	Arithmetic operation	2.5	0.6	2.8	0.4	4.2	***
	49	Copy a formula	2.5	0.6	2.9	0.4	4.2	***
	50	Edit a table	2.6	0.6	2.9	0.3	3.7	***
	51	Width change of rows and columns	2.7	0.6	2.9	0.3	2.5	*
	52	Centering	2.8	0.5	3.0	0.0	2.4	*
	53	Change the number of digits	2.5	0.6	2.9	0.4	4.2	***
	54	Draw a ruled line	2.7	0.6	2.9	0.3	3.2	**
	55	Designate print range	2.7	0.6	2.9	0.4	2.7	**
	56	Print preview	2.8	0.5	3.0	0.2	2.2	*
	57	Print	2.9	0.4	3.0	0.0	2.1	*
	58	Make a graph	2.7	0.6	2.8	0.4	1.2	
	59	Editi a graph	2.7	0.6	2.8	0.4	1.8	+
	60	Function	2.4	0.6	2.7	0.5	2.9	**

*** $p<.001$, ** $p<.01$, * $p<.05$, + $p<.1$

3. Conclusion

Students were instructed to create slides about storytelling of contents related to their self-understanding and self-analysis for their career decision, using PowerPoint to strengthen computer literacy. This article described the practice method and measured literacy for computer functions before and after the course. Results were reported herein.

Literacy in functions of PowerPoint was lower than that in either Word or Excel before the course. However, through practice, literacy in functions of PowerPoint was enhanced significantly. Students reported that they had learned to use PowerPoint as they had Word and Excel, which suggests that the functions that are necessary for a presentation were more useful after the practice than before the practice. When making slides, reports, and evaluation sheets during the creation of storytelling, students were instructed to produce a work so that it might be well understood by everyone, which seems to have enhanced effectiveness. Our future plans include analysis of results obtained from an attitude survey to clarify differences in the effects anticipated from three themes in storytelling.

Acknowledgements

The author appreciates the support of the Grant-in-Aid for Scientific Research, foundation study (C22500949) provided by the Ministry of Education, Culture, Sports, Science and Technology, Japan for this research.

References

[1] Susono, H., Simomura, T., Oda, K., Ono, E. (2006). “Production of digital storytelling by K-12 and college students,” *Proceedings of the 22nd Annual Conference of JSET*, pp.911-912.

[2] Sadik, A. (2008). "Digital storytelling: A meaningful technology-integrated approach for engaged student learning." *Educational Technology Research and Development*, Vol.56, No.4, pp.487-506.

[3] Zipes, J. (1995). "*Creative storytelling: building community, changing lives*," New York & London, Routledge.

[4] Tsou, W., Wang, W., Tzeng, Y. (2006). "Applying a multimedia storytelling website in foreign language learning", *Computers & Education*, Vol.47, No.1, pp.17-28.

[5] Kelleher, C., Pausch, R. (2007). "Using storytelling to motivating programming", *Communications of the ACM*, Vol.50, No.7, pp.59-64.

[6] Figg, C., McCartney, R., Gonsoulin, W. (2010). "Impacting academic achievement with student learners teaching digital storytelling to others: The TTTCSE digital video project", *Contemporary Issues in Technology and Teacher Education* (CITE Journal), Vol.10, No.1, pp.38-79.

[7] Ballast, K.H., Stephens, L., Radcliffe, R. (2008). "The effect of digital storytelling on sixth grade students' writing and their attitudes about writing," *Proceedings of Society for Information Technology & Teacher Education International Conference 2008*, pp.875-879.

[8] Ku, Y., Chen, Y., Chang, W. (2012). "Development of early literacy skills: The nature and relationalships between oral and written language comprehension," *Proceedings of the 10th Hawaii International Conference on Education*, pp.1-7.

[9] Hirose, H., Yamamoto, Y. (2004). "Relation between individual characteristics concerning personal computer operation and use frequency of WBT as a supplementary teaching material," *Japan Journal of Educational Technology*, Vol.28, No.2, pp.141-147.

[10] Yamagishi, M., Kimoto, M., Meguro, K., et al. (2009). "Consideration of the application software operation achievement by the change of the PC user's anxiety," *Research Report of JSET Conferences*, JSET09-1, pp.347-352.

[11] Miyaji, I. (2011). "Comparison between Effects in Two Blended Classes Which E-learning is Used inside and outside Classroom," *US-China Education Review*, USA, Vol.8, No.4, pp.468-481.

[12] Nakamura, Y., Kojima, H., Kikugawa, M. (2003).Measurement of basic knowledge about information technology using questionnaire on computer technical terms, and its application to evaluation of information education, *Information Processing Society of Japan Journal*, Vol.45, No.4, pp.1222-1231.

Comparing Enactments of a Collaborative Writing Activity in a Networked Language Learning Classroom

Yun Wen*, Wenli Chen, Chee-Kit Looi
Learning Sciences Lab, National Institute of Education, Singapore
*yun.wen@nie.edu.sg

Abstract: Good learning activity designs do not guarantee effective classroom orchestration by the teacher. Enactments of the same learning activity design may vary greatly among different teachers. This study compares two teachers' enactments of a collaborative learning activity in a L2 writing classroom supported by a networked technology called Group Scribbles (GS). Plausible factors of teacher's moves and actions that impact the different enactments are identified and discussed, including articulating the objective of activity explicitly, providing improvised formative assessment and scaffolding to support students' work on an ongoing basis, and controlling the tempo of the activity and maintaining students' enthusiasm.

Keywords: Enactment; Collaborative writing; Networked language classroom

1. Introduction

In a computer-supported collaborative learning (CSCL) environment, teachers are required to shift their role from being a dominator to being facilitator, guiding and helping students with their learning. Despite the realization that the way teachers enact the instructional practices is essential for the success of collaborative learning, much research focuses on the interactions among students when they are engaged in collaborative learning [2, 9]. Relatively less attention is paid to teacher practices in the networked classrooms where collaborative learning tasks place. A small group of researchers seek to analyze how teachers create opportunities for student interaction through analyzing teacher discourse [7, 8]. These studies investigate how teacher-led discussions have affected classroom interactions and identified successful strategies that teachers used in an inquiry classroom [11]. In addition, findings reveal that enactment styles vary amongst teachers with different beliefs, pedagogy and content knowledge [3]. Puntambekar et al. [6] claim that few studies have investigated "how difference in enactments of collaborative activities might impact students' learning outcomes" (p.82). Enactments may vary greatly amongst different teachers, even though they address the same activity design. Puntambekar et al. [6] compare classroom enactments of an inquiry science curriculum by two teachers and suggest the importance of teachers in helping students make connections between activities such as brainstorming, generating questions, finding and applying information in an inquiry unit. The findings of studies in the context of science and mathematics may not be applicable for language learning. Research on teacher's enactment of pedagogy design in productive collaborative learning for language classrooms is still lacking.

Collaborative activities have been widely used in language learning [1]. A growing number of researchers work on developing technological environments to provide explicit scaffoldings for language learners, as well as visualization of and feedback on group work process. These studies focus on technology design instead of CSCL pedagogical design and enactment in authentic classroom environments. This paper uses a comparative study approach to examining the differences of teacher enactment of the same collaborative writing activity, seeking to explain these differences, and how they might impact language learners' learning outcomes. It is a collaborative second language (L2) writing lesson in a networked classroom supported by a collaborative technology called Group Scribbles (GS). The findings shed light on how to address the teacher's role and challenges in enacting well-designed CSCL activities successfully in real classroom settings.

2. Context of Study

The study described here is part of a 3-year project introduce Rapid Collaborative Knowledge Improvement (RCKI) to language learning classrooms in a secondary school of Singapore [5]. The school provides a technology-rich environment for students. Each student is equipped with a laptop. In Singapore schools, English as the first language is the main teaching language in schools, whereas Chinese is taught as a second language (L2) for the Chinese ethnic students. This paper focuses on the enactment of collaborative L2 argumentative writing lessons in secondary grade 2 (14-16 year old) higher Chinese classes.

The two experimental classes are selected because the students' writing ability of the *two classes is at a similar level* in term of the scores of writing in school's examination on Chinese subjects (t=-0.265, p>0.05). One of them, Class E1 (N=20) is taught by *Chin*, who is a female teacher with approximately 10 years of teaching experiences. The other one is Class E2 (N=16) taught by *Judy*, who has about 5 years of teaching experiences. Both Chin and Judy have had experiences of studying Chinese language abroad (Chin in Taiwan for 4 years and Judy in mainland China for 3 years). Compared with other local Chinese language teachers without overseas educational backgrounds, these two teachers are fully aware of the necessity and importance of Chinese ethnic students in mastering the Chinese language well. They are willing to try new teaching approaches to arouse students' interests in Chinese language learning. Both of them fully believe that every student has potential, and what teachers need to do is to assist students to reach their potential.

3. Intervention

GS is a software platform designed for supporting students to create lightweight multimodal representations for mediating collaborative activities. Its workspace is divided into private and public spaces presented in a two-paned window (Figure 1). The lower pane of the GS is the user's personal workspace or private board whereas the upper pane is the public board or public board. The private workspace was provided with a virtual pad of fresh scribble sheets on which the user could draw or type. The students can share the scribbles sheets by dragging them from private space to public space. A student can select any group board by clicking the board number on the right-top, and browse all other groups' postings posted on the public board. GS hence promotes and facilitates intra- and inter-group interactions.

When exploring the affordances of GS for Chinese language learning, we have proposed the rapid collaborative knowledge improvement (RCKI) concept and related 9 principles. The concept of RCKI refers to the notion of democratizing participation and idea refinement in the context of live dynamic classroom settings, that is, face-to-face (FTF) collaborative knowledge construction and improvement over the duration of a class session, and supported by certain technologies for lightweight instant interaction (see [9, 12]).

Before implementing the GS-based collaborative writing classes, both teachers and students had been familiar with its function. A series of professional development sessions (1 hour per week, 5 weeks) were held by two researchers to ensure the teachers' belief and understanding about GS-based language learning and RCKI principles. After that, a GS-based Chinese writing lesson (60 minutes) was co-designed by the teachers and researchers, guided by RCKI principles and argumentative writing strategies.

The main learning objective of the lesson was to help students understand that an argumentative essay can be written from discussing a phenomenon followed by finding effects and providing solutions. The topic for the writing was "*Guilty? Plastic survey*". A template (Figure 2) was uploaded as the background of each GS group board for assisting learners to perform tasks. Chin and Judy enacted the same lesson plan for their classes. Table 1 presents the main phases of the designed collaborative writing activity. Both classes of students were heterogeneously organized into groups of 4 members each.

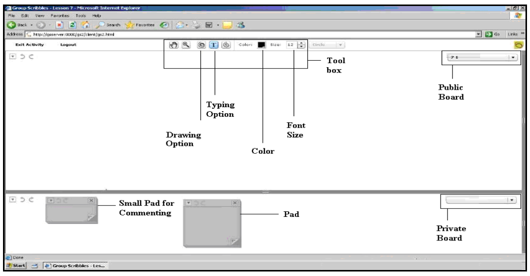


Figure1. The user interface of GS

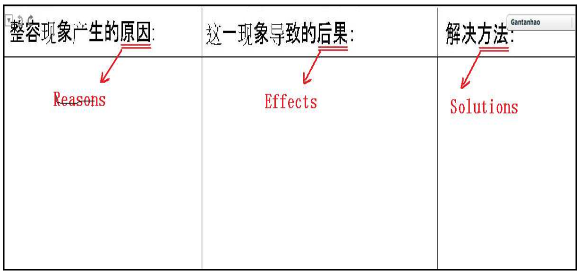


Figure2. A graphic organizer for the activity

Table 1: Overview and main phases of the activity

Phases	Teacher-directed moves
P1	Introduce to students the main purpose of the GS activity; help them recap strategies for argumentative writing; encourage them to think of phenomenon around the topic and brainstorm reasons for the phenomenon.
P2	Facilitate students to perform the task (providing "results" based on given "reasons"); ask them to explain the results and give comments for group artifacts.
P3	Facilitate students to perform the task (thinking about "solutions" based on existing "reasons" and "results"); ask them to explain the results and give comments for group artifacts.
P4	Encourage students to do FTF discussion to improve their group artifacts, synthesize and extract big ideas about their group writing.
P5	Facilitate and ask students to present group final artifacts, provide comments and summarize the whole lesson.

4. Method

This study compares two teachers' enactment of the lesson plan and explores its relation with students' learning outcomes. There are three main sources of data: 1) results of students' subsequent individual writing; 2) individual student feedback; 3) teachers' instructional discourse in the classroom. The quantitative results of subsequent individual writing and students' qualitative feedback together help ascertain any difference in learning outcomes between Chin and Judy's class. The data regarding the two teachers' instructional discourses are analyzed to find out if and how the two teachers' enactments differ, in an effort to understand what might have affected the effects of collaborative writing enactment.

When collecting data on teachers' momentary instructional practices, two researchers observed Chin and Judy's GS lessons, took notes and captured the whole class process by video cameras. The researchers used a chronological representation tool called *Studio Code* to provide visual patterns of their instructional discourse. First, for the sake of consistency, two researchers watched and transcribed all the video data about teacher discourse in two classes. Next, the transcribed data was segmented into units of "theme" by using semantic features such as ideas, discussion topic, or by regulative actions such as asking for an explanation or explaining on specific point. Finally, the teacher instructional discourse was coded again on Studio Code to present the patterns of the teacher practices visually.

The coding scheme was designed to capture five aspects of enactment. It was developed through an iterative process of creating codes, coding, modifying and refining codes, and recording consistent with Miles and Huberman's [4] recommendations for rigorous and meaningful qualitative data analysis. We did open coding as many existing coding schemes are mainly about inquiry-based learning and thus they are not applicable for language learning. For instance, we added the item "Relating to language" borrowed from Swain and Lapkin [10] who coined language-related episodes when they studied L2 learners' language use in collaborative dialogues. Parts of the coding categories of the study came from research on enactments of inquiry lessons [6]. For example, "*Relating to topic*" referred to the way in which students were encouraged to think back about the topics they had already learned, and to connect that learning to the new topic that was being introduced. In this study, it referred to relating knowledge and strategies of argumentative writing to the current topic. "Focus on goals" referred to the speech that teacher expressed the aim of lesson or the activity design explicitly to students. The details are listed in Table 2.

Table 2: Coding schemes with examples

Categories	Items	Interpretations	Examples
1: Ensure the activities can be completed as designed	Task introduction	Introducing how to complete the tasks	"In the task 2, your group should give 'effects' based on 'reasons' that have been posted."
	Time management	Reminding students to pay attention to time	"Last 2 minutes. Seize the time to complete your group's task."
	Providing encouragement	Encouraging or giving praise to students' performance	"Group 2 did quite well."
2: Ensure the activities can be completed with a high quality	Protocol	Rules for group task complement in GS-based learning environment	"Each group leader help teacher to monitor the procedure of your group work."
	FTF discussion	Encouraging students to do FTF discussion	"Communication. Let me hear your voice of the group verbal discussion."
	Ideas	Encourage students to share ideas and improve them consistently	"After visiting other groups' board, you need to improve own existing postings"
	HOT	Encouraging skills like analysis, synthesis, categorizing, evaluation	"If you disagree with the comments from your peer group, please give your reasons."
3: Ensure students can understand the significance of the activity design	Relating to topic	Help students think back about the topics that they had learned, and connect that learning to the new topic that was being introduced.	"We have learnt about the 'Five Fingers' which taught us that an argumentative essay could be written from: individual, family, friend, county and society. You can give reasons from these 5 perspectives."
	Focus on goals	Helping student keep the overall goal of the challenge in mind while being engaged in activities.	"Today, the ultimate goal of our study is to mater the RES model for your argumentative essay writing."
4:Help students master language knowledge	Relating to language	Episodes in which the teacher deliberate over lexical or grammatical choices	"The idiom should be 'Ru Huo Ru Tu'. The last character should be pronounced as 'Tu'. "
5: Help students improve cognitive skills	Asking for further think	Opportunities for students to explain their thinking	"Here, what does 'others' refer to?"
	Explaining specific point	Comment and elaborate on student ideas	"Social influence here might be understood as plastic surgery rampant in the community"
	Assessing or summarizing students' work	Providing evaluative response to students' group artifacts	"Group 5 provides reasons for the phenomenon of plastic surgery mainly from individual and society these two aspects"

5. Results and Discussion

5.1 Students' Learning

We started our analysis of student learning by examining whether the two classes performed differently in the subsequent individual writing. Each student took 50 minutes to complete the writing with the same topic in the following Chinese language class. Their compositions were marked according to writing rubric from Ministry of Education of Singapore, which has 5 parameters: 1) Solid writing material; 2) In line with topic and the goals; 3) Consistence of reasoning and focus 4) Diversity and innovation of ideas 5) Creative imagination. The maximum score for a composition is 70. All compositions from both classes were marked by two teachers. The strength of association between scores marked by them was high (r=0.727). The result of t-test shows a significant difference of students' subsequent writing scores between two classes (t=3.153, p<0.01). The mean score of Chin's class is 47.53, higher than 41.88 of Judy's class. It indicates that the writing score of Chin's class after GS-based collaborative activity is significantly higher than that of Judy's class.

In addition to academic performance, students were required to reflect on their learning experiences which help researchers to explain why the results of sequential writing differ between Chin and Judy's classes. Two semi-constructed questions were provided: Q1) Collaborative writing and individual writing, which one do you prefer? Why? Q2) What did you learn from the last GS lesson? Towards Q1 Almost all of them were positive about collaborative writing. Two students from Chin's class and three students from Judy's class stated that they prefer individual writing to collaborative writing. Students from different classes provided different reasons. Students from Chin's class emphasized that they enjoyed the process of creating alone more, compared with completing a collaborative writing task. Students from Judy's class merely emphasized that they felt it was easier to do individual writing. Of those who preferred to collaborative writing, the predominant reason given (by 12 students from Chin's class and 9 students from Judy's class) was that it provided them with an opportunity to compare and exchange ideas with each other. For example, students from Judy's class pointed out that:

> "I prefer to writing together with my group members, as we can share our ideas, help each other to finish our writing. " "I like group writing as we come to know others' views towards the topic, and we can choose the most suitable opinions through discussions."

Chin's students however, provided wider and deeper reasons to explain why they like collaborative writing. In addition to sharing ideas, they noted that:

> "I like to write together in a group, because we can discuss with each other and provide ideas and suggestions for each other. My own idea might not be the best, but we can keep improving these ideas in group work." "…students own different ideas toward the same topic, thus everyone will be involved in intensive discussion on how to write an article. During this process, we can learn from each other, and in this way, our abilities of critical thinking and collaborative learning get opportunities to be improved."

The most obvious difference is from two class students' answers for the second questions. All 20 students from Chin's class noted that they learned that "reasons, effects and consequences (RES) are three essential parts of an argumentative essay." Just like one student expressed that:

> "…RES indeed help us better understand the process and the theme of essay writing. It portrays the whole thinking process and the way to find out the answers. Making use of this model, students can get the right way of writing with guidance."

However, no student in Judy's class mentioned RES. Their feedback focused on "how to write together" or "how to complete a task with group members within a short period of time". It seems that students in Chin's class could better understand the intent of teacher's activity design. This may explain the differences on students' academic performance between two classes, although in both classes, students hold a positive opinion of their group work. It also seems that the students in Judy's class were highly engaged but did not learn as much as students in Chin's class. What might account for this difference? The researchers try to further find out if there is a difference between teachers' enactment that may cause the differences in students learning

5.2 Comparing Teacher Discourse through Chronological Representation

We used Studio Code to represent teacher discourse chronologically to understand how they facilitated GS activities. This method enables a graphic representation of the chronology of discourse, allowing an understanding of how it changed over time [7]. Figure 3 and Figure 4 describes Chin and Judy's discourse as they occurred respectively. The top line of each figure indicates the time period of every teacher-directed phase (Table 1). The line of the figure depicts a single category with the incidence of teacher discourse in that category represented along the horizontal line. Each action is represented using a bar code.

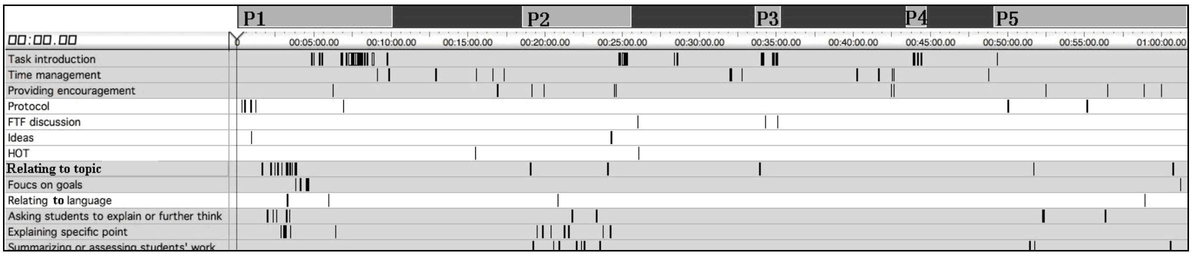


Figure 3. Chronological representation of Chin's discourse

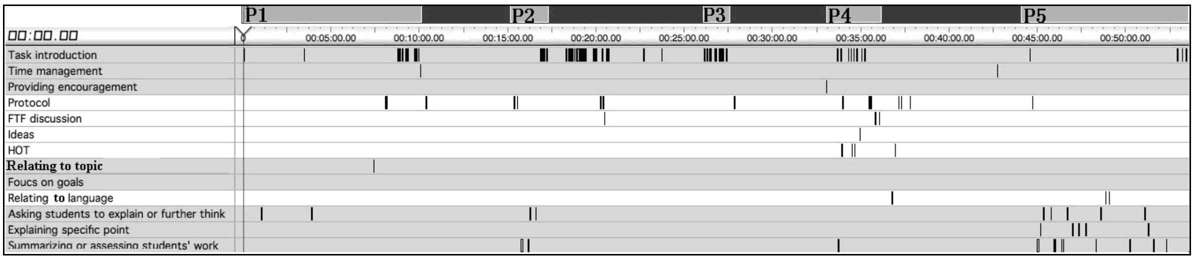


Figure 4. Chronological representation of Judy's discourse

Generally speaking, both Chin and Judy did enact the lesson plan with some fidelity. They spent almost the same time at the beginning phase of orientation and introduction of the activity (P1), and the last phase of evaluation and students presentation (P5). But the patterns of two teachers orchestrating the activity are different. At the orientation phase (P1), Chin made the goals of the GS-based activity explicit. Rather than assigning the task directly, she spent 5 minutes in introducing the RES and helping students to recap other related skills for argumentative essay writing. Judy, however, spent about 7 minutes to talk about the existing phenomenon of plastic surgery and to give examples in daily life. Judy spent a lot of time seeking to arouse the students' interests of the topic discussion, but without making the goals of the activity clear to students.

Figures 3 and 4 indicate that Chin spent more time than Judy at phase 2 of activity (P2). When students were brainstorming reasons for the phenomenon of the popularity of plastic surgery, Chin said "If your group has posted sufficient 'reasons', now you need to group these given reasons. Do remember to think of the five perspectives mentioned in 'Five Fingers' that we have learnt". After the students had completed the first phase of the activity, Chin selected some of the group works to provide comments and further explanations (see Figure 3). However, Judy did not asked her students to categorize their brainstormed ideas and she seldom provided improvised feedback at the class level, though, like Chin, she kept passing between groups and monitoring the state of group work as well.

Below figure 5 visualizes the differences between the number of instances of Chin and Judy's discourse occurring in each category. It is observed that Judy spent more time on task introduction than Chin. In other words, both teachers gave their task instruction clearly to ensure students follow the designed procedures, but Chin's introduction was more concise. Compared with Judy, Chin managed class time more strictly, and she praised and encouraged her students more frequently. The similarity is that both encouraged students to do FTF discussion to improve their ideas consistently, and to give comments/suggestions for others at the group or class level. Except for the number of providing protocol, there is no obvious difference in category 2. Judy liked to provide more specific instructions regarding to the GS group work protocol. For example, she required each group to pick a different color to represent their group. She thought this might help proceeding round robin smoothly. This could potentially cause students losing their individual identity. In contract, Chin preferred that group students could generate their own group protocol.

The most significant difference between Chin and Judy's enactment discourse can be found in category 3. Chin helped students make clear the learning objective and make explicit the connection between the meaning of collaborative writing activity design and learning content, rather than merely gave commands for students to follow procedures automatically. This sort of discourse often happened at the beginning of the activity in Chin's class (figure 3). She used this approach to help students to complete subsequent phases of the activity effectively. This result is consistent with findings drawn from students' feedback, that students from Chin's class had a better understanding on the objective of the designed collaborative writing activity. The difference between two teachers' enactments in this category is probably one of the main reasons leading to the difference of students' performance in the subsequent writing between two classes.

As for category 4 & 5: Help students master language knowledge & improve cognitive skills, there is not much difference on the discourse frequency between two teachers. Both teachers' instructional discourse did not focus on correcting grammatical or syntactical errors. They sought to scaffold students' cognitive development and problem-solving approach. The difference exists in the timing of providing scaffoldings. Chin monitored the progress and quality of group work and reacted immediately, whereas Judy concentrated her explanation and assessment when students had completed the activity thoroughly.

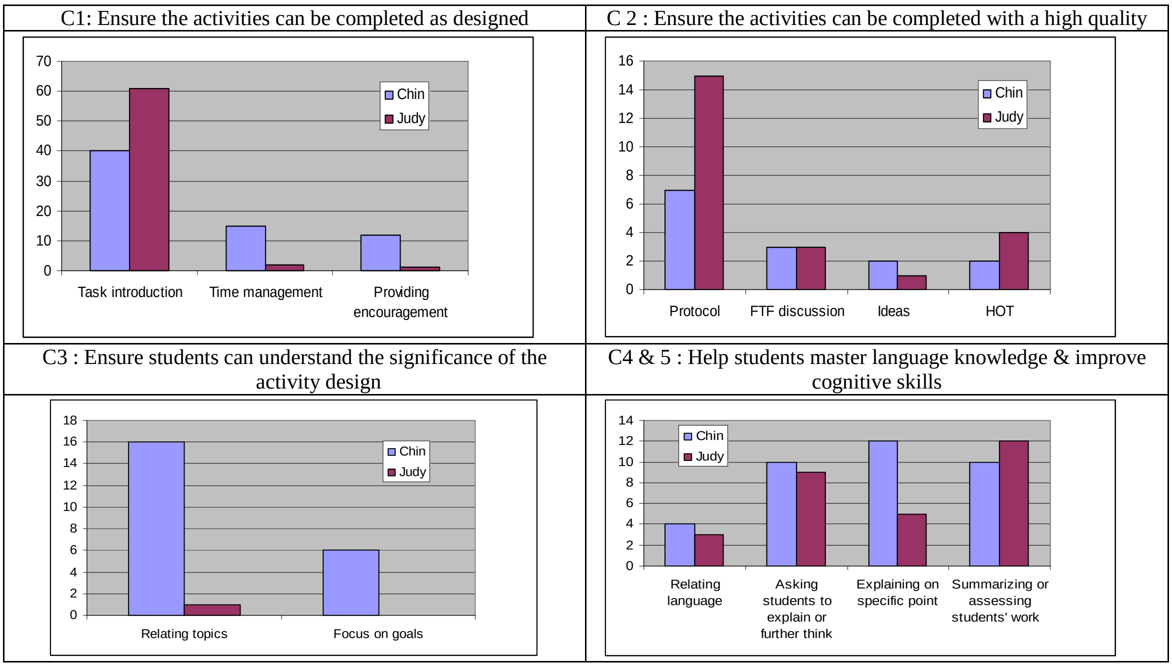


Figure 5. Frequency chart comparing Chin and Judy's discourses

6. Conclusion and Limitations of This Study

This study pointed out three major differences in the enactments of lessons of two language teachers. Firstly, the learning objectives and the connection between the purpose of writing activity design and learning content, were made explicit in Chin's class, whereas it was not evident in Judy's class. Secondly, Chin was able to improvise teaching in the light of dynamic formative feedback from students' group works at each phase of the activity. Judy, however, commented students' group work only at the end of the activity. Thirdly, Chin strictly controlled the time of each segmented activity and consistently encouraged and praised students. Although Judy also did well at the beginning trying to arouse students' interests in participating in the activity, she neglected to maintain this enthusiasm. The results suggest that language teachers should 1) articulate the objective of the collaborative activity on language learning explicitly; 2) provide improvised formative assessment and scaffolding to support students' cognitive development and problem-solving in student working process, rather than being entangled on correcting grammatical or syntactical errors; 3) control the tempo of the activity and maintain students' enthusiasm.

There are limitations in this study that need to be addressed in further research. For example, this study focused on analyzing the differences of enactments between two teachers. The commonalities between them are not been taken into consideration, though they may influence the effects of the enactment. The present comparative study is conducted under the assumption that two teachers have the similar beliefs about student language learning after a series of professional development sessions. We acknowledge that the teachers' perceptions about the students as well as their knowledge and beliefs about learning influence their enactments. Thirdly, due to the page limit, we are not able to present the details about the process of students' collaborative learning in relation to the teacher's instructional discourse.

References

[1] Blake, R., Wilson, N. L., Cetto, M., & Ballester, C. P. (2008). Measuring oral proficiency in distance, face-to-face, and blended classrooms. *Language Learning & Technology*, *12*(3), 114-127.

[2] Cakir, M. P., Zemel, A., & Stahl, G. (2009). The joint organization of interaction within a multimodal CSCL medium. *Computer-Supported Collaborative Learning, 4*, 115-149.

[3] Dimitriadis, Y. A., (2010). Supporting teachers in orchestrating CSCL classrooms. In A. Jimoyiannis (Ed.), *7th Pan-Hellenic Conference with International Participation ICT in Education* (vol.I, pp.33-40). Korinthos, Greece.

[4] Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.) Thousand Oaks, CA: Sage Publications.

[5] Looi, C. K., Chen, W., & Patton, C. (2010). Principles and enactment of rapid collaborative knowledge building in classrooms. *Educational Technology Magazine*, September.

[6] Puntambekar, S., Stylianou, A., & Goldstein, J. (2007). Comparing classroom enactments of an inquiry curriculum: Lessons learned from two teachers. *The Journal of the Learning Sciences*, *16*(1), 81-130.

[7] Sandoval, W. A., & Daniszewski, K. (2004). Mapping trade-offs in teachers' integration of technology-supported inquiry in high school science classes. *Journal of Science Education and Technology, 13*(2), 161-178.

[8] Schneider, R. M., Krajcik, J., & Blumenfeld, P. (2005). Enacting reform-based science materials: The range of teacher enactments in reform classrooms. *Journal of Research in Science Teaching, 42*, 283–312.

[9] Suthers, D. D. (2006). A Qualitative analysis of collaborative knowledge construction through shared representations. *Research and Practice in Technology Enhanced Learning, 1*(2), 115-142.

[10] Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*(iii), 320-338.

[11] Tabak, I., & Reiser, B. (1997). Complementary roles of software-based scaffolding and teacher-student interactions in inquiry learning. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Proceedings of the 2nd International Conference on Computer Support for Collaborative Learning* (pp.289–298). Toronto, Ontario, Canada: AACE.

[12] Wen, Y., Looi, C.K., & Chen, W. (2011). Towards a model for rapid collaborative knowledge improvement in classroom language learning. In Hans, S., Gerry, S., Naomi, M., Nancy, L. (Eds.) *Connecting Research to Policy and Practice: Proceedings of CSCL 2011* (pp. 836-851). Hong Kong.

Do students and lecturers actively use collaboration tools in learning management systems?

Nastaran ZANJANI, Shaun NYKVIST[*] & Shlomo GEVA
Queensland University of Technology, Australia
*s.nykvist@qut.edu.au

Abstract: In recent years there has been a large emphasis placed on the need to use Learning Management Systems (LMS) in the field of higher education, with many universities mandating their use. An important aspect of these systems is their ability to offer collaboration tools to build a community of learners. This paper reports on a study of the effectiveness of an LMS (Blackboard©) in a higher education setting and whether both lecturers and students voluntarily use collaborative tools for teaching and learning. Interviews were conducted with participants (N=67) from the faculties of Science and Technology, Business, Health and Law. Results from this study indicated that participants often use Blackboard© as an online repository of learning materials and that the collaboration tools of Blackboard© are often not utilised. The study also found that several factors have inhibited the use and uptake of the collaboration tools within Blackboard©. These have included structure and user experience, pedagogical practice, response time and a preference for other tools.

Keywords: Learning management systems, Communities of Practice, e-learning

Introduction

In recent years, the necessity for Higher Education Institutions (HEI) to invest in learning management systems (LMS) that provide a platform for e-learning (electronic learning) has increased. This has often been seen as an attempt for these institutions to be more competitive and to capture a larger market share of students [33]. Initially, the idea of using e-learning systems was focused around the ability to connect with external and distance education students and provide greater access and flexibility to these students [2; 24]. However, e-learning has now become a core component of the education experience for many students in higher education and an ever-increasing combination of face-to-face (F2F) learning and e-learning is now occurring [5; 20]. This learning, referred to as blended learning, uses technology to expand the physical boundaries of the classroom, providing access to learning content and resources and enhancing the instructor's ability to receive feedback on learners' progress [21].

In creating this blended learning environment in higher education, an LMS (such as Blackboard© or Moodle) is often used to access inbuilt collaboration tools such as blogs, wiki's and discussion forums. These tools, often referred to as web 2.0 or e-learning 2.0 tools, are most common to these environments and touted as having the ability to empower educators to facilitate a sense of community through the possible interactions that could occur in these environments. Consequently, it is this buoyant relationship between the use of Internet collaboration tools and people that has the potential to create powerful online learning communities [8; 16].

In an attempt to further explore the relationship between current research into the benefits of using online collaborative tools to create a sense of community, this paper will describe and report on a small scale-study (N=67) into the use of collaborative tools within higher education. This study is specifically based on the collaborative tools available as default within the LMS, Blackboard©.

1. Blended Learning and Collaborative Tools

Blended learning (or hybrid learning) combines e-learning with other, usually more traditional forms of teaching and learning [21]. Bielawski and Metcalf [3] described it as "blending classroom, asynchronous and synchronous e-learning, and on-the-job training" (p. 71). It is generally held that blended learning "combines the advantages of two learning modalities [34], p.157 with Bowles [6] suggesting that "when classroom instruction is combined with self-paced instruction via the Internet, for example, the face-to-face contact makes for easy social interaction and allows for instant feedback" (p. 47). The advantages of blended learning may be summarised as follows:

> *The blended learning approach helps to create a shared understanding of concepts important to the learning culture and provide opportunities to reinforce them in a live classroom setting. Leveraging the convenience and accessibility of online components with traditional classroom instruction also expands the curriculum without increasing programme completion time.*

It is this strong relationship between the F2F interactions and online collaborative tools in a blended learning environment that has the potential to move educators from a didactic approach of teaching and learning to an approach that is based on building a sense of community through computer mediated communications (CMC). CMC is a term referring to the interpersonal discourse between users with computer-based media. CMC extends from discussion boards/forums through to contemporary Web 2.0 applications [36] and is said to enable collaborative reflection, which, in turn, prompts the conceptualisation and re-conceptualisation of ideas [10; 25]. It is these conversations and interactions between students that strengthen their deeper understanding of the topic [29].

Learning management systems within the higher education sector provide educators with an environment containing inbuilt collaborative tools (e.g. discussion forums, blogs and wiki's) to use for their teaching purposes. These collaborative tools can be used for computer-mediated communication where communities of practice can be supported and envisaged. When these tools are coupled together with F2F teaching the notion of blended learning can be realized. In realising this notion of blended learning, a widely used LMS such as Blackboard©, is often used in the higher education sector [27].

The Blackboard© website publishes a number of case studies which highlights the possible advantages of using the Blackboard© LMS within higher education. According to Blackboard© [4], the University of Cincinnati has been a Blackboard© customer since 1999 and uses their LMS for distributing learning resources, podcasts of recorded lectures and announcements while the University of North Carolina provides customized library content for students using their LMS [27]. However, despite the large number of clients using Blackboard©, Heaton-Shrestha [17] found that learning resources and announcements are the most valued tools of Blackboard© used by students and lecturers of the Kingston University, not collaboration tools. Current literature highlights the importance of these tools, however, further research [12; 22] identifies a lack of active participation by students and teaching staff with these tools in the Blackboard© learning environment.

While there is little empirical research surrounding the lack of use of these tools in this environment, Alexander and Boud [1] claim that the potential for online learning is not being realised due to traditional didactic approaches being transferred to the online environment. This approach merely mimics the traditional classroom with lecture notes and resources being placed online and the LMS is seen as a web-based delivery of course resources or as a communication tool. The need to understand the issues surrounding the limited use of these collaborative tools within an LMS such as Blackboard© is essential for a blended learning environment to exist. One of the major issues facing researchers is the rapid advancement of technology used within these environments and the ability for research to keep abreast of it [15; 28].

According to Greenagel [13] the development of collaborative learning systems that ignore users learning styles could be one contributing factor to their failure to engage students and staff in their use, while Everson [11] and Wallace [35] argue that the user friendliness and interface design need to be considered. Everson [11] advises not to "waste valuable time preparing tools that will only frustrate and disenchant your students". Romiszowski [29] further claims that these systems should focus attention on efficient learning materials and not just deal with indexing, coding and tagging teaching objects to facilitate using digitized learning materials.

These factors of design and usability may contribute to some reluctance by students and teachers to use the systems, however other factors such as increased workload may also contribute to the lack of their use [26; 30; 32]. The administration of students and the monitoring of their interactions can contribute to an increased workload for an educator. The issue of workload was not only evidenced by educators, but also students, who often complained when asked to use the collaborative tools as part of their learning experiences [30]. In contrast to this, Jones, Blackey, Fitzgibbon and Chew [19] claim that students with individual interests attempted to use the available collaborative tools when afforded to them.

Given that collaborative tools within an LMS such as Blackboard© offers a means by which blended learning can occur, current research highlights the challenges that educators within higher education institutions face in actively using these collaborative tools effectively. Consequently, the following study reported on here investigates the use of collaborative tools within Blackboard© at a university that encourages blended learning across all disciplines.

2. Research Design

2.1 Research Objectives

The main purpose of this small-scale study was to investigate how collaborative tools are being used within Blackboard© for teaching and learning at a major Australian University, and to explore the factors that influence their usage. Blackboard© is a conventional learning management system used for a period of five years at the university where the study was conducted.

2.2 Research Methodology

The study employed a mixed methods approach that combined both qualitative and quantitative strategies. Simultaneously using both approaches allows for a holistic view of the problem to be generated and provides a 'comprehensive analysis of the research problem' [9]. Open-ended interviews were conducted with participants and usage data pertaining to the use of collaborative tools in Blackboard© were collected and analysed. The open-ended interviews cater to more substantial information being generated by allowing respondents to state their own perceptions with their own expressions [31], while the combined usage statistics build a more holistic view of the study.

2.3 Participants

The participants (N=67) of the study consisted of both teaching staff (n=9) and students (n=58) from the faculties of Science and Technology, Law, Business and Health, all studying at the university where the study was conducted. The participants were informed of the study through emails sent to each of the faculties and they individually volunteered and gave full consent to participate in the study. Participants who volunteered to be a part of the study were of varying ages and of mixed sex.

2.4 Interviews

Interviews were focused on the way students and lecturers use the collaboration tools of Blackboard© and were open ended in nature. The number of participants interviewed was brought to an end once a saturation point had been reached where no new data was collected from participants. Guest [14] demonstrated that saturation often occurs within the first twelve interviews and that this is sufficient to obtain a reliable conclusion. The participants interviewed in this study were 67.

3. Research Findings and Discussion

The two main sources of data were the responses to the interviews and the Blackboard© usage statistics of the collaborative tools. This section will present the findings in each of these areas.

3.1 Interviews on the use of collaborative tools

The participant interviews (N=67) revealed that 33% staff interviewed used collaborative tools within Blackboard© while 51% students interviewed indicated that they had used collaborative tools as part of their learning experience. All the students that indicated that they had used some of the tools, also stated that they only briefly used them as an add-on to their existing learning experiences.

An analysis of student (n=58) and staff (n=9) responses from the interview on factors pertaining to why they did not use collaborative tools in Blackboard© could be grouped under six main categories; structure and user experience, availability of time, preference for other tools, lack of knowledge about tools, pedagogical practice and response time.

3.1.1 Structure and User Experience

The structure and user experience of collaboration tools within Blackboard© accounted for one of the major reasons why collaborative tools were not used. Over two thirds of student participants (67.35%) indicated that this was an issue while 67% of lecturers also found structure and user experience to be a factor. Structure and user experience related to the ease at which participants could easily navigate and find functions and use the collaborative tools. General consensus from student participants is that Blackboard© is hard to navigate and it is not user friendly while staff participants also found complicated procedures associated with using the tools.

3.1.2 Availability of Time

The availability of time was highlighted as a contributing factor by 13% of the student participants and 55% of the staff participants. Student participants indicated that they struggled to find time to keep up with the other requirements of the unit and learning how to use the tools or to participate online was seen as another burden on time. Supporting this finding one of the students stated that *"I have no time to do this, because you have to spend a lot of time to understand how to set up them [the forums]" (S-6).* Lecturers were also concerned about the time needed to structure the use of the tools and be actively involved with either synchronous or asynchronous discussion.

3.1.3 Preference for other tools

A preference for other tools that students (39.5%) were already accustomed to was indicated as another factor in the student use of collaborative tools in Blackboard©. While students indicated a preference for other tools such as Skype or MSN Messenger to discuss issues or topics pertaining to their study or a particular unit, no lecturers indicated a preference for other collaborative tools. One of the students questioned *"why you would learn something new when there is already something else available just as good … is wasting time" (S-5).*

3.1.4 Lack of Knowledge about Tools

Lack of knowledge about the functionalities of the various collaborative tools or their existence within Blackboard© was identified as a factor affecting their use. This was identified as a factor by 48.2% of students and 75% of lecturers. A number of students referred to Blackboard© as merely a platform to access learning materials and receive announcements. It is within this context that research conducted by Bradford et al. [7] supports these findings in associating complexity and knowledge of LMS tools as a limitation of these environments.

3.1.5 Pedagogical Practice

The pedagogical practices of 50% of lecturers were seen as a factor impinging upon the use of collaborative tools in their teaching. These lecturers indicated that they were comfortable with their traditional approaches and that shifting to new practices was difficult and time consuming. This approach is reiterated by Alexander and Boud [1] who claim that these environments are not being used to their full potential and that didactic teaching practices have become a part of these online environments.

3.1.6 Response Time

Response time refers to the length of time students had to wait to receive a response using asynchronous collaboration tools within Blackboard©. This was indicated as a contributing factor to why 48% of students did not use or continue to these tools. This confirms findings of a number of studies [23] where a lengthy response or no response discourages the student to use the collaboration tools in Blackboard©.

3.2 Statistical Usage Data

The interviews conducted in this study provided qualitative data on factors pertaining to why participants did not use the collaborative tools in the Blackboard© environment. The statistics presented in this section demonstrate the average time that students in the entire university spent engaged with Blackboard© (Figure1) and further confirmed that the Blackboard© LMS is mostly used as an online repository for teaching resources (Figure 2).

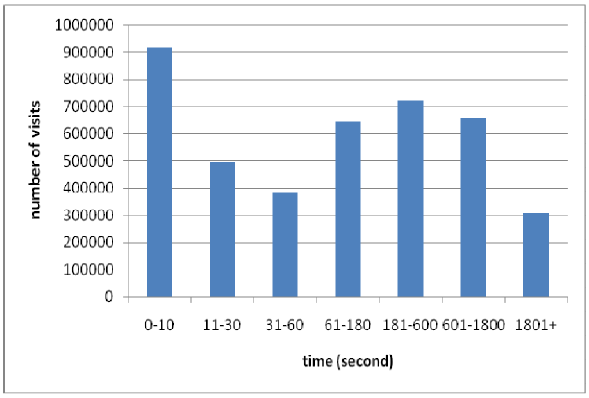


Figure 1: User engagement

Figure 1 shows the rate of university-wide user engagement in Blackboard© for one semester. This user engagement does not distinguish between a user reading and viewing learning resources or creating resources and the use of collaborative tools. We can ignore visits shorter than 30 seconds as this is too short a time to indicate a significant collaboration attempt if at all. A normal distribution of visit durations is observed.

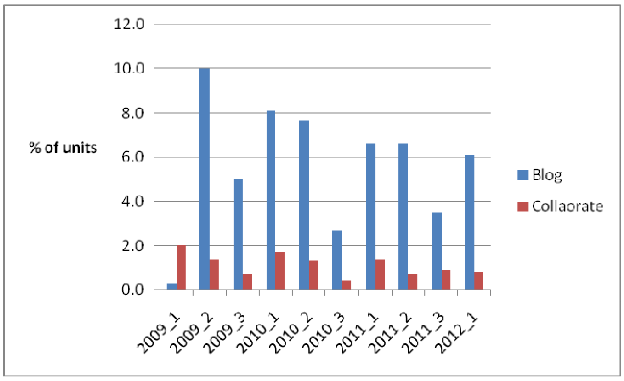


Figure 2: Percentage of courses that use at least one collaborative tool

Figure 2 displays the percentage of units that used Blackboard© collaboration tools in the university. These statistics are captured over a three year period for each teaching semester and indicate that 10% or less of the units offered at the university use some form of collaborative tools.

The data presented in Figure 2 not only demonstrates a low percentage of units using Blackboard© collaboration tools but also indicates a decline in the use of these collaboration tools. These findings further support and confirm the results of previous studies [12; 17; 18] and beliefs that learning management systems such as Blackboard© are used mainly as a content delivery mechanism and not used to their full potential.

The most prominent factor contributing to the lack of use of the collaborative tools in Blackboard© resulted from a negative user experience with the tools with over two thirds of both students and lecturers indicating that this was an issue. The next most important factor for all participants was related to an understanding of the available collaboration tools, followed by availability of time for lecturers and response time for students. While literature [8; 16] touts the importance of using these tools for building communities of practice, it is evidenced in these findings that there would be 10% or less units in the university with the ability to build these powerful learning communities within Blackboard©, due to the absence of students and teaching staff actively using collaborative tools.

4. Conclusion

There is no doubt that e-learning is a significant part of higher education teaching and learning, however it is vitally important that it is used in ways that promote and encourage positive learning experiences for all and build communities of learners. The mere existence of collaborative tools in an LMS such as Blackboard© does not automatically equate to them being used for successful teaching and learning purposes. While the study presented here is a small-scale study of one typical large university and one LMS, it demonstrates the need to address the key factors that act as barriers to the use of collaboration tools in higher education. The most significant factor in the study was that of structure and user experience. It highlights the need to design computer supported collaboration tools that encourage student interaction to produce collaborative knowledge building through communities of practice.

References

[1] Alexander, S. A. & Boud, D. J. (2001). Learners still learn from experience when online. *Teaching & Learning Online - Pedagogies for New Technologies*, 3, 15.

[2] Allen, I., & Seaman, J. (2007). *Online Nation: Five years of growth in Online Learning*. The Sloan Consortium: Needham, MA.

[3] Bielawski, L. & Metcalf, D. (2003). *Blended e-learning: Integrating knowledge, performance, support, and online learning*. Amherst, MA: HRD Press.

[4] Blackboard© INC. (2008). Blackboard© & the University of Cincinnati [Online]. Available: http://www.Blackboard©.com/CMSPages/GetFile.aspx?guid=90ffd3e6-1073-4815-9ff8-5e34975696b4 [Accessed September 2010].

[5] Borden, J. (2011). *The Future of Online Learning*.

[6] Bowles, M. S. (2004). *Relearning to e-learn: Strategies for electronic learning and knowledge*. Melbourne, Australia: Melbourne University Press.

[7] Bradford, P., Porciello, M., Balkon, N. & Backus, D. (2007). The Blackboard© Learning System: The Be All and End All in Educational Instruction? *Journal of Educational Technology Systems*, 35, 301-314.

[8] Brown, A., & Campione, J. (1990). Communities of learning and thinking, or a context by any other name. *In D. Kuhn (Ed.), Developmental perspectives on teaching and learning thinking skills* (Contributions to Human Development Series) (pp. 108-126)). Basle, Karger.

[9] Creswell, J. (2009). *Reaserch design: Qualitative, quantitative and mixed methods approaches*. Los Angeles: Sage Publications.

[10] Dunn, T. (2003). Preparing preservice teachers for diverse classrooms: Influences of CMS on beliefs and instructional practices. *Proceedings of World Conference of E-Learning in Corporate, Government, Health Care and Higher Education*, 2003 444-447.

[11] Everson, M. (2009). 10 Things I've learned About Teaching Online. Retrieved 7/04/2012 http://elearnmag.acm.org/featured.cfm?aid=1609990

[12] Green, S. M., Weaver, M., Voegeli, D., Fitzsimmons, D., Knowles, J., Harrison, M. & Shephard, K. (2006). The development and evaluation of the use of a virtual learning environment (Blackboard© 5) to support the learning of pre-qualifying nursing students undertaking a human anatomy and physiology module. *Nurse education today*, 26, 388-395.

[13] Greenagel, F. (2002). *The illusion of e-learning: Why we are missing out on the promise of Technology*. Phoenix: League for Innovation in the Community College White Papers.

[14] Guest, G., Bunce, A. & Johnson, L. (2006). How many interviews are enough? *Field methods*, 18, 59-82.
[15] Guri-Rosenblit, S. (2005). Eight paradoxes in the implementation process of eLearning in higher education. *Higher Education Policy*, 18, 5-29.
[16] Hartnett, M., Bhattacharya, M. & Dron, J. (2007). Diversity in online learners: Searching for differences that may matter. In, 2007. *IEEE*, 899-900.
[17] Heaton-Shrestha, C., Gipps, C., Edirisingha, P. & Linsey, T. (2007). Learning and e-learning in HE: the relationship between student learning style and VLE use. *Research Papers in Education*, 22, 443-464.
[18] Heirdsfield, A. M., Walker, S., Tambyah, M. M. & Beutel, D. A. (2011). Blackboard© as an online learning environment: what do teacher education students and staff think? *Australian Journal of Teacher Education*, 36, 1-16.
[19] Jones, N., Blackey, H., Fitzgibbon, K. & Chew, E. (2010). Get out of MySpace! *Computers & Education*, 54, 776-782.
[20] Keengwe, J. & Kidd, T. T. (2010). Towards Best practices in Online Learning and Teaching in Higher Education. *MERLOT Journal of ONLINE Learning and Teaching in Higher Education* 6, 533-541.
[21] Klein, H. J., Noe, R. A., & Wang, C.W. (2006). Motivation to learn and course outcomes: the impact of delivery mode, learning goal orientation, and perceived barriers and enablers. *Personnel Psychology*, 59(1), 665–702.
[22] Landry, B. J. L., Griffeth, R. & Hartman, S. (2006). Measuring student perceptions of Blackboard© using the technology acceptance model. *Decision Sciences Journal of Innovative Education*, 4, 87-99.
[23] 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, 864-873.
[24] Mason, R. (2006). The university – current challenges and opportunities. *S. D'Antoni The Virtual University: Models & Messages*.
[25] McLoughlin, C., & Luca, J. (2000). Cognitive engagement and higher order thinking through computer conferencing: We know why but do we know how? Retrieved 28 July, 2007, from htpp://lsn.curtin.edu.au/tlf/tlf2000/mcloughlin.html
[26] Pirrone, R., Cannella, V. & Russo, G. (2009). A map-based visualization tool to support tutors in e-learning 2.0. In: *2nd Conference on Human System Interactions*, HSI '09, 2009. 482-487.
[27] Pishva, D., Nishantha, G. & Dang, H. (2010). A survey on how Blackboard© is assisting educational institutions around the world and the future trends. In: *The 12th International Conference on Advanced Communication Technology (ICACT)*, 2010. IEEE, 1539-1543.
[28] Robertson, H. J. (2003). Toward a theory of negativity teacher education and information and communications technology. *Journal of Teacher Education*, 54, 280-296.
[29] Romiszowski, A. J. (2004). How's the e-learning baby? Factors leading to success or failure of an educational technology innovation. *Educational Technology-Saddle Brook Then Englewood Cliffs Nj-*, 44, 5-27.
[30] Schroeder , A., Minocha, S. & Schneider†, C. (2010). The strengths, weaknesses, opportunities and threats of using social software in higher and further education teaching and learning. *Journal of computer assisted learning*, 26 159-174.
[31] Teddlie, C. & Tashakkori, A. (2009). *Foundations of mixed methods research*. Thousand Oaks, CA: Sage.
[32] Trentin, G. (2009). Using a wiki to evaluate individual contribution to a collaborative learning project. *Journal of Computer Assisted Learning*, 25, p 43-55.
[33] Turner, W. L. & Stylianou, A. C. (2004). The IT advantage assessment model: Applying an expanded value chain model to academia. *Computers & Education*, 43, 249-272.
[34] Voci, E. & Young, K. (2001). *Blended learning working in a leadership program*. Industrial and Commercial Training.
[35] Wallace, R. M. (2003). Online Learning in Higher Education: a review of research on interactions among teachers and students. *Education, Communication & Information* 3, 241-280.
[36] Wood, A., & Smith, M. (2005). *Online communication. Linking technology, identity and culture* (2nd Ed.). London: Laurence Erlbaum Associates.
[37] Yang, Y. & Cornelius, L. F. (2004). Students' Perceptions towards the Quality of Online Education: A Qualitative Approach. *Association for Educational Communications and Technology*, 2, 861-877.

Teacher Agency and Student Autonomy in Inquiry-based Mobile Learning Trail

Esther TAN[a*], Hyo-Jeong SO[a] & Xujuan ZHANG[a]
[a]*National Institute of Education, Nanyang Technological University, Singapore*
*esther.tan@nie.edu.sg

Abstract: This research study investigates students' capacity at greater autonomy in an inquiry-based mobile learning trail in relation to the agent of the teacher (e.g., instructional design, facilitation and scaffold support). To afford a more coherent study, narrative interviews and web-based data capturing student-student and teacher-student interaction were obtained for analysis. Data were analysed based on three key areas: (a) the design of the mobile learning trail in facilitating autonomous learning, (b) the interaction with the teachers and (c) the collaboration with peers. Overall findings showed that students' capacity to engage in autonomous learning rests on learning trail design, collaborative efforts and an awareness of teachers' "presence". The teachers cited students' profile, the motivational factor and instructional design as important determinants for autonomous learning. In conclusion, we argue that students' capacity for more autonomy in mobile learning does not necessarily lie in a decrease in teacher's control, but rather, it is contingent on student readiness, learning design, technological mediation, as well as, the community of learners.

Keywords: Teacher agency, student autonomy, inquired-based mobile learning

Introduction

The advent of mobile technologies has dramatically revolutionized the conventional role of teachers and students. Harnessing the affordances of technology-mediated cognitive tools to engage learners, enhance learning effectiveness, empower and enable synchronous and asynchronous interaction and collaboration is believed to bring about greater student autonomous learning. Teachers presumably best function as facilitators to scaffold the learning milestones and to mediate technological support to enhance learner autonomy. However, what essentially facilitates the occurrence of more student autonomy and/ or how teachers can orchestrate such learning situations, are needful areas for more intense research and investigation. On supporting learner autonomy, Black and Deci (2000 as cited in [1]) liken this phenomenon to a situation where the learners are equipped and empowered to make autonomous decisions in the learning process given the accessibility and availability of "pertinent information and opportunities for choice" (p.28). And In theorizing mobile learning, Sharples, Taylor, and Vavoula [2] surface "control and context" as two of the key areas for reflection: *control* is distributed across multiple elements from teacher, peers, technologies to environmental artefacts, and *context* is constructed by the learners interacting with the environment, which comprises of communities of learners and all mediating technologies. And in our context of inquiry-based mobile learning trail, the imminent challenge would be to apportion the right measure of teacher "presence" without jeopardizing student's capacity at autonomous learning. Hence, supporting learner autonomy is not a simple equation of decentralizing teacher agency and control; rather, it calls for an informed action taking into account all contextual elements in the said learn setting.

1. Theoretical Framework

To encapsulate the nature of the inquiry-based mobile learning trail, the contextual elements and the social actors (i.e., the teacher and the students), we employ situated cognition to make sense of how learning takes place for two fundamental reasons. First, the key theoretical premises of situated cognition afford an insight into the dynamic interplay of critical constructs such as the learning activities, all mediating “tools” (e.g., physical environment, social actors, artefacts, etc.), and importantly, the cultural and social practices in the learning context. According to Brown, Collins and Duguid [3], “Knowledge is situated, being in part a product of activity, context, and culture in which it is developed and used” (p.32). Second, the theoretical underpinnings of situated cognition provide a conceptual framework to make sense of student autonomy and teacher agency in a mobile learning trajectory for it is impossible to discuss sensibly the changing roles of teachers and students without making reference to the contextual configurations and their relations. Brown and Duguid [4] contend that, "One of the powerful implications of situated learning is that the best way to support learning is from the demand side rather than the supply side…” (p.8). This has strong implications on the role of the teacher and the place for students’ autonomy in the learning process. On the design of the learning environment, Choi and Hannafin [5] advocate a shift from organizing and sequencing content to creating and designing environments that “induce, then facilitate, understanding” (p.67). The functional role of the teacher here would be to allow an unstructured space within the structured learning environment, whereby learners have the liberty to exercise judgment, set new learning intent and pursue new inquiries/ interest areas. On this note, Snow (1994 as cited in [6]) posits, “we must not only learn in context but also by context” (p.84). Learners are empowered to respond to contextual changes within the framework that guides their inquiry process.

Apart from a theoretical emphasis on learning in an authentic platform as against “decontextualized contexts”, situated cognition also exemplifies the importance of “cultivating learning processes versus learning outcomes” (p. 53) [5]. Here, it presupposes two significant groups of players in the learning process. One is the teacher-student and two, the student-student. First, it inherently implies a marked change in the role of the teacher – from a knowledge dispenser to a facilitator of students’ learning processes (Bednar et al., 1991; Duffy & Jonassen, 1991; Winn, 1993 as cited in p. 67) [5]. And facilitation can take on varying forms such as “modeling, scaffolding, coaching and guiding, collaborating, fading” and via different technology-mediated cognitive tools and resources (p. 63) [3]. It is also the onus of the facilitator to develop in the learners the capacity and the ability to perform a knowledge and skill transfer across varying contexts. Second, the individual learner’s interaction and collaboration with his/ her counterparts form a critical phase of this collective learning enterprise. Thus, we recognize that student autonomy is both enabled and shaped by a host of factors at play in a learning situation, of which, the mediation of cognitive tools and collective cognition play a definitive role. The teacher, thereby, assumes a more significant and complex function - a designer, a mediator, and sometimes, a participant and collaborator of the learning enterprise. In a nutshell, the fundamental role and responsibility of the teacher would be to “design the situation” (p.5) [7]: engineering the learning environment and ensuring the availability and the accessibility of technology-mediated cognitive tools and resources to bring about the desired learning outcomes.

In the context of an inquiry-based mobile learning trail, the purpose of this qualitative research study is to examine how the agent of the teacher in learning design, in appropriating technology-mediated cognitive tools to support collaborative mobile learning, and in apportioning the measure of teacher "presence" can impact students' capacity to exercise autonomous learning. We are also interested to identify which of those above-mentioned elements (e.g., learning design, technology, facilitation, student readiness) form the key determinants that shape student autonomy and to what measure.

2. Methodology

2.1 Research Background

Building on our previous research efforts to promote collaborative knowledge co-construction on mobile learning trails, the present research study seeks to explore the teacher's role in facilitating more student autonomy in an inquiry-driven mobile learning trajectory leveraging on the rich affordances of the physical environment and technology mediation. Aligned with the theoretical premises on situated learning, the mobile learning trail was designed with a focus on inquiry task-type and space for more independent learning via collective effort, technological mediation and teacher facilitation. Figure 1 illustrates the three-pronged approach - F.A.T (Facilitation, Activity Design, Technology), a design framework we conceptualized to guide our trail design and implementation. This holistic approach sees activity design as the primary driver of the other two equally critical components - facilitation and technological mediation in the design of the learning situation.

Figure 1: The 3-pronged Approach toward Mobile Learning Trail Design

2.2 Design Consideration

The trail tasks were co-designed by the researchers and collaborating teachers in the integrated humanities department. The F.A.T. design framework guides the design process (i.e., the overall focus on inquiry-based learning) and the range of activities primarily determine the type of technological tools and the features of web-based platform to support the collaborative learning space, the facilitation and the communication process. The design of the trail seeks to promote interdisciplinary inquiry-based discourse. This move sees an unprecedented rich integration of History and Geography with the intent to develop a holistic understanding of the body of cognitive and procedural knowledge and skills in the integrated humanities. All trail task questions point to an ultimate problem statement where learners will need to see relationships across the findings to the various task questions and eventually evaluate and synthesize shared knowledge and understanding as a collective body. As illustrated in Table 1, trail tasks range from performative (application) to knowledge generative and knowledge synthesis where the findings and inferences should enable learners to respond to the overarching big question on Sentosa's role in British defence plan.

Table 1: Examples of Tasks at Fort Siloso Mobile Learning Trail

Big Question: What is the role of Sentosa in the British's big plan of defence?		
Learning Station	***Task Type***	***Task Description***
A	Performative	T1. Determine the direction of the guns using the iPad compass.
	Knowledge Generative	T2. Describe the dimension of the tunnel and state its purpose.
B	Performative	T3. Locate the "Stealth" boat entering the harbor entrance.
	Knowledge Generative & Synthesis	T4. Explain why the previous artillery gun (Area A) and this one are pointed in the same direction.
		T5. Give reasons for the British's plan to locate the tower at area B. Describe the role and purpose of the tower and the guns.

Facilitation forms a critical determinant on students' capacity for autonomous learning. Three main modes of facilitation were put in place. First, a web-based platform was designed and developed to host all trail activities with embedded apps (e.g., digital map) for students to carry out their activities. Each team (max. 4 students) had an iPad, and students were also able to upload their findings and collated artifacts onto their teams' respective web pages. Further, they were also able to communicate with other teams through the feedback feature; giving comments and/ or suggestions. Second, trained facilitators were assigned to all four activity-stations and teachers were also present to monitor students' progress. The physical presence serves mainly to provide students a sense of assurance of aid should they be confronted with any major difficulties in an outdoor situation. Facilitators were also briefed to adhere to more unstructured questioning techniques to avoid constricting students' capacity to leverage on situational resources in the learning process. Third, apart from face-to-face facilitation, virtual teacher facilitation was deemed necessary to provide immediacy of facilitation, which is of significance in a mobile learning trail. Virtual facilitation via the public space feature in the web-based platform affords teacher-student interaction and communication.

2.2.1 Participants

The collaborating institution is also a member of the FutureSchools@Singapore project. The school leverages on its 1:1 computing initiative to create a technology-rich learning environment and advocates a small class size of 20–25. The mobile learning trail took place at Fort Siloso, Sentosa Island in March 2011. Participants of the trail were two classes of secondary one students - Class A (total class size = 20) included mostly high-achieving students while Class B (total class size = 22) included mixed-ability students.

2.3 Data Collection and Analysis

The focused group interviews with teachers and students were rendered necessary to find out about students' and teachers' perceptions of their inquiry-based mobile learning experience, as well as, the actual occurrence of discourse amongst students and between students and teachers. Post-trail focus group interviews with ten students (randomly selected five from each of the two classes) and a one-to-one interview with six collaborating teachers were held. The interview questions were semi-structured to solicit feedback on three critical areas, namely, trail activities and collaboration efforts, facilitation and technological mediation. In addition, relevant excerpts of data captured on the web-platform of students' interactions with other groups and the teacher facilitators were also cross-examined to afford a more accurate insight into the research inquiry on student autonomy and teacher agency in inquiry-based mobile learning.

3. Findings

Aligning with the conceptual framework on situated cognition, we examined and analyzed the corpus of data in relation to the three key themes: (a) the activities in relation to contextual elements (b) the mediating tools and (c) cognitive apprenticeship.

3.1 Students' Narratives

3.1.1 Impact of Trail Design and Collective Efforts on Students' Capacity for Autonomous Learning

Students felt that task questions integrating Geography and History, opened up the platform for further inquiries, generation of ideas and hypotheses; triggering a chain of discussion. One participant voiced that the course of finding answers to the history questions, had enabled them to see how geographical factors affected human decisions – 'why they did what they did' and this provided the explanation to the historical events back then. Students felt that there was better engagement with the abstract concepts and a stronger sense of ownership of their learning during the trail without the constant physical presence and supervision of the teachers.

Next, for the majority of the students, the inquiry-based approach lends itself better for collaborative efforts over individual undertaking. On this note, Mark contended that "if you have multiple people working on the same problem …so even if you get stuck, maybe another person know how to do it …". Another participant, Cayden concurred that the very act of coming together to resolve an issue would inevitably give rise to a convergence of human thinking resources. It promotes distribution of the think processes and gave rise to the possibility of multiple perspectives on a common task and also assistance within group when one is confronted with tougher task questions. Other participants felt that it had increased their overall learning and thinking capacity when group converged again to share their respective findings; this allowed them to learn from the explanation of the fellow team mates who were better with a particular task and /or subject area. However, not all students were optimistic about collaborative efforts, Tiffany recounted that, "some teams are not very receptive to ideas… don't disturb us…go away…we are doing our work …". Some teams conceived of inter-group collaboration as a form of interference; delaying their work processes. Another reservation about collaborative learning lies in the issue of *reciprocity* where students perceived that they could be short-changed by another group, "some of my silly group members go and give them the answers...and some groups also, they have nothing to say". We attribute this to the gap in belief and actual practices. That is, although students believe in the benefits of collective cognition, in practice, they are more attuned to individual display of effort and performance.

3.1.2 Technological Mediation and Student Autonomy

The availability of the feedback and comment functions in the web-based platform enabled the immediacy of facilitation and asynchronous collaboration. Isaac related their team's experience with feedback and comments from other teams and how it gave them some form of 'directions' to re-attempt task questions and re-work through their own work processes: "there's was one point we got stuck then we resort to getting inspiration from other teams…we try to understand how they got the answer and then incorporate it …". Likewise, the provision of feedback allows students a second chance to rethink through their findings amid the rich physical affordances. Lucas recalled, "let's say you make any mistake, the teacher will send you a message". The 'alert' function cum instant feedback from teachers permitted a review of work processes. Immediacy of teacher facilitation enlarges students' independent learning space and thereby increases students' capacity to take control of their own learning journey in a mobile learning environment. Students are given more autonomy to re-evaluate their initial findings and re-negotiate meaning.

3.2 Teachers' Narratives

3.2.1 Activity Design Shapes Inquiry-based Learning Processes

At the cognitive level in relation to greater student autonomy in inquiry-based learning, the collaborating teachers spoke on the significance of 'situated learning' experiences to foster autonomous learning and inquiry-based discourse. Mr. Loh explained, "Ground experiences can never be replicated… important to view the authentic documents at the site rather than online – to develop empathy and multiple perspectives; giving space to the internal voice, queries, hypotheses..." Ms. Lee further added, "the mobile device also increased the proximity of the learners to the object of inquiry". This allows the students an up close and personal encounter in the course of their inquiries. Mr. Seah commented that "the use of iPad allows students' mobility and accessibility to any work tools and at the same time, enables students to communicate and collaborate, in turn, teachers were also able to capture their learning process".

On the notion of mobile learning and inquiry-driven curriculum, Ms. Lee felt that "mobile learning on its own, cannot be a stand-alone instructional tool. The basics should be done in class ..., pre-trail lessons…and after that trail, post-trail". Further, Mr. Loh stressed that it is needful to strengthen that link to "review how this trail fits into the larger picture of things, i.e., the entire curriculum". Hence, for students to take on an inquiry-driven learning trajectory on the day of trail, teachers rendered preparatory work as a necessary phase to equip and empower students to fully benefit from the mobile learning scenario, taking charge of their own learning.

3.2.2 Technological Mediation and Facilitation as a Means to Enhance Student Autonomy

One of the means of increasing student autonomy and appropriating teacher involvement was the provision of technological cognitive tools. Ms. Ang observed that the web-based platform was "a good communication platform for teachers to be engaged in the whole collaboration - heighten interactions, give instant feedback and able to gauge students' progress, difficulties". To which, Mr. Yeh concurred, "technology makes possible a virtual facilitator, which made possible for students to receive guidance and on the spot to re-look at their options. Mr. Yeh continued, "the broadcast feature helped them stay on the right track when they are pre-occupied or intentionally go off tangent". Another teacher commented that "interaction between groups made possible and between teachers and students". Ms. Lee shared that "the web-based platform has made it very viable for learning, and enables interaction and the immediacy of facilitation. Layout also enabled teachers to participate in the trail, and I was able to monitor all their responses from where I was, without moving around."

Teachers felt that the design of the inquiry-driven trail, the provision of technological tools and online facilitation make it possible for teachers to "take a step back, observe how kids work and give them a chance to maneuver their way through and reach the destination – give them more ownership of their own learning. Students take pride in their work", remarked Ms. Teh. On the self-same note, Ms. Lee felt that there was "less front loading and transmitting of content - more self-directed learning. This experience changed our roles as practitioners in the classroom – even seeing their responses (if incorrect) and even if they fail it doesn't make me anxious." The teachers felt that letting go of control and the expectation for correctness and performance would liberate that space for greater student autonomy in learning.

3.2.3 Impact of Teacher Presence on Student Motivation and Autonomous Learning

On the significance of teacher presence, be it virtual or physical, Mr. Yeh observed that "broadcast messages create the kind of atmosphere for students – motivate them as they know there's someone out there responding to them unlike cyberspace". Likewise, Mr. Seah observed that the high level of engagement from the students had to do with the manageability of tasks and the awareness that teachers are "present" (virtual and face-to-face) to assist them should they stumble in the course of accomplishing their set goals. This explains their eagerness and motivation moving from one learning station to another. Mr. Yeh noticed another interesting phenomenon, "they do not ask for answers through the platform, physically yes, if they meet you, they ask for clues." He felt that virtual facilitation renders a different form of teacher involvement, "If answers seem general, try to find out whether they understood task requirements, try to elicit more from them and then guide them." Overall, teachers were positive that students obtained a greater sense of ownership of their work processes in the learning trail experience. However, teachers expressed the need for more concrete measures to be taken in order to increase student autonomy and learning effectiveness, namely, soft skills in collective undertaking of tasks and, questioning and inferential skills in engaging their counterparts during the collaboration process.

4. Discussion and Conclusion

As exemplified from the findings, the measure of teachers' presence and participation ought to be weighed in the context of the learning situation, the prevailing socio-cultural practices and the profile of the participants. In sum, student readiness to assume more autonomy in inquiry-based mobile learning hinges on a variety of contextual factors. First, the relevance of the learning activities ought to align with the desired learning outcomes, as well as, the students' profile and capacity for autonomous learning in such situated learning contexts. And of significance would be to see the one-day learning trail as a continuum of the entire curriculum- teacher support and scaffolds to increase student capacity at autonomous learning commences at the very stage of the larger curriculum design. For students to be able to benefit from autonomous learning in a mobile learning context, autonomy support begins with the day-to-day instructional program. Students' confidence and comfort level to initiate and pursue inquiries, to make informed decisions, and to conduct constructive interaction discourse is a gradual developmental process, orchestrated by the teacher. Second, the vicarious presence of the teacher – virtual and face-to-face still has its rightful place. Students still need the endorsement of an authoritative figure - the teacher. As Ellis (1993 as cited in [8]) puts forth, “The teacher thus has the role of a group member that has the option to qualify the dialogue through questions” (p. 22). And importantly, in an outdoor learning situation, students commence at different levels of understanding in the collective meaning-making process before converging at a common shared understanding. In the research study, the appropriation of the measure of assistance and scaffolding was made possible based on the kind of answers and feedback students pose on the web-based platform. And further, the teacher’s presence as a participant and a collaborator serves as a form of facilitating and regulating the exercise of autonomous learning on the part of the students. Third, student autonomy also rests largely on the presence of the collective body of their fellow workmates and the collaborative learning space. As evident in the narratives, students (esp. the high ability group) felt that they were not very comfortable with the idea of collaborative efforts as some still held on to the notion of individual merit and performance. Conversely, mixed ability group was more open to sharing of ideas and findings. This inadvertently implied that the socio-techno learning space to a considerable measure dictates student readiness to become agents of their own learning.

To conclude, the analysis and synthesis of the data findings surface significant implications pertinent to the design of the learning situation, the agent of the teacher in shaping student autonomy in such situated learning contexts. The artful balance of teacher agency and student autonomy requires a sound understanding of the content and context of learning, and the appropriation of relevant technological mediated tools and facilitation.

Acknowledgements

This research is supported by the FutureSchools@ Singapore project under the National Research Foundation’s (NRF) Interactive and Digital Media (IDM) in Education Research and Development (R&D) Programme.

References

[1] Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 33-42.

[2] Brown, J. S., & Duguid, P. (1993). Stolen knowledge. *Educational Technology, 33*(3), 10-15.

[3] Choi, J.I. & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research & Development*, *43* (2), 53-59.

[4] 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.

[5] Jonasson, M. (2011). Framing learning conditions in Geography excursions. *International Education Studies,* 4 (1), 21 - 29.

[6] Sharples, M., Taylor, J. & Vavoula, G. (2005, October). *Towards a theory of mobile learning*. Paper presented at 4th World conference on mLearning, Cape Town, Africa.

[7] Stefanou, C. R., Perencevich, K. C., Dicintio, M. & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist*, *39*(2), 97–110.

[8] Uden, L. (2007). Activity theory for designing mobile learning. *International Journal, Mobile Learning and Organisation, 1*(1), 81-102.

Using Kinect to Track Learning Behavior of Students in the Classroom as Video Portfolio to Enhance Reflection Learning

Chuan-Ling LIEN*, Ching-Yi HUANG, Chin-Yeh WANG & Gwo-Dong CHEN
Department of Computer Science and Information Engineering, National Central University, Taiwan
*ayaling77@gmail.com

Abstract: We propose a novel approach to produce video portfolios in the classroom, allowing student's whole body, which was captured by a Microsoft Kinect appear in a rich and context-sensitive background. The purpose of the study was to investigate this proposed system could impact on various learning style students' cognition and metacognition. The system makes the student have chance to examine and evaluate on one's own and then promotes metacognition. Experimental results indicated a variety of learning styles do affect the performance of learners, especially reflective and sense of style, metacognitive have a significant upgrade capacity by using the system.

Keywords: Learning style, video portfolio, metacognition, authentic learning

Introduction

Video Portfolio has been widely used to improve the quality of supervision, evidence and transfer in education. Video recorders can capture many interactions. Recorded video can be assessed formatively with the goal of improved teaching, and they may also be assessed to yield a summative score or judgement [2]. Students can also reflect more deeply on their performance by reviewing a particular segment many times at a different location.

Metacognition is an important skill that could be improved by video portfolio. Black & William [3] stated that the four pillars of the e-portfolio (metacognition, authentic tasks, contextual feedback and student responsibility) seem to clear up the effectiveness of formative evaluation. Metacognition is a skill of self-monitoring and reflecting on our own mental process, and could contribute to critical-thinking skills. With good metacognition skills, students can find out important information needed to solve a problem by analysis and inference. They can also know their own demands and resources and consider using the appropriate strategies at different times to effectively learn. With the support of video portfolio, students can have opportunities to practice self-monitoring and reflect critically on their experience.

Currently, video portfolio has been adopted to solve this problem. However, in the classroom, it is difficult to include authentic context in the video recordings that represented students' performance. To ensure the recorded videos look good with rich and proper background for review and reflection, the background needs to change accordingly depending on what themes and students are going on.

We propose a novel approach to let students' whole body, which was captured by a *Microsoft Kinect* appear in a rich and context-sensitive scene. Students could watch themselves appeared in the screen with a specific background that is designed by instructors, and perform what they learnt. Later, students can review their own performance under the guidance of their teachers or alone. Rich background with their figures of the video portfolio could allow them perform reflection and enhance metacognition. Moreover, learning style was taken seriously in recent years. When teaching style coincides with students' learning style, their learning could be easier and effective [19]. We also examine what kind of learning style that students have could benefit from the approach significantly.

1. Related Work

1.1 Metacognition

Metacognition is extremely important, but really difficult to teach and assess. Part of the reason why student cannot effectively use strategies and cannot become an active and independent learner is the lack of metacognition. To become self-directed learners, learners should develop their metacognition [22].

Flavell [10] found out from the study of human cognition that learners usually displayed bad cognition due to their inability to monitor their cognition process and adjust it later to match specific objectives. Metacognition is the ability of an individual to self-supervise, in terms of their problem-solving ability, and is an important prerequisite [13]. With good metacognition, students can clearly understand self-monitoring, self-discipline, self-correcting, and self-assessment while they are solving problems. Schraw, G. & Dennison, R.S. [20] divided the concept of metacognition into five steps which could be improved individually:

- Planning: Planning, goal setting, and allocating resources prior to learning.
- Information Management Strategies: Skills and strategy sequences used to process information more efficiently (e.g., organizing, elaborating, summarizing, selective focusing).
- Comprehension Monitoring: Assessment of one's learning or strategy use.
- Debugging Strategies: Strategies used to correct comprehension and performance errors.
- Evaluation: Analysis of performance and strategy effectiveness after a learning episode.

Nevertheless, metacognition is situation dependant. Each individual is required to manage specifications while considering their strategy and adjusting their train of thought toward the situation that their strategy is going to used [13]. The video portfolio of what they performed could be useful if students can perform in authentic learning environments.

1.2 Authentic learning

On the other side, authentic tasks are vital to the e-portfolio that could result in the effectiveness of formative evaluation [3]. Suchmon [21] found that humans' cognitive activity is limited by the social context of their activities. However, knowledge learned in a school system is different from the cognitive ability that is obtained from reality [13]. The true meaning must be realized via practical actions, and students should know how to solve unknown problems by using environmental resources [14]. Therefore, concepts and rules must be understood through real experiences, which is not easy to be fulfilled in the classroom.

Similarly, metacognition should be cultivated in authentic environment. Brown, Collins & Duguid [4] pointed out that people invent effective strategies to solve problems in situations due to interacting constantly with a specific situation.

When students perform situational tasks, they can reflect on their learning results after learning through video portfolios. From video tape to digital memory, video recorders have been widely used to improve teaching and learning in the classroom. To change the method of cognition and enhance learning effect, authentic learning environments should be added to the classroom, allowing students perform in authentic environments and producing video portfolios for enhancing metacognition.

1.3 Portfolio

Portfolios are widely used to help students learning. Portfolio assessments of learning outcomes are more realistic and dynamic presentation than the written records examination, and the students can know their results clearly via portfolio [13]. Albert Bandura's Social learning theory mentioned that learners change individual behavior through observing and imitating others. Learners can better understand and analyze their advantages and disadvantages via comparing themselves to others. Through portfolios, teachers and parents can also view the child's advances, and give them timely help [5],[15],[16],[23]. In short, Paulson Leon & Paulson Pearl [17] mentioned that promotion of reflection, self-evaluation, self-understanding and the ability of metacognition is the most important objective that learning portfolios could be.

Portfolio is individualized data [1], purposeful data collection [18], visual process presentation [9], and authentic performance results [18]. Chang [6] revealed that the system of network learning files can produce many effects such as helping to control the learning process, reflecting on the advantages and disadvantages, and enhancing the growth, progress, and benefit of learning. Currently, video portfolio has also been frequently utilized to assist students learning and its effectiveness has been proved hugely. This study try to enhance the way that video portfolio was taken in the classroom. Students' performances and the according backgrounds are mixed together and recorded simultaneously as the video portfolio. Comparing to the portfolio recorded by DVRs, the video portfolio which is generated by our proposed approach could provide students authentic context.

1.4 Learning Style

Besides examine the impact that the proposed idea could be, this study also wants to realize if there is any difference while students with different learning styles. Here gives a brief description of learning style. Learning styles are various approaches or ways of learning. They involve educating methods that are presumed to allow individual to learn best. Several dozen learning style models have been developed. One of the most common and widely-used categorizations of the various types of learning styles is Fleming's VARK model [12]: (1) visual learners have a preference for seeing, (2) auditory learners best learn through listening, and (3) kinesthetic or tactile learners prefer to learn via experience—moving, touching, and doing. One learning style is neither preferable nor inferior to another, but is simply different, with different characteristic strengths and weaknesses [7].

Category of learning style allows teachers to prepare classes to satisfy students' preferences. Students can also use the model to identify their preferred learning style and maximize their educational experience by focusing on what benefits them the most. The proposed method to produce video portfolios might be only suitable for students with specific types of learning style. This study also statistically examines what kind of students could benefit from the new approach.

2. System Design and Implementation

2.1 System Architecture

In order to record student’s learning activities, we used an L-shape authentic learning platform with a vertical screen to display situation and a horizontal interactive touch table to allow students to immerse into situation and perform interactive tasks.

When we put teaching materials into the L-shape platform to do authentic learning, Kinect let the image of students into the vertical screen, and students can see their true performance. While learning activity to proceed, HyCam2 (screen recording program) is called additionally to videotape the screen. In this way, we have the materials for students to review and reflect after class.

To make students reviewing those recorded video portfolios easily, learning objectives and the Learning adventure book videos were cut into small episodes. Besides, those videos and meta-information were stored in a database. Then, we used the programming languages, PHP and MySQL, to store those data and provide students personalized review.

Fig.1 L-shape platform using situation


Fig. 2 The system using situation

2.2 Instructional Design

Our teaching material is a designed situational game which makes students feel immersive in learning and interact frequently with their peers. Under this learning environment, teachers and students can carry out the group activities on the both sides of the platform. Through the help of teachers and the cooperative learning with peers, students can engage in learning activities and gain learning outcomes. The game allows students to view their own learning process to enhance metacognition. Here we choose English as second language learning as objectives, and integrate the materials into game-based learning activities.

After students complete the learning activities, we classify and cut the videos propriately, and let students have their personal video portfolios. Those videos are divided into three parts: evaluation, films, learning adventure book. Evaluation is categorized by learning objectives of self-evaluation and self-performance. Films are related to learning objectives, personal films and situations. The personalized video portfolios present the films of students plus narrative description. Following are the steps that students review the video portfolios. Firstly, they login the system via entering their own accounts and passwords. There is an easily understood instruction which contained all the evaluation, films, and learning adventure book on the left side of the screen. We reminded students to fill out the “learning objectives of self-evaluation” to realize how much they have learned in that objective. If there is inadequate, the system will recommend students intimately to review the video portfolio again. The demonstration of teachers and outstanding films of peers will be marked with special tags for students' references. Secondly, students watch their own videos, then complete the assessment, evaluate and reflect on their academic performance. After finished the two stages, students could choose any video which they were interested in or want to enhance. They can click the Learning adventure book to review the content they have learned quickly and completely. Finally, if students feel that the review and reflection completed, they can end the learning.

2.3 The design to boost metacognition

The experience of metacognition usually happened in the moment of cognitive failure. In the process of trying to solve the problems, we sometimes would know whether our performance was smoothly or not [11]. If students can make reflection to realize their own shortcomings, they would find out that there were many shortcomings can be improved. In this way, students can trigger the learning motivation. However, students usually forget or neglect their personal performance due to tension. Therefore, we should have a complete record to allow students review their personal performance.

The system could provide a possible solution. The system used the screen recording program to record the performance when a student carries out activities. In this way, we don't be afraid not filming to the front of students and don't need extra-manpower to control the camera. Students can use the system after class on their own computers. The films of the record of student carrying out activities will import into the system. We utilize the "video categories" to meet the demand of viewing learning performance. Students can click their peers' films to exchange study, or click the sample films of teacher in the "learning objectives" to review again. Furthermore, we make books of "Learning adventure book" by their learning record in situational stage. Using these books, it makes students clarify the overall concept. Besides, we added the learning adventure book and evaluation into our tools which included students' self-reflection and feedback of teachers.

Through the image recording and complete planning, the system brings some new possibilities for situational learning.

- Self-assessment of the effectiveness of learning: By fill out the "self-assessment of learning objectives," students can reflect on their learning situation, and more understanding of what objectives should be completed, and then click on the "learning target film to strengthen their own lack of part. In this way, students can promote their metacognitive skills of self-reflection, self-evaluation, and self-understanding.
- Self-monitoring:"Personal videos" provide students watch their performance and observe something they did not notice before and the part that they did not perform well on. Fill out the "Assessment of self-expression" assists students to view their performance again, and to stimulate students' understanding and reflection.
- Imitation and learning: Students see videos of other peers can learn the advantages of others, strengthen the impression of learning content, reflect on their own whether the same needs to be improved.
- Integration concept: An "learning adventure book" that presents the whole activity in a storybook style not only increases the lesson's appeal but also enhances the students' level of understanding. The student is the protagonist, providing them with a spectacular sense of actually being there. It can also be provided for the parents, so they could know what their children have done.

3. Experiment and Results

Here we would like to examine whether the proposed system could impact on student's cognition and metacognition. Besides, we also examine what kind of learning styles that students processed could significantly benefit from the proposed system.

3.1 Procedures

We designed an experiment to find out the outcome of our system. The experiment was held at a university in Taiwan. The subjects were 15 university students. Before conducting the teaching activities, they first do a Metacognitive Awareness Inventory (MAI) scale of the Regulation of Cognition to detect their habits and the degree of metacognitive awareness [20]. This is a 35-question questionnaire. In addition, there is another questionnaire (Learning Styles) with 44 questions to understand the students' behavior and habits of learning [8]. Then the subjects conducted a 40-minute learning activity. Five days later, we invited them to use the system again. Before using the system we provided 5-minute introduction to the system and experimental procedures. Next, we let them use the system for 20 minutes. After using the system, the subjects conducted the same MAI questionnaire that was to understand the effectiveness of learning after using the system and compare the correlation between effectiveness of metacogniton and learning style. Besides, we let them write satisfaction questionnaire to realize which functions are efficacious. Besides, we also videotaped the whole learning process to gain the observational data.

3.2 Result Analysis

Table.1 Learning styles and its corresponding point of pretest and posttest questionnaire

		Planning		Information Management Strategies		Comprehension Monitoring		Debugging Strategies		Evaluation		Overall	
		M	P	M	P	M	P	M	P	M	P	M	P
Active	pre	24	0.788	34.7	0.458	22	0.143	18.7	0.532	20.5	0.346	23.9	0.554
	post	24.5		36.2		25.7		17.5		21.8		25.1	
Reflective	pre	21.8	0.03*	33.2	0.021*	21.2	0.005*	19.4	0.313	19.8	0.086	23.1	0.011*
	post	26.3		38.4		26		20.2		22.2		26.6	
Sensing	pre	22.8	0.023*	33.5	0.019*	21.4	0.001*	19.2	0.8	19.9	0.009*	23.4	0.01*
	post	26.4		38		26.5		19.4		22.5		26.6	
Intuitive	pre	22	0.742	35.5	0.874	22.5	0.874	19	0.656	21	0.626	24.0	0.144
	post	20.5		34.5		22		17.5		19		22.7	
Visual	pre	22.9	0.141	34.6	0.151	22.1	0.023*	19.2	0.676	20.9	0.216	23.9	0.103
	post	24.7		37.4		25.5		18.8		22		25.7	
Verbal	pre	22	0.228	31.5	0.158	20	0.029*	19	0.68	17.8	0.131	22.1	0.144
	post	26		38		26.8		20		22.3		26.6	
Sequential	pre	22.5	0.05*	33.3	0.109	21.2	0.005*	18.9	1	19.4	0.037*	23.1	0.046*
	post	16		37.2		25.9		18.9		21.8		26.0	
Global	pre	22.5	0.537	33.8	0.187	22.5	0.178	19.6	0.907	21.3	0.741	24.0	0.479
	post	24.6		37.5		25		19.5		22.3		25.8	
Overall	pre	22.7	0.044*	33.8	0.033*	21.5	0.001*	19.1	0.878	20.1	0.041*	23.4	0.028*
	post	25.6		37.5		25.9		19.1		21.6		25.9	

*stand for significant different

Table 1 show that five abilities in the questionnaire: *planning*, *information management strategies*, *comprehension monitoring*, *debugging strategies*, and *evaluation*. A paired *t*-test was used to examine the effect. There is a significant difference between pretest and posttest in *Comprehension Monitoring* ($p=0.001$, $p<0.05$) indicating that assessments of learning strategies were effective after using the system. There is a significant difference between pretest and posttest in evaluation ($p=0.041$, $p<0.05$) indicating that their performance and ability of strategy analysis were improved after using the system. There is a significant difference between the pretest and the posttest in planning ($p=0.044$, $p<0.05$) indicating that *ability of goal setting* and *planning* was improved after using the system. There is a significant difference between pretest and posttest in *information management strategies* ($p=0.033$, $p<0.05$) indicating that students can improve their abilities of organization and information reconstruction effectively. However, there is no significant difference between pretest and posttest in *debugging strategies* indicating this system cannot assist students in correcting their strategies.

In five parts and the overall effectiveness of metacognition compared with learning styles, general discussion excluding no significant differences in the debugging strategies. The data show that *sensing style* has significant differences in all metacognitive skills. Reflective styles for planning, information management, strategies and comprehension monitoring have significant differences. Sequential style for the planning, comprehension monitoring and evaluation has significant differences. Visual and verbal style only in comprehension monitoring has significant differences.

4. Discussion

We find that the system has significant differences in information management strategies. We explore from the data of MAI and system usefulness questionnaire, and we find that there is effective assistance for students via teaching by situational games and films of peers who perform well or teacher's demonstration. Nevertheless, the significance of evaluation is barely satisfactory, exploring several students whose achievement is relatively low in this part. We find that they are not accustomed to reflect on their performances.

About learning style, experimental results show that there is mostly significant in reflection. This means that students who usually had a habit of independent thinking and judgment are able to exert the skills of self-reflection and learning from peers. Students with Sequential style have significant effectiveness as a whole. This may be because the system designed as entire teaching content but planned completely as different learning objectives. This makes students who accustom to understand and learn individually can easily get started. Students with Intuitive style have almost no significance in all respects. This may be that these films are too much alike to these students who don't like repetitive things. Therefore, they won't click these films respectively to do comparative learning. Furthermore, there is a particular finding. The Sensing style has quite significance in all respects except in the part of Debugging Strategies. The part of Debugging Strategies has quite demand of thinking and changing the learning strategy. Wherefore students who are thoughtful (EX: students with Reflective style), their performance will be relatively better. According to the satisfaction questionnaire of system, students who have been learning styles with Sensing and Visual have relatively high preference for the system. This is obviously due to this system that is suitable for their usual customary way of learning, so there is no obstacle in the learning.

5. Conclusion

We propose a novel approach to produce video portfolios in the classroom, allowing student's whole body, which was captured by a *Microsoft Kinect* appear in a rich and context-sensitive background. Students can review their own performance under the guidance of their teachers or alone to enhance their cognition or metacognition. Experimental results show that the proposed approach could significantly enhance students' metacognition on Reflective, Sensing and Sequential styles. In the future, we would suggest that providing references to students whose self-reflection ability is relatively inadequate. We also suggest that adding some interactive elements in the system to make students not lose their attention while reviewing these video portfolios.

Acknowledgements

This work is partially supported by the National Science Council, Taiwan, under the grant number NSC 101-2631-S-008-001

References

[1] Collins, A. (1992). Portfolios for science education: Issues in purpose, structure, and authenticity. *Science Education*, *76*, 451-463.
[2] Berg M. H., & Smith J. P. (1996). Using Videotapes to Improve Teaching. *Music Educators Journal*, *82*(4), 31-37.
[3] Black, P., & William, D. (1998). Inside the white box: raising standards through classroom assessment. *Phi Delta Kappa*, *80*(2), 139–148.
[4] Brown, J. S., Collins, A. & Duguid P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, *18*(l), 32-42.
[5] Carpenter, C. D. et al. (1995). Portfolio Assessment: Opportunities and Challenges, *Intervention in School and Clinic*, *31*, 34-41.
[6] Chang, C. C. (2001). A Study on the Evaluation and Effectiveness anAlysis of Web-based Learning Portfolio, *British Journal of Educational Technology*, *32*, 435-458.
[7] Felder R.M., & Brent R. (2005). Understanding student differences, *Journal of Engineering Education*, *94*(1), 57-72.
[8] Felder, F.M., & Soloman, B.A. (1991). Index of Learning Styles. North Carolina State University.
[9] Fischer, C. F., & King, R. M. (1995). Authentic Assessment: A Guide to Implementation. The Practicing Administrator's Leadership Series, Corwin Press.
[10] Flavell, J. H. (1976). Metacognition aspects of problem-solving,. In L.B. Resnick(Eds), *The nature of intelligence* (pp. 231-235). Hillsdale, NJ：Lawrence Erlbaum.
[11] Flavell, J. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry. *American Psychologist*, *34*, 906-911.
[12] Fleming, N. D. (2001) *Teaching and learning styles: VARK strategies*. Christchurch, New Zealand: N.D. Fleming.
[13] Huang, C.Y. et al. (2011). Using "Learning Video Portfolio" to Enhance Students' Metacognition in Authentic Learning. International Conference on Advanced Learning Technologies. Rome, Italy.
[14] Lave, J & Wenger, E (1991). Situated learning: Legitimate peripheral participation, Cambridge NJ: Cambridge University Press.
[15] Lambdin, D. V., & Walker, V. L. (1994). Planning for Classroom Portfolio Assessment, *Arithmetic Teacher*, *41*, 318-24
[16] Moss, P. A. (1992). Shifting Conceptions of Validity in Educational Measurement: Implications for Performance Assessment Review Of Educational Research, *62*, 229-258.
[17] Paulson, F. L., & Paulson, P. R. (1994). Assessing Portfolios Using the Constructivist Paradigm
[18] Paulson F. L., Paulson P. R., & Mayer C. A (1991). What makes a portfolio a portfolio. *Educational leadership*, 60-63.
[19] Renzulli, J. S. (1994). Schools for talent development: A practical plan for total school improvement. Mansfield Center, CT: Creative Learning Press.
[20] Schraw, G., & Dennison, R.S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475.
[21] Suchmon, L. A. (1987). Plans and Situated Actions: The Problem of Human-Machine Communication. New York: Cambrige University Press.
[22] Turner, G. Y. (1989). Promoting learning autonomy: helping students become independent learners, *Reading Horizons*, *29*(2), 110-116
[23] Nolet, V. (1992). Classroom-Based Measurement and Portfolio Assessment. *Assessment for Effective Intervention*, *18*, 5-26.

A Science Learning Environment using a Computational Thinking Approach

Satabdi BASU[a*], John S. KINNEBREW[a], Amanda DICKES[b], Amy Voss FARRIS[b], Pratim SENGUPTA[b], Jaymes WINGER[a] & Gautam BISWAS[a]
[a]*Department of EECS/ISIS, Vanderbilt University, U.S.A.*
[b]*Mind, Matter & Media Lab, Vanderbilt University, U.S.A.*

Abstract: Computational Thinking (CT) defines a domain-general, analytic approach to problem solving that combines concepts fundamental to computing, with systematic representations for concepts and problem-solving approaches in scientific and mathematical domains. We exploit this trade-off between domain-specificity and domain-generality to develop CTSiM (Computational Thinking in Simulation and Modeling), a cross-domain, visual programming and agent-based learning environment for middle school science. CTSiM promotes inquiry learning by providing students with an environment for constructing computational models of scientific phenomena, executing their models using simulation tools, and conducting experiments to compare the simulation behavior generated by their models against that of an expert model. In a preliminary study, sixth-grade students used CTSiM to learn about distance-speed-time relations in a kinematics unit and then about the ecological process relations between fish, duckweed, and bacteria occurring in a fish tank system. Results show learning gains in both science units, but this required a set of scaffolds to help students learn in this environment.

Keywords: Computational thinking, agent-based modeling, visual programming, simulation, scaffolding, inquiry learning

1. Introduction

Computational thinking (CT) describes a general analytic approach to problem solving, designing systems, and understanding human behavior [10,15]. It draws on fundamental concepts in computing and computer science to support practices (e.g., problem representation, abstraction, decomposition, verification) that are central to modeling, reasoning, and problem solving in scientific and mathematical disciplines [10,11].

Developing scientific practices and problem-solving skills requires sustained, immersive educational experiences, which can be implemented as learning progressions for science in K-12 classrooms [7,9]. CT can support such progressions in the K-12 science curricula by integrating its inherent domain generality with support for domain specific representations, reasoning, and analysis of real world problems across multiple domains [10,11,15]. Balancing and exploiting this trade-off between domain-generality and domain-specificity, however, presents an important educational design challenge.

Previous studies on integrating programming with K-12 science have pointed out a variety of similar challenges [3,4]. Curricula that have effectively addressed these challenges take advantage of *reflexivity*, which hypothesizes that learning programming in concert with concepts in another domain can be easier than learning each separately [6]. Several researchers have shown that programming and computational modeling can serve as effective vehicles for learning challenging science and math concepts [2,5,8]. Further, many programming and CT concepts parallel important aspects of STEM (Science, Technology, Engineering, & Mathematics) learning. For example, the creation of coherent, formal representations of scientific phenomena and mathematical representations of scientific laws are similar to object-oriented programming concepts of encapsulation, abstraction, and generalization. Conversely, the biological concepts of taxonomy and inheritance are the inspiration for class inheritance concepts in programming.

Leveraging the synergy between CT, science, and math learning, we have designed the Computational Thinking in Simulation and Modeling (CTSiM) learning environment and are implementing it using a learning-by-design progression. The learning environment combines visual programming and simulations to allow for flexible iterations between initial instruction in the science topic; modeling the appropriate entities and processes using a visual, agent-based computational framework; simulating and studying the behavior of the model; using explanation and argumentation skills to understand and verify the model; and, applying the developed model and science understanding to problem-solving tasks. The learning progression can be implemented over a progression of topics.

This paper presents the rationale for the learning environment design and a two-unit science sequence (kinematics and ecology) to demonstrate our computational thinking approach across domains. We describe an initial CTSiM study with 6th-grade students in a middle Tennessee public school. The results demonstrate the effectiveness of our approach, supporting the premise that students' conceptual understanding of science topics improves after engaging in CT-based curricular units. Finally, we present the categories of scaffolds used in the study and discuss the role these scaffolds likely played in the students' learning.

2. The CTSiM learning environment and curricular units

To support learning-by-design activities, CTSiM comprises three primary components [11]: (1) the Construction (C) world, (2) the Enactment (E) world, and (3) the Envisionment (V) world. The C world provides a visual programming interface where students build computational models for the science topics they are learning. It includes a library of visual primitives corresponding to agent actions, sensing conditions qualifying agent actions, and controls for regulating the program's flow of execution (e.g., conditionals and loops). Each visual primitive is defined in terms of an underlying domain-independent computational primitive. Students drag and drop these primitives and arrange them spatially to generate their computational models, as illustrated in Figure 1 for an ecology unit.

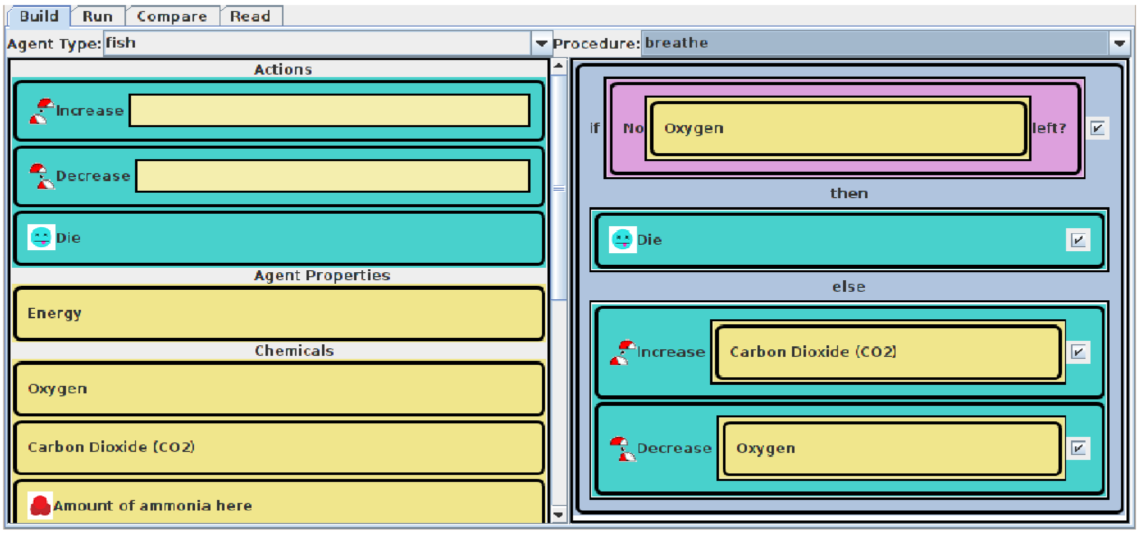


Figure 1: The Ecology unit Construction world with a 'breathe' procedure for 'fish' agents

In the E world, illustrated in Figure 2, students set initial parameter values and observe the Netlogo-based simulations corresponding to their models. NetLogo visualizations and plotting functions [15] provide students with a dynamic, real-time display of how their agents operate in the microworld simulation, thus making explicit the emergence of aggregate system behavior (*e.g.*, from graphs of a species population over time).
The V world provides students the opportunity to perform systematic experiments to compare their models' behavior against behavior generated by an "expert" model. This side-by-side comparison of plots and microworld visualizations for the two models makes it easier for students to investigate and revise their models. With proper scaffolding, we believe that the overall process of model construction, analysis, comparison, and refinement will help students gain a better understanding of science phenomena, mathematical concepts (e.g., rates), and computational constructs and methods.

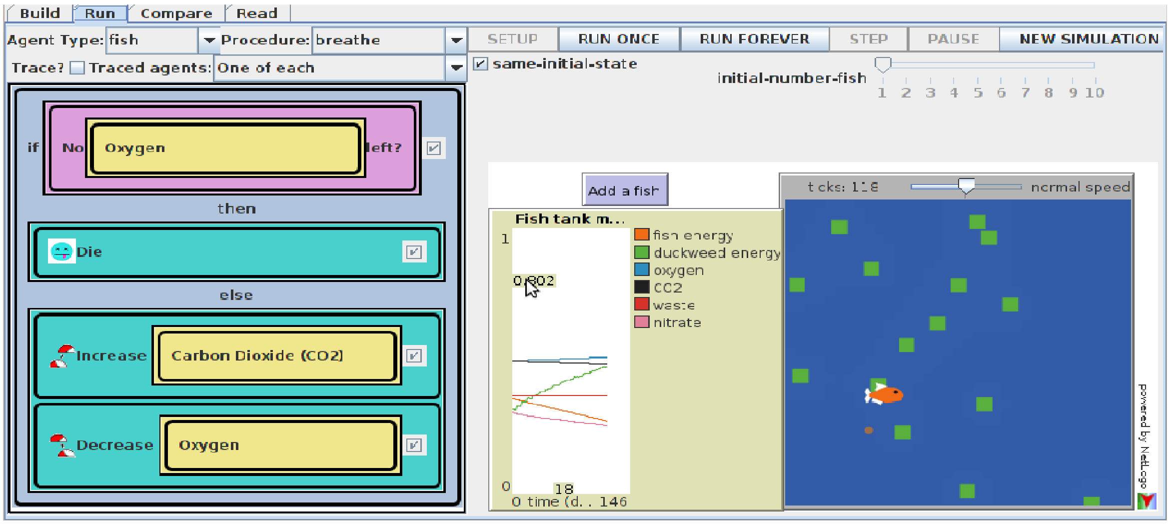


Figure 2: A screenshot of the Enactment world for the Ecology macro-unit

2.1 Kinematics Unit

Activities in the Kinematics unit were divided into three phases [12, 13].

Phase 1: Turtle Graphics for Constant Speed and Constant Acceleration – We introduced students to programming commands by showing them how to manipulate different elements in the user interface. Then, we asked them to generate algorithms to draw simple shapes (squares, triangles and circles) to familiarize them with programming primitives like "forward", "right turn", "left turn", "pen down", "pen up" and "repeat". Next, we asked students to modify their algorithms and generate spiraling shapes in which each line segment is longer (or shorter) than the previous one. This exercise introduced students to the "speed-up" and "slow-down" commands, and it gave them a chance to explore the relationship between speed, acceleration, and distance.

Phase II: Conceptualizing and re-representing a speed-time graph - In this activity, students generated shapes such that the length of segments in the shapes were proportional to the speed in a given speed-time graph. Figure 3 depicts the speed-time graph provided to all students, along with a sample student output where the initial spurt of acceleration is represented by a small growing triangular spiral, the gradual deceleration by a large shrinking square spiral, and constant speed by a triangle. The focus was on developing mathematical measures from meaningful estimation and mechanistic interpretations of the graph, and thereby gaining a deeper understanding of concepts like speed and acceleration.

Phase III: Modeling motion of an agent to match behavior of an expert model - For this activity, students modeled the behavior of a roller coaster as it moved on different segments of a track: up (pulled by a motor), down, flat , and then up again. Students were first shown the simulation results produced by an 'expert' roller coaster model in the V world. Then, they were asked to conceptualize and build their own agent model to match the observed expert roller coaster behavior for all of the segments.

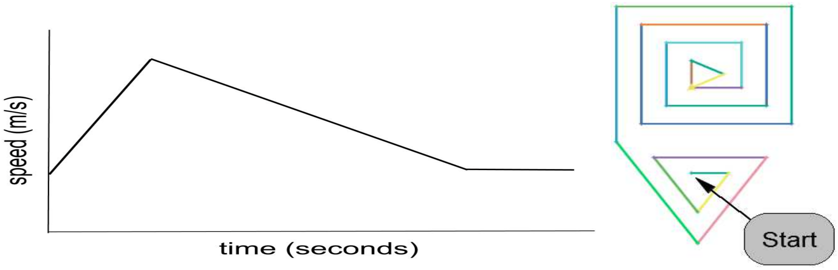


Figure 3: Acceleration represented in a speed-time graph and sample student output

2.2 Ecology Unit

For the Ecology unit students modeled a closed fish tank system in two steps: (1) a macro-level semi-stable model of the behavior of fish and duckweed; and (2) a micro-level model of the waste cycle with bacteria. The macro model included several key relations: (1) the food chain and respiration cycles of the fish and duckweed, (2) the macro-level elements of the waste cycle (fish produce waste, duckweed consume nitrates), and (3) the reproduction of duckweed.

The non-sustainability of the macro-level model (the fish and the duckweed gradually died off), which students built first, helped them reflect, which, in turn provided the transition to the micro model. When prompted to think about why the system was not self-sustaining, students could identify the continuously increasing fish waste as the culprit, and this provided the trigger to introduce the role of bacteria in the system.

At the micro level, students modeled the waste cycle with bacteria and the related chemical processes that converted the toxic ammonia in the fish waste to nitrites, and then nitrates, which sustained the duckweed. The graphs generated from the expert simulation helped students understand the producer-consumer relations: (1) *Nitrosomonas* bacteria consume ammonia and produce nitrites; (2) nitrites are consumed by *Nitrobacter* bacteria to produce nitrates which provide food for the duckweed.

2.3 Sequencing of units and activities

Within each unit, the learning activities were designed to introduce students to (1) *the agent-based program structure*, *i.e.*, agents and their actions, and (2) *the general computational constructs, i.e.*, conditionals to model situation-based interactions, loops to capture repeated agent behavior, and mathematical operations. The units provided a natural sequencing in which students first learned to model and reason with a single agent in kinematics and then went on to model multiple agents and their interactions in ecology. The focus in the kinematics unit was on modeling real-world phenomena by generating computational abstractions. In the ecology unit, students had the more complex task of modeling multiple agent types (e.g., fish and duckweed at the macro level) and composing several procedures to define the behavior associated with each agent type (e.g., swim, eat, and breathe procedures for fish agents).

3. Method

The study was conducted with 6th-grade students from an ethnically diverse middle school in middle Tennessee. 15 students worked on the system outside the classroom with one-on-one guidance from members of our research team (Scaffolded or S-Group), while the remaining 9 students in the class worked on the system in the classroom (Classroom or C-Group) with some instruction from the researchers and the classroom teacher. The C group also received individual help from the researchers if they raised their hand and asked for help. The students were assigned to the groups by their classroom teacher. During the intervention, five interviewers worked one-on-one with the S-Group students and provided verbal scaffolds. In the C-Group, students received minimal one-on-one scaffolding. We formulated two research hypotheses:
1. The intervention will help both groups improve their understanding of science concepts as demonstrated by their pre-to-post-test learning gains.
2. The one-on-one scaffolding will help the S-Group learn more than the C-group.
As part of the design-based research, we collected and characterized the scaffolds provided by the researchers to inform future system development.

All students worked on the three phases of the kinematics units before the ecology macro and micro units. After completing the ecology micro model, the S group received an additional scaffold: they discussed the combined micro-macro model with their assigned researcher and were shown how the two models were causally linked to support sustainability. Students were given the paper-and-pencil task of building a causal model of the cycles, and then prompted to use this representation to explain the effects of removing one agent on the stability of the cycle.

Students worked on the two science units in hour long sessions for three days each. On day 1 of the study, we administered pre-tests for both units. Students worked on the kinematics unit from day 2 to 4, and then took the kinematics post-test on day 5. This was followed by work on the ecology unit from day 6 to 8, and the ecology post-test on day 9.

4. Results and Discussion

4.1 Learning Gains in Kinematics and Ecology

The Kinematics pre/post-test assessed students' abilities to reason causally about mathematical representations of motion and determined whether agent-based modeling improved their abilities to generate and explain these representations. Specifically, the questions on the test required interpretation of speed versus time graphs and generating diagrammatic representations to explain motion in a constant acceleration field. For the Ecology unit, the pre- and post-tests focused on students' understanding of roles of species in the ecosystem, interdependence among the species, the waste and respiration cycles, and how a specific change in one species affected the others. Some of the questions checked students' declarative knowledge about the fish tank system (Declarative Knowledge Check or DKC), others required performing Causal Reasoning about entities using the Declarative Knowledge (CRDK), and a Transfer Question (TQ) required students to reason about the carbon cycle. An example DKC question asked was *"For each of the following species in the fish tank,: a) Goldfish, b) Duckweed, c) Nitrosomonas, d) Nitrobacter, mention the chemical(s) it directly needs to stay alive"*, while an example CRDK question asked was "*Your fish tank is currently healthy and in a stable state. Now, you decide to remove all traces of nitrobacter bacteria from your fish tank. Would this affect a) Duckweed, b) Goldfish, c) Nitrosomonas bacteria? Explain your answers."*

Students in both groups had higher percentage scores on the kinematics pre-test than the ecology pre-test, indicating that they had a better initial understanding of the kinematics domain. However, pre-test scores and the mean TCAP (Tennessee Comprehensive Assessment Program) science scores suggested differences in prior knowledge and abilities of the S and C groups [significant differences (t=3.15, $p<0.005$) in mean TCAP science scores between the two groups]. Hence we computed a repeated measures ANCOVA with TCAP science scores as a covariate to study the interaction between time and condition. There was a significant effect of condition on pre-post learning gains in ecology ($F(1,21)=37.012$, $p<0.001$), and a similar trend was seen in kinematics ($F(1,21)=4.101$, $p<0.06$). The plots in Figure 4 show that the S group's adjusted gains were higher than the C group in both units.

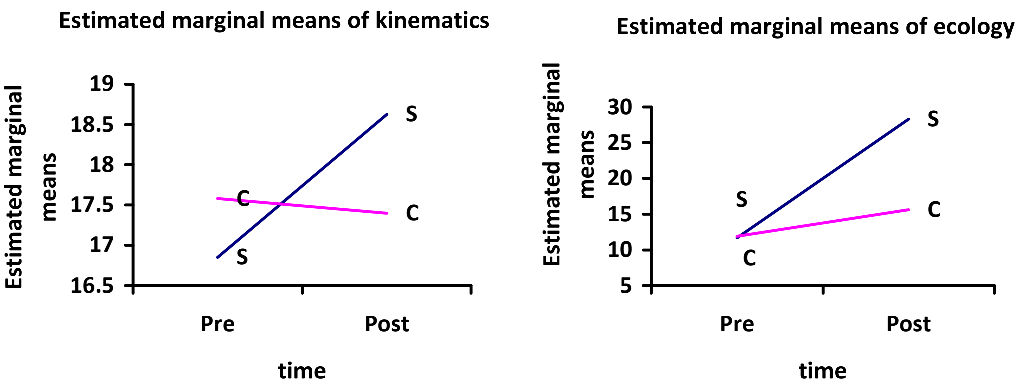


Figure 4: Comparison of gains between groups using TCAP scores as a covariate

Given the significant interaction between time and condition, we performed paired t-tests on pre-to-post gains for each condition. Table 1 shows that the intervention produced statistically significant gains for ecology unit, but not for the kinematics unit. However, for both units, the S group, which received direct one-on-one scaffolding, showed higher learning gains than the C group.

Table 1: Paired t-test results for Kinematics and Ecology pre and post test scores

	Kinematics				Ecology			
	PRE (S.D.) (max=24)	POST (S.D.) (max=24)	t-value	P-value (2-tailed)	PRE (S.D.) (max=35.5)	POST (S.D.) (max=35.5)	t-value	P-value (2-tailed)
S-Group (n=15)	18.07 (2.05)	19.6 (2.29)	2.699	0.017	13.03(5.35)	29.4(4.99)	8.664	<0.001
C-Group (n=9)	15.56 (4.1)	15.78 (4.41)	0.512	0.622	9.61(3.14)	13.78(4.37)	3.402	<0.01

The lack of statistical significance in the kinematics unit may be attributed to a ceiling effect in the students' scores. One exception was a question that asked students to diagrammatically represent the time trajectory of a ball dropped from the same height on the earth and the moon. The students were asked to explain their drawings and generate graphs of speed versus time for the two scenarios. The S group showed significant gains ($p<0.0001$) on this question, while the C group showed an increasing trend, although it was not significant ($p=0.16$).

For the ecology unit, the S-Group students gained on all categories of questions, though all of the gains were not statistically significant. Table 2 reports normalized learning gains (gain/maximum possible gain) by question category for both the groups. Significant gains were observed on the DKC and CRDK questions, which can be attributed to an increased awareness of the entities in the fish tank and their relations with other species. For example, pre-test results indicated that the students did not initially know about the bacteria and their roles. Though students in both groups were told about the role of bacteria during the intervention, the supplementary causal-reasoning activity helped the S-group students gain a better understanding of the interdependence among the species. The S group's gains on the TQ were not significant due to a ceiling effect (most students had strong prior knowledge about the carbon cycle). On the contrary, the C-Group gained only on the CRDK questions, though less than the S-Group ($F(1,21)=21.06$, $p<0.001$). This can be explained by the C group's minimal scaffolding and, especially, the absence of scaffolds targeted towards causal reasoning.

Table 2: Normalized learning gains on categories of Ecology questions

S-Group normalized gains (P-value)			C-Group normalized gains (P-value)		
DKC	CRDK	TQ	DKC	CRDK	TQ
.865 (<0.0001)	.725 (<0.0001)	0.495 (.11)	0 (NA)	.192 (<0.01)	0 (NA)

4.2 Types of scaffolds provided and their effectiveness

A preliminary post hoc analysis of the recorded interviews was used to categorize the different scaffolds provided to the S group, as summarized in Table 3. The SS helped students become familiar with the different modeling primitives, the interface elements, and the modeling task. The MS helped with specifying correct parameter types for the modeling blocks, prompting reflection about how the commands in the model corresponded to the observed simulation behavior, and describing how different functionalities of an agent needed to be separated into different procedures. The TS were specific to the unit the students were working on, and helped clarify the task for the student. The DS suggested that students slow down the simulations to study differences between their model's behavior and that of the expert model using the simulations and plots as guides. The students had to explain possible reasons for the observed differences, and then identify procedures and primitives to modify in order to match expert behavior. The CRS prompted students to reason in causal chains about different entities of the system to help understand global system behavior. For example, as mentioned in Section 3, the S group investigated a combined model of the fish tank and then re-represented the fish-duckweed cycle on paper to reason about the roles of different species in the fish tank.

Table 3: Categories of scaffolds provided to the S-Group of students

Scaffold Category	**Description**
System-use Scaffolds (SS)	Familiarizes students with the UI and use of modeling primitives
Model-building Scaffolds (MS)	Helps students correctly parameterize modeling primitives, modularize code, and correlate models with the resultant simulations
Task-based Scaffolds (TS)	Helps clarify the particular activity being worked on in a unit
Debugging Scaffolds (DS)	Helps students identify differences between their model-generated simulations and expert simulations, elicits explanations for the reason(s) behind the differences and the methods for rectifying them
Causal Reasoning Scaffolds (CRS)	Encourages reasoning about system entities in causal chains to understand global system behavior

Although the results in Section 4.1 illustrate the overall effectiveness of the scaffolding, they do not identify the utility of particular scaffolds. However, some scaffolds likely had a larger impact than others. For example, after adjusting for variations in TCAP scores, there was a significant difference in performance between the S and C groups in the

Ecology Unit on the CRDK questions (F(1,21)=21.06, $p<0.001$), which may be a result of the CRS provided only to the S-Group. Reasoning causally through multiple agent interactions in the fish tank provided a global view of the ecosystem dynamics [1], which was essential for understanding concepts of balance and interdependence covered by the CRDK questions.

5. Conclusion

We have presented a learning environment which integrates computational thinking, visual programming, and agent-based modeling and simulations to help middle school students learn science across multiple domains. Our results indicate that the learning environment helped produce significant learning gains, as measured by pre- and post-test scores, for both the Kinematics and Ecology units. We also demonstrated the necessity of scaffolding and some types of scaffolds required in such an environment. As next steps, we will integrate such scaffolds into the CTSiM environment by building scaffolding tools and providing feedback via a virtual mentor agent. Also, now that we have developed common computational constructs for modeling in different domains, one of our design goals is to help students realize and exploit these commonalities.

References

[1] Basu, S., & Biswas, G. (2011). Multiple Representations to Support Learning of Complex Ecological Processes in Simulation Environments. In *Proceedings of the 19th International Conference on Computers in Education. Chiang Mai, Thailand.*

[2] Blikstein P. & Wilensky, U. (2009). An Atom is Known by the Company it Keeps: A Constructionist Learning Environment for Materials Science Using Agent-Based Modeling. *Int Jounal of Computers for Math Learning, 14*:81–119.

[3] diSessa, A., Hammer, D., Sherin, B., and Kolpakowski, T. (1991). Inventing Graphing: Meta Representational Expertise in Children. *Journal of Mathematical Behavior*, 10, 117-160.

[4] Guzdial, M. (1994). Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments*, 4(1), 1-44.

[5] Hambrusch, S., Hoffmann, C., Korb, J.T., Haugan, M., and Hosking, A.L. (2009). A multidisciplinary approach towards computational thinking for science majors. In *Proceedings of the 40th ACM technical symposium on Computer science education . ACM, New York, NY, USA, 183-187*

[6] Harel, I. & Papert, S. (1991). Software design as a learning environment. *Constructionism*, 51–52.

[7] Krajcik, J., McNeill, K. L. & Reiser, B. (2008). Learning-goals-driven design model: Curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, 92(1), 1-32.

[8] Kynigos, C. (2007). Using half-baked microworlds to challenge teacher educators' knowing, *Journal of Computers for Math Learning*, *12*(2), 87-111.

[9] Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development*, 23 (4), 512-529.

[10] NRC (2010). Report of a Workshop on The Scope and Nature of Computational Thinking.

[11] Sengupta, P., Kinnebrew, J., Biswas, G., and Clark, D. (2012). Integrating Computational Thinking with K-12 Science Education: A Theoretical Framework. *Proceedings of the 4th International Conference on Computer Supported Education*

[12] Sengupta, P., & Farris, A.V. (2012). Learning Kinematics in Elementary Grades Using Agent-based Computational Modeling: A Visual Programming Based Approach. In *Proceedings of the 11th International Conference on Interaction Design & Children*, pp 78 – 87.

[13] Sengupta, P., Farris, A.V., & Wright, M. (2012). From Agents to Continuous Change via Aesthetics: Learning Mechanics with Visual Agent-based Computational Modeling. *Technology, Knowledge & Learning.* Vol. 17 (1 - 2), pp 23 - 42.

[14] Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

[15] Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society*, vol. 366, pp.

Student Learning Behavior in an Unsupervised Learning Environment

Paul Salvador INVENTADO[ab*], Roberto LEGASPI, Rafael CABREDO[ab] & Masayuki NUMAO
[a]*The Institute of Scientific and Industrial Research, Osaka University, Japan*
[b]*Center for Empathic Human-Computer Interactions, De La Salle University, Philippines*
*inventado@ai.sanken.osaka-u.ac.jp

Abstract: Learning is commonly associated with knowledge transfer involving guidance from a teacher. However, as people grow older they are expected to know how to learn by themselves. In this research, we analyzed student learning in an unsupervised learning environment, i.e., performing academic research, wherein students have complete control over their learning thus requiring them to manage it. Transition likelihood metrics were used to analyze the interplay between emotion, learning and non-learning related activities while students did research. Several observations were seen from students learning in this environment such as students experiencing cognitive disequilibrium but experiencing disengagement faster. Non-learning related activities were also shown to have the potential of motivating students to resume learning. Lastly, user-specific traits and context seem to affect the interplay between learning and non-learning activities in an unsupervised learning environment. This highlights the need to not only create general models to predict student behavior but also user-specific models to allow future systems to provide appropriate feedback in this environment.

Keywords: unsupervised learning environment, cognitive and affective state transition, likelihood metric

Introduction

Learning is an essential part of human life. It starts from the very beginning of our lives when we start learning simple actions and then eventually moving on to very complex processes. We usually learn through knowledge transfer from a teacher whose role is fulfilled by a parent, a sibling, a peer or a teacher in school. Over time however, we are expected to learn on our own so that we no longer need constant guidance from teachers in acquiring knowledge and applying it to solve the problems we wish to address. In most cases, we become the teachers too and it becomes our turn to share our knowledge to others.

Many computer systems have been created to aid the process of learning so that in cases when a human teacher is unavailable, a computer system can take on that role and provide support for students while learning a particular topic [8]. Although these systems model the students' knowledge of the topic being learned, these do not consider how the student interacts with elements outside of the learning environment and its effects on the learning session.

In this work, we focused on *unsupervised learning environments* which we define as an environment characterized by the absence of supervision from either a human or an automated teacher. As a result, this environment requires students to manage their learning apart from accomplishing their learning goals. Some elements that need to be managed include the goals for the session, the amount of time spent in completing goals, the affective states experienced, the transitions to non-learning related tasks and the avoidance of distractions. Support is important in this kind of environment because students unable to manage learning will most likely perform poorly or fail to complete their requirements.

Our goal for this work was to analyze and model students' learning behavior in an unsupervised learning environment so we can enable future systems to provide appropriate feedback in such an environment.

1. Related Work

Most existing systems designed to support learning provide cognitive and/or affective feedback to students as they solve problems or explore environments generated by the system [1][6]. Tracking both cognitive and affective states have allowed these systems to model students more completely and also provided a better understanding of student learning. This in turn was used as basis for designing and identifying the appropriate feedback for students using these systems for learning.

D'Mello and Graesser [4] developed a model that used both cognitive and affective elements to explain how students' emotions transitioned while learning. This model can be used to predict student behavior and provide appropriate feedback in cases when students engage in activities or experience emotions that are not helping them learn. Students using computer-based learning environments also engage in non-learning related activities and certain cognitive and affective states have been reported to lead to off-task activities like gaming the system and talking to seatmates [2][13]. Although learning environments prevent or discourage students from engaging in non-learning related activities, students learning by themselves can freely do so and are left to deal with them on their own. If they are unable to manage their learning, they may learn less and may need to spend more time to learn [9].

Not much work has been done to investigate the role of non-learning related activities while learning especially when these are not purposely prevented or discouraged. We believe that a better understanding will allow us to design systems that can leverage from its benefits and try to reduce its negative effects.

2. Data Gathering

We gathered data from four students who performed academic research. We considered this an unsupervised learning environment because the students did not receive supervision and they needed to manage their own learning. When students worked on their research, they spent time developing their ideas, performing analysis and making conclusions. They usually consulted or got suggestions and ideas from their supervisor but when they did actual research work, it was their responsibility to identify their goals for the day, which activity to prioritize and how much time to spend on an activity. Research required the students to spend much time working and it was common for them to engage in non-research related activities such as reading an email from their supervisor, exchanging instant messages with a friend and viewing a video aside from others.

The participants consisted of one male undergraduate student, one male master's student and two female doctoral students. All of these students were required to do research as a requirement for their degrees and they commonly did their research with the help of a computer using different applications such as web browsers, word processing software and programming environments. Students were given the freedom to engage in both learning and non-learning related activities either on or our outside of the computer. The annotation software used and the methodology for collecting data are described in the following subsections.

2.1 Sidekick Retrospect Software

The Sidekick Retrospect Software was developed by the authors to help students manage learning and annotate their behavior. It encouraged the use of self-regulation strategies specifically goal setting, self-monitoring and self-reflection [14]. When students began their learning session, they first identified their goals for the session and inputted each one on the interface. While learning, students were allowed to add more goals in case they found the need to. The software also logged the applications used by the students and took screenshots of their desktop and webcam feed throughout the session.

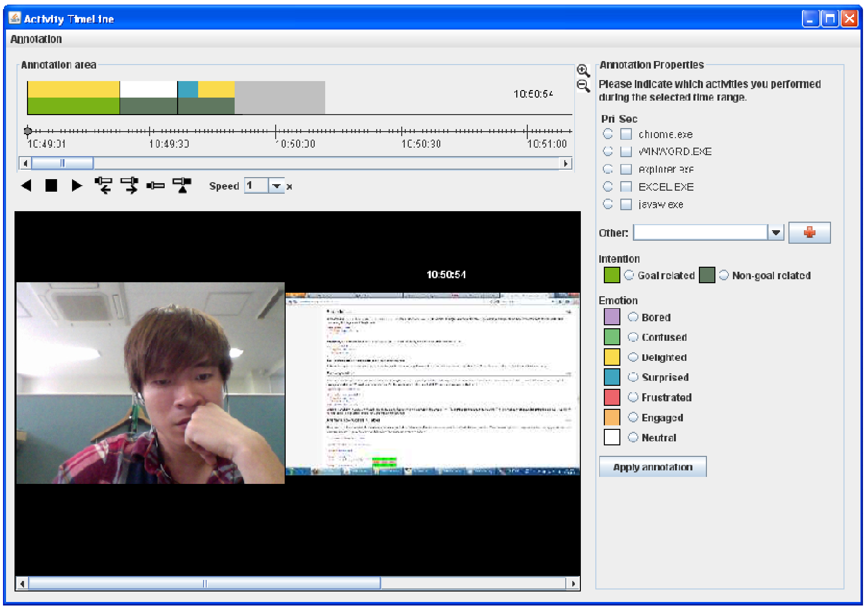


Figure 1. Interface for annotating learning behavior.

When the students ended the session, they were presented with a timeline representing their entire learning session. The corresponding desktop and webcam screenshots were shown whenever the mouse hovered over the timeline to help students recall what activity they did at that moment. Students annotated their behavior by clicking and dragging the mouse to select a time span and then identifying their corresponding intention, activity and emotion label using the interface shown in Figure 1. Intention is the student's purpose for performing an activity which was either goal related or non-goal related. In this work, goal related and non-goal related activities are synonymous to learning and non-learning related activities because the goals defined by the students were for learning. An activity is what students did at a certain point in time such as using the Google Chrome browser to search for information, chatting using Yahoo Messenger, or reading a technical paper. Lastly, emotion is the student's experienced emotion while performing the activity. Two emotion sets were used during annotation. When students performed goal related activities, they selected one emotion from: delight, engagement/flow, confused, frustrated, surprised, afraid and neutral. The list comprised of academic emotions which are commonly observed when students learn [3]. When students performed non-goal related activities, they selected one emotion from Ekman's six basic emotions namely angry, disgusted, sad, delighted, afraid, sad and neutral [7]. A different emotion set was used for learning related activities because academic emotions do not only refer to an affective state but are also indicative of certain cognitive states which are useful for providing feedback. Academic emotions do not have special meaning in non-learning related activities thus Ekman's basic emotions were considered sufficient.

After annotation, students were asked to identify how much of their goals were completed, which activities helped them complete their goals and how productive they felt the learning session was. This helped students self-reflect to identify which activities were helpful to learning or which were not and also for evaluating their performance.

At the end of the entire session, the software generated a log file containing the intention, activity and emotion labels with their corresponding time stamp. Instances in the log file were one second apart.

2.2 Data Gathering Methodology

Before data gathering, each participant was first given a short tutorial on how to use the Sidekick Retrospect software and an explanation of the different annotation elements. They were told that all activities they did which were related to their research goals for that day would be considered goal-related and any other activity would be non-goal related. They were also asked to indicate the different activities they did regardless if it was done on or outside the computer. Lastly, each emotion was explained to them so they would know how to differentiate them. In situations when they experienced more than one emotion, they were asked to identify the most prominent one.

After the tutorial, the software was installed on the participants' computers and they were taught how to start and stop the data gathering module and create annotations in the annotation module. They were then asked to use the software in five separate sessions, wherein each session lasted around two hours. Students had the freedom to choose when, where and which materials to use for learning.

At the end of every session, the students were asked to annotate their learning behavior and then answer a survey. The survey contained questions regarding the students' realizations from their learning behavior. Data was gathered from each student over a period of one week, with five two-hour sessions per day resulting in a total of 40 hours of data from all participants.

3. Analysis

Majority of the students used the software and annotated their learning behavior when they worked at their own table in their respective research laboratories. This was where students commonly did their everyday work wherein other research students were also present in the same room interacting with each other and sometimes with the participant. The students learned in the same setting as they always did and the software did not require them to change their behavior while learning. Although students were asked to annotate their data, this was done after the learning session. We believe that the methodology we used resulted in naturalistic data.

The data showed that students spent about one hour and 30 minutes engaging in learning related activities out of the entire two hour session. The participants set three goals on average throughout the duration of the session. They rarely completed all of the goals they defined but in most cases they completed at least 75% of one of the defined goals. This further indicates that the data gathered from the participants was naturalistic because the students did not avoid non-learning related activities nor favored it. It also shows that they were capable of managing their learning behavior because they were able to identify their goals, perform activities to accomplish their goals, and spend majority of their time in learning related activities. Students experienced engagement, delight, boredom and confusion over prolonged periods of time when learning. On the other hand, students experienced delight and were in the neutral affective state over prolonged periods of time when non-learning.

We further investigated the interplay between learning and non-learning related activities using a transition likelihood metric to identify what caused students to shift between these activity types. D'Mello et al. [5] introduced a metric to measure the likelihood of transitioning from one affective state to another while students used an intelligent tutoring system. However, in our research we did not only consider activities related to learning but also non-learning related activities. We modified the likelihood metric to include the type of activity done by the student which is shown in Eq. 1. The equation measures the likelihood of transitioning from a state pi to state pi+1 where a state consists of the activity intention I, which may either be learning or non-learning and an emotion E, which may either be delight, engagement/flow, confused, frustrated, surprised, afraid and neutral for learning activities and angry, disgusted, sad, delighted, afraid, sad and neutral for non-learning activities. In our analysis, we were only concerned with transitions into a different state since these described instances when certain factors affected the student to move out of the current state. Assuming that a student performs the following transition: State A → State B → State B → State C, the prolonged state is treated as a single state resulting in the following transition: State A → State B → State C. The metric's resulting value ranges from 1 to -∞. When the value is above zero, it indicates a likely transition with increasing likelihood as it approaches 1. A transition likelihood of zero indicates that the transition is equal to chance and values below zero indicate that the transition is less likely to occur compared to the base frequency of the succeeding state.

$$L(p_i, p_{i+1}) = \frac{\Pr(p_{i+1} \mid p_i) - \Pr(p_{i+1})}{(1 - \Pr(p_{i+1}))} \qquad (1)$$

$$\text{where}: p = \langle I, E \rangle$$

All possible transitions can be generated by combining all possible pairs resulting in a 14x14 matrix. To get a view of the students' learning behavior, we averaged each of the student's transition likelihood values and used a t-test to identify statistical significance. Table 1 shows the resulting matrix of likelihood values containing only the transitions that were more likely than chance (i.e., L>0) with their corresponding p values in parentheses. Likelihood values without p values indicate that there was only one instance of that transition over all sessions from all students. Significant transition likelihood values (i.e., p<0.05) are highlighted. Transitions to the same state were omitted since the data used for the likelihood matrix only contained state changes and not prolonged states.

As many researches have already shown, emotions play an important role in learning which were also observed from the data we gathered. Specifically, there were highly significant and likely transitions from engagement to confusion (L=0.15) which is indicative of cognitive disequilibrium. Cognitive disequilibrium occurs when students encounter errors or anomalies in what they are learning requiring them to either accept new information or to alter their current understanding to return to an equilibrium state [12]. Cognitive disequilibrium is commonly attributed to learning and usually co-occurs with feelings of confusion [10][11]. We also observed that students likely transitioned from a bored learning state to a neutral non-learning state (L=0.40). Boredom is a state where students have already disengaged from the learning activity and it is quite understandable for them to simply shift to a non-learning activity most especially because they are not prevented from doing so. A similar behavior was also seen in the shift from a confused learning state to a neutral non-learning state (L=0.20). Although the transition's p value was p=0.054 it can still be considered statistically significant. This may partially explain why students did not experience hopeless confusion. D’Mello and Graesser [4] described hopeless confusion as a state wherein students in a confused state are unable to resolve the problem and eventually become frustrated. In the case of an unsupervised learning environment, it seemed that when students were not able to resolve the cause of confusion, instead of being frustrated they simply disengaged. On one hand, this is disadvantageous as students would more likely disengage instead of solving the problem. On the other hand, this could have led to lesser frustration, resulting in less stress and allowing them to possibly have more motivation to continue learning at a later time.

Table 1. Transition likelihood matrix with corresponding statistical significance

		L						**NL**			
		EN	FR	CO	BO	DE	NE	DI	DE	SU	NE
L	EN			0.15 (0.001)			0.04 (0.390)		0.05 (0.502)	0.05 (0.172)	0.17 (0.037)
	FR										0.74
	CO				0.03 (0.558)		0.03 (0.607)		0.09 (0.303)		0.20 (0.054)
	BO						0.03 (0.691)		0.08 (0.241)		0.40 (0.016)
	DE						0.14 (0.514)		0.50 (0.336)		0.26 (0.610)
	NE	0.41 (0.031)	0.11 (0.351)			0.01 (0.422)					
NL	DI	1.00									
	DE	0.35 (0.013)				0.02 (0.402)	0.13 (0.106)				
	SU			0.43	0.23						0.14
	NE	0.25 (0.024)	0.02 (0.334)	0.06 (0.454)	0.01 (0.895)		0.09 (0.212)	0.02 (0.334)			

Intentions: **L** – Learning; **NL** – Learning; Learning emotions: **EN**gaged, **FR**ustrated, **CO**nfused, **BO**red, **DE**lighted, **NE**utral; Non-learning emotions: **DI**sgusted, **DE**lighted, **SU**rprised, **NE**utral

As mentioned earlier, students were likely to resume learning as shown by the transitions from a delighted and neutral state while non-learning to an engaged learning state (L=0.34 and L=0.26 respectively). This can be interpreted in two ways. It can be that students were capable of fending off distractions in an unsupervised learning environment or, the presence of non-learning related activities reduced stress or helped students maintain a level of motivation which helped them relax more and resume learning at a later time.

Lastly, the likely shift from a neutral learning state to an engaged learning state (L=0.26) also indicated that students were not always engaged. Presumably, they first read and understood content before getting immersed in the activity they were trying to accomplish. Over time, they did become engaged which is the ideal state for learning.

Some students transitioned to and from other states however these were not observed in other students or were only seen in a small number of sessions. This brought up two concerns that needed to be considered when dealing with unsupervised learning environments. First, personality or some other user-specific trait may have an effect on a student’s learning behavior that was observed in one student but not others. A more in depth data analysis showed that only one student transitioned to a neutral learning state before experiencing other emotions while learning. Consider the following transitions observed from the student while engaging in a learning activity: EN→NE→CO; EN→NE→BO; EN→NE→EN; CO→NE→EN; BO→NE→EN. The transition to a neutral state indicates that the student might have lost focus first then eventually experienced confusion or boredom. While in the state of confusion or boredom, the student seemed to get back his learning momentum first before transitioning to an engaged state again. Other students transitioned directly between these emotions without passing through a neutral state. Secondly, context may also play a big part in these transitions. For example, when a student had an upcoming deadline, his transitions to a delighted learning state or even to non-learning states were not as frequent as the time when he did not have a deadline.

Both of these concerns indicate that although we were able to get a good idea of how non-learning states interplay with learning states in general, the differences in both context and user-specific traits require a more user-specific and context-dependent analysis. It is also important to observe learning behavior over a longer period of time to gather data on more scenarios that students may encounter.

When the students made annotations using the software, they were also able to reflect about their own learning behavior. The students' answers to the survey reflected the learning behaviors that were uncovered from the data. For example, one student said "I got distracted a lot but I needed it as it helps me get back into focus" showing the value of engaging in non-learning related activities. Another student said "I felt bored before transitioning from a goal related to a non-goal related activity" which explained the high likelihood of transitioning from a bored learning state to performing a non-learning related activity.

4. Conclusion and Future Work

In this research we defined an unsupervised learning environment as a learning environment wherein students do not receive guidance from a human teacher or an automated system and wherein they freely control their activities. Students learning in this environment need to be capable of managing their goals and controlling the attention they give to non-learning related activities.

An analysis of students' learning behavior was conducted which showed the occurrence of cognitive disequilibrium. This indicated that the environment was challenging enough for the students to engage in learning. The accessibility of non-learning related activities also seemed to cause students to disengage faster but also minimized their frustration or stress. Non-learning activities also seemed to help students maintain or regain motivation to continue learning.

The data showed that although students were capable of learning in this environment as well as fend off distractions from non-learning related activities, there were still instances when they needed help in managing them. Students commented that they spent too much time in non-learning related activities and easily gave in to distractions, highlighting the importance of support in such an environment.

The data gathered from the students showed transitions shared among all students and also transitions which were unique only to some students. This indicated that user-specific traits and context may have affected how students transitioned between learning and non-learning activities. This requires not only a general model but also a user-specific model to predict their learning behavior. A general model can be used as a basis for designing feedback which is applicable to all students, while a user-specific model can be used to adapt feedback over time and to provide more appropriate feedback based on their traits and the contexts which they commonly learn in. The creation of such models can lead to the development of systems that can provide support for students learning in unsupervised learning environments. Such a system will also help students become more equipped for learning on their own in the future.

Acknowledgements

This work was supported in part by the Management Expenses Grants for National Universities Corporations from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) and JSPS KAKENHI Grant Number 23300059. We would also like to thank all the students who participated in our data collection.

References

[1] Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K., and Christopherson, R. (2009). Emotion sensors go to school. In *Proceedings of the 2009 Conference on Artificial Intelligence in Education*, pages 17-24, Amsterdam, The Netherlands. IOS Press.

[2] Baker, R. S., Rodrigo, M. M., and Xolocotzin, U. E. (2007). The dynamics of affective transitions in simulation Problem-Solving environments. In *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*, ACII '07, pages 666-677, Berlin, Heidelberg. Springer-Verlag.

[3] Craig, S. D., Graesser, A. C., Sullins, J., and 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.

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

[5] D'Mello, S., Taylor, R. S., and Graesser, A. (2007). Monitoring affective trajectories during complex learning. In *Proceedings of the 29th annual meeting of the cognitive science society*, pages 203-208, Austin, TX, USA. Cognitive Science Society.

[6] D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., Kaliouby, E. R., Picard, R. W., & Graesser, A. (2008). Autotutor detects and responds to learners affective and cognitive States. In *Proceedings of the Workshop on Emotional and Cognitive Issues at the International Conference on Intelligent Tutoring Systems*. Montreal, Canada.

[7] Ekman, P. (1992). Are there basic emotions*? Psychological Review*, 99(3):550-553.

[8] Graesser, A. C., Conley, M. W., and Olney, A. M. (2012). *Intelligent Tutoring Systems*, volume 3 of *Applications to Learning and Teaching*, pages 451-473. American Psychological Association, Washington, DC.

[9] Kirschner, P. A. and Karpinski, A. C. (2010). Facebook® and academic performance. *Computers in Human Behavior*, 26(6):1237-1245.

[10] Lehman, B., D'Mello, S., and Graesser, A. (2012). Confusion and complex learning during interactions with computer learning environments. *The Internet and Higher Education*, 15(3):184-194.

[11] Lehman, B., D'Mello, S., and Person, N. (2010). The intricate dance between cognition and emotion during expert tutoring. In Aleven, V., Kay, J., and Mostow, J., editors, Intelligent Tutoring Systems, volume 6095 of *Lecture Notes in Computer Science*, pages 1-10. Springer-Verlag, Berlin, Heidelberg.

[12] Piaget, J. (1952). *The Origins of Intelligence*. International University Press, New York.

[13] Rodrigo, M. M., Baker, R. S. J. D., Lagud, M. C. V., Lim, S. A. L., Macapanpan, A. F., Pascua, S. A. M. S., Santillano, J. Q., Sevilla, L. R. S., Sugay, J. O., Tep, S., and Viehland, N. J. B. (2007). Affect and usage choices in simulation Problem-Solving environments. In *Proceedings of the 2007 conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, pages 145-152, Amsterdam, The Netherlands. IOS Press.

[14] Zimmerman, B. J. (2002). Becoming a Self-Regulated learner: An overview. *Theory Into Practice*, 41(2):64-70.

A New Course Design for Advanced Communication: "Debate and Beyond…"

A Technology Enhanced Course for Communication Incorporating Empathy

Tosh YAMAMOTO[a*], Masanori TAGAMI[b] , Chiaki IWASAKI[c] & Minoru NAKAZAWA[d]

[a~c]*The Center for Teaching and Learning, Kansai University, Japan*

[d]*Dept. of Information Engineering, Kanazawa Inst. of Tech., Japan*

*ctltosh@kansai-u.ac.jp

Abstract: Based on the fundamental concept of the communication being mainly composed of logic in information, the university in Japan has just begun offering such courses as logical thinking, critical thinking, and debate as subjects for the general education to freshmen and sophomores before they begin their special areas of study. However, after the natural disasters since the last year, the society has realized that the basis of communication may be supported not mainly by logic or critical thinking but more or less by some empathy driven factor. Because the university has the mission to raise promising potential members for the future society, a new concept for a course has come out in which the successful communication consists of both logical and critical thinking on the one hand, and emotion or empathy on the other. This paper attempts to propose a new course to such issue enhanced by the information technology available to us.

Keywords: communication, empathy, negotiation, ICT, SNS

Introduction

Since the Benesse Corporation reported in its journal "VIEW" in 2011, it has been revealed that most universities do not satisfy the needs from the society in their curricula or diploma policies, [1]. For example, after three years after graduation, most university graduates feel the need for such fundamental skills as problem solving skills (data collection, analysis, problem solving), continuous learning (intellectual curiosity and active learning), independence (exercising independence and displaying leadership in projects), teamwork (collaborative attitudes & perform one's duties and responsibility in projects), self-management (setting goals, planning well, and working accordingly), problem setting and solving (analyzing the situation to identify the problem to be solved), and logical thinking. It follows that the universities in Japan do not provide the education necessary for the graduated students. In this way, such fundamental skills have been beginning to be incorporated in the freshman and sophomore curricula. The fundamental academic skills are generally considered as the skills necessary to do well in an educational setting as well as in a social situation after graduation. Thus, most universities thought that such incorporation of the fundamental skills into the curriculum was enough to meet the required needs by the graduates.

However, after the natural disasters happened on March 11 the last year, the society has realized, in the process of recovering while demonstrating resilience, that the basis of human communication to build the strong connections between people may be supported not mainly by logic or critical thinking but more or less by some empathy driven factor. This means that although most universities have been offering courses such as logical thinking, critical thinking, and debate, students graduated without having good communication skills to do well in the society. As a matter of fact, in his book titled “Become an Effective Leader,” Dale Carnegie emphasizes that most part of communication to reach the heart of others is composed of affection to them instead of the logic or rhetoric of the language that people use, [2].

Is there any way to combine the traditional communication course of teaching logical and critical thinking with empathy to build a long-term good relationship with others? It is proposed that a new course may be designed with the help of a new approach to communication as well as the cutting-edge technology.

1. A New Communication Course Incorporating Empathy

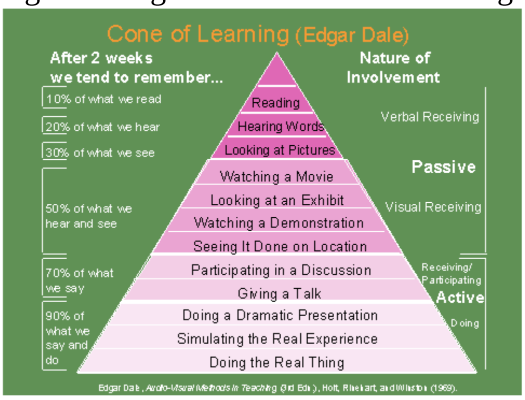
1.1 Background and the Problem Defined

Kansai University has in the curriculum the basic study skills courses including problem-identifying, problem-solving, note-taking, report writing, presentation, computer literary, and debate. Debate is placed as one of the most advanced courses in the basic study skills courses, which incorporates the logical thinking and meta-cognition skills. However, the debate course does not offer anything to build trust and good relation between students. Further, it has been found that emotion or empathy was not in any part of the communication in the debate. Rather, it is a virtue to hide emotion while debating.

1.2 Kansai University’s Challenge to Design a New Course

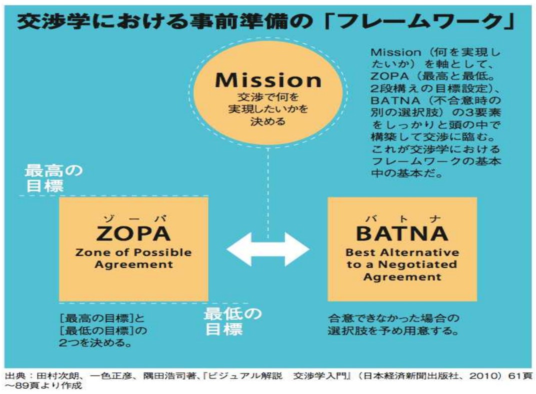
In order to remedy the lack of empathy or emotion in the course, it is proposed that a new course for communication incorporating empathy be introduced in the curriculum to enhance the series of courses for the basic academic skills. In the course of such incorporation, trust building through communication is fostered through negotiation seminar and exercises in groups. In such a course, the simulation experience is crucially employed. The rationale behind this is that the simulation of the real experience will have extremely high learning outcome close to the real life experience, as Edger Dale claims, [3]. See Figure I.

Figure I. Edger Dale’s Cone of Learning



From: www.cals.ncsu.edu/agexed/sae/ppt1/sld012.htm

Figure II. The Framework for Negotiation



From: J. Tamura, et. al. (2010). “Visual Explanation: An Introduction to Negotiation”, Nihon Keizai Shinbun.

1.3 Incorporating Trust Building into the Basic Communication Courses with Empathy

The main goal of the methodology for communication proposed by Roger Fisher at Harvard University is generally called “negotiation to lead to say yes.” The negotiation requires prior preparation to set a mission, which is the optimal goal to be aimed at, n this case, for maintaining a long-term good relationship, [5]. In order to accomplish such mission, some preparation is needed. Since it is not usually possible to achieve the 100% of the mission through negotiation, a zone of possible agreement is set to limit the range between the maximum and the minimum goals prior to the actual negotiation. This zone of goals is called zopa, [6]. Furthermore, if all the options for possible agreements prepared for the zopa fail, the best alternative to the zopa, called batna, is further planned, as shown in Figure II, [8], [9].

The negotiation skill described above demonstrates the skill for gathering information for a certain topic and then making a decision based on the information. By understanding the common framework of negotiation consisting of the mission, zopa, and batna, and by conforming to them, it is possible to demonstrate the logical thinking skill, the critical thinking skill, as well as the skill to conform to the rules and conditions, [8]. Further, by exercising the framework of negotiation, students can develop their communication skills to elicit information necessary for the negotiation as well as the arrangement skill for conducting the negotiation, [9]. This can be only possible with the mission aiming at a long-term good relationship.

1.4 New Course Design

Having the basic concepts of negotiation, the negotiation methodology has been developed incorporating a case study and a role-play simulation, [10]. In order to generate the maximum learning outcome for the designed negotiation course, it requires at least three hours in a session. However, the traditional time frame or contact hour for the course is 90-minute long, meeting only once a week. Thus, it requires two weeks of class contact hours in order to have a session of three hours. The problem is how to maintain students’ motivation and attention to be stretched to the class on the second week. In other words, the retention of the learning experience from the within group discussion will be lost on the second week if nothing is done by the students. Thus, by the time when the class resumes on the second week, the students would need to review what they learned on the previous week. In other words, much time would end up being wasted to refresh what the students learned on the first week.

Table I. One Session Ranging Over Two Weeks

[first week]

0~20 min.	Explanation of the Negotiation Exercise The class is split into two groups.
20~50 min.	Reading the case synopsis common to both groups and synopses for their own Groups
50~90 min.	Within group discussion: prep session Setting the mission, the zopa, and the batna.

[second week]

0~30 min.	One-to-one peer negotiation session
30~60 min.	Reflection session in peers
60~90 min.	Feedback & Conclusion

Figure III. Discussion Activities on SNS

One may think that the students could meet face-to-face in groups after the class or later in the week to continue the discussion for the second week. However, taking into an account that the students from the thirteen colleges are taking the course as an elective for the general study in addition to their required courses of 15 to 24 credit hours per semester, it is impossible for the students to get together face-to-face except for the class contact hours.

2. Proposed Solution to the Problem

After the within group discussion on the first week, the students are asked to post the minimum of five times to post their own strategic plans to the SNS designed for the class and to give comments to other members' posts. The two within groups have their own communities and thus they do not have cross-membership to both within groups. In this way, the SNS can be accessible for 24/7 for the students, [11], [12]. Refer to Figure III.

3. Conclusion

A new course incorporating empathy or emotion to the traditional communication course contents was developed and elaborated in this paper. In the course of development, Dr. Roger Fisher's methodology for negotiation was employed to satisfy the needs. However, in order to achieve a high learning outcome, at least three hours are required to one session for negotiation, which conflict with the traditional class schedule (a 90 minute-session per week times 15 weeks). Thus, one negotiation session (i.e., three hours) must be extending for two weeks. This means that the students' retention to the course contents must be maintained till the following week so that the first half of the three-hour session be smoothly fed into the second half. It was proposed that the SNS be employed as out-of-class activities between the segments of the negotiation session. With the help of information technology, the psychological interruption has been overcome.

References

[1] The Benesse Corporation. (2011). *The View. Vol 21.* Tokyo, 2011.

[2] Carnegie, Dale. (2012). *Become an Effective Leader,* New York: Dale Carnegie & Associates, Inc. In *Become an Effective Leader*, Tokyo: Sogensha.

[3] Dale, Edgar. (1969). *Audio-Visual Methods in Teaching, 3rd ed.,* (p. 108). New York: Holt, Rinehart & Winston,

[4] Bruner, Jerome S. (1966). *Toward a Theory of Instruction*. (p.49). Cambridge, MA: Harvard University Press.

[5] R. Fisher, L. U. William, & I. B. Patton. (1991). *Getting to Yes: Negotiating Agreement Without Giving In*. New York: Penguin Books.

[6] R. Fisher & S. Brown. (1989). *Getting Together: Building Relationships As We Negotiate*. New York: Penguin Books. 1989.

[7] J. Tamura. (2004). *Negotiation Strategies, An Introduction to Negotiation Based on Critical Thinking.* Tokyo: Diamond Press.

[8] M. Isshiki, J. Tamura, & K. Sumida. (2010). *An Introduction to Negotiation*. Tokyo: Nihon Keizan Shinbun Shuppansha.

[9] Diamond, Robert M. (1989). *Designing and Improving Courses and Curricula in Higher Education.* San Francisc: Jossey-Bass.

[10] B. E. Drushel & K. German. (2011). *The Ethics of Emerging Media: Information, Social Norms, and New Media Technology*. New York: The Continuum International Publishing Group.

[11] C. Fuchs. (2011). *Foundations of Critical Media and Information Studies*. New York: Routledge.

Virtual Learning Environments in Primary Schools–Using Learning Theory to Develop an Interactive Virtual Medium

Eiman Tamah Al-Shammari
Kuwait University, Kuwait
Eiman.tamah@gmail.com

Abstract: Virtual Learning Environments (VLEs) bring traditional educational strategies into the 21st century by using technology as a deployment method. Online environments can be utilized alone or in conjunction with face-to-face instruction to develop a robust educational paradigm that is tailored to the students' needs. This paper intends to describe how learning theories recently applied in traditional classrooms can be transferred to an interactive virtual medium to achieve this paradigm. In addition, this paper also demonstrates how the practical model of Piaget's Constructivist Theory of Learning and Vygotsky's Zone of Proximal Development, and their related strategies, could be applied to achieve a set of specific and effective e-learning objectives. This paper will also discuss identify how these strategies can be used to develop a VLE appropriate for primary schools.

Keywords: E-learning, Piaget's constructivist theory, Virtual Learning Environments, Zone of Proximal Development.

Introduction

With the available e-learning tools in hand and their perceived deficiencies on record, we will attempt to design a new primary school VLE e-learning model that takes into account the principles of constructivist theory in conjunction with Vygotsky's theory [7] of the Zone of Proximal Development. We will also try to assess and evaluate how we can create effective VLEs to integrate seamless pedagogical values into the e-learning system.
Effectively speaking, the launching of VLEs in both primary and secondary schools could be extremely effective and have tremendous positive results [2], in its monumental report in the effectiveness of VLEs, suggests that:

"To set up an efficient VLE to implement school curriculum, the school management may need to think of its effectiveness that in turn is influenced by three important factors: knowledge management, pupils' approach to learning, and academic performance (p32)."

In other words, any VLEs used in a primary school should satisfy the above-mentioned pedagogical parameters in a way that eventually leads to success of the e-learning initiative. An effective VLE design should provide seamless knowledge and skill management. Before designing a prototype model, we need to look at various factors that influence the e-learning outcome. The VLE prototype should meet the following criteria:

1. It should consider various issues that relate to learners' mental development stages, cognitive development and user preferences.
2. It should be flexible, easy to use and have an effective interface.
3. It should address the concerns of instructors cited previously in this paper. Using effective VLEs to integrate seamless pedagogical values

1. Pedagogical Framework to the Proposed Model

Swiss developmental psychologist Jean Piaget dedicated his life to studying child development and is credited with the development of the constructivist theory of knowing. His theory acknowledges that people bring previous knowledge to each knew encounter, and as they take in new information it is assimilated with that previous knowledge in a continuous process to gain a higher level of understanding.

Constructivism (also commonly known as social constructivism) recognizes that each student brings a unique and diverse perspective to the learning environment. This is especially beneficial in a VLE since there are no physical boundaries as to where the learners may be; instead of a traditional classroom where the students may all share a culture or socioeconomic background, virtual learners might be scattered across the globe.

Piaget's constructivist theory links up well with another prominent learning theory, the Zone of Proximal Development (ZPD), which was developed by Lev Vygotsky in the early 20th century. ZPD promotes the idea that primary-age children learn first and best through social interactions, with the teacher serving as observer. The ZPD theory is based on the concept that learners are motivated to reach slightly outside their comfort zone to further develop their knowledge and skills. The concepts of constructivism and ZPD form the pedagogical framework to our proposed VLE model for primary school students.

2. Description and salient features of the new e-learning model

Incorporating both constructivism and ZPD into our e-learning prototype for primary school students, we have established the following methodology for learning interactions:

1. Socialization
2. Individual Pre-assessment
3. Active Learning
4. Individual Post-assessment

This methodology is intended to be circular, with the experience of knowledge transfer a cyclical process. In other words, the learning experience does not end. It is built upon by repeating the process to continually increase knowledge. Let's take a look at each step of the process and how it can be applied to the primary school VLE.

Both Piaget and Vygotsky understood the critical role that socialization plays in the learning process, and that is why this prototype includes socialization as its initial step. Using a simple chat-like interface that will likely be familiar to most primary school students, each learner should be encouraged to share information: a personal detail, such as their favorite color; an interesting experience, such as a recent vacation; or any other information that the student wants to share with his or her peers. Not only does this provide the basis for peer-to-peer relationships, it also provides the foundation for learning in a constructivist methodology. Students take in information from their peers and assimilate it with their own knowledge and experiences. For example, when Sarah shares that her family took a vacation to Disney World in Florida and also visited the ocean, the other learners may assimilate that Florida is near an ocean.

At this point, the VLE shifts into a pre-assessment, designed to evaluate the existing level of development on the topic at hand. For example, if the lesson is covering colors, the VLE may provide a technology-based quiz that assesses the student's knowledge of various colors. Keeping in mind that some primary students may not yet be able to read, the VLE should incorporate audio as well as visual effects to meet each learner's needs. The pre-assessment provides a baseline for the instructor before moving into the active learning phase of the methodology. The baseline identifies each student's existing zone of knowledge, which is necessary to understand the ZPD--that area that is just slightly outside the existing zone. If a student is pushed to make too large a leap between the existing zone of knowledge and the next zone, they are likely to back away from the learning process altogether. The pre-assessment encourages the effective building of skills.

With the information gained in the pre-assessment, teachers can adapt an active learning environment to each student's needs. Continuing with the color example, if a teacher determines that most of the students can identify primary colors but have difficulty identifying secondary and tertiary colors, the teacher can tailor the learning experience to meet those needs. Likewise, if a pre-assessment identifies a student who needs remedial or more challenging work, the instructor can accommodate those needs. The teacher can also group learners based on level of ability to encourage peer-to-peer socialization and learning, and to make sure the student continues to build on existing knowledge.

This building process has come to be known in recent years as "scaffolding." This term was developed in the 1950s by cognitive psychologist Jerome Bruner, and it originally related to the development of language skills in young children. However, it has grown to include all cognitive development based on a model of helpful interactions between an adult and a child that build the child's skills. The adult provides boundaries and perhaps even a template for the learning experience. As it relates to our methodology, scaffolding takes place in the third stage as the instructor provides an active learning experience.

Instructors may have to use some creativity when designing active learning experiences delivered in a VLE. In his constructivist learning theory, Piaget noted three types of learning in children: functional play, symbolic play, and games with rules. Traditional classrooms allow for physical functional play, such as running or jumping up and down, and symbolic play, such as making paper dolls. In the VLE, however, instructors may have to be creative when developing play-based learning experiences that lack the physical component. For example, instructors can encourage role-playing as a type of symbolic play that can be accomplished in a VLE.

The key to the active learning step is to continue to provide the socialization that is so critical in both Piaget's and Vygotsky's methodologies. It is important to note that "active" learning refers to active cognition, rather than active physical behavior. A student can be physically calm sitting in a chair in front of a computer and still be engaged in active learning. The critical piece of active learning is that the student is engaged in the learning process rather than passively receiving facts and information.

Following the active learning step in our methodology, the student returns to a brief individual post-assessment to evaluate the effectiveness of the learning experience. This stage should parallel, although not exactly mirror, the pre-assessment, to fairly evaluate what was learned. Integrating the individual stages of this methodology, the pre- and post-assessment, with the more socially interactive stages provides a balance to the learning experience and gives the student time to reflect on what has happened and what will happen next. With this VLE model, the primary advantages include the following:

1. It is a user-centric model. According to Anghern [1], the learner should be at the center of the e-learning model and this approach becomes very critical for managing individual knowledge capital and competence.
2. It is extremely interactive. The model proposed here is very dynamic and interactive. The instructor is continuously available to the student and the peers are in ready contact as well. Primary school learners need constant monitoring and this model helps the instructor mentor the young learners on a consistent basis.
3. The scaffolding effect allows the learners to use a rich learning environment to try out exploration, manipulation and construction of new ideas and learning fundamentals

Table 1: Constructivism Learning Model in a VLE

Stage	Description	Example
Socialization	Teacher gives a topic; students share by typing their responses in the chat area of the VLE.	"Today we're going to talk about USA, Who wants to share something about a state they've been to, other than the one we live in?"
Individual Pre-assessment	Online assessment of ability	Each student individually sees a map of the U.S. & a state name, the student should click on the state outline that matches the state name.
Active Learning	During the lesson, the teacher, using the VLE, gives the lesson. Each student must participate by responding to questions & Interacting in the VLE.	As the teacher talks about a state, it lights up on the screen and icons appear to illustrate the discussion. For example, as the teacher discusses Iowa's farming & biofuel industries, an icon of corn appears.
Individual Post-assessment	Online assessment of learning	Each student sees the same map as seen in the pre-assessment, but now must match icons to the states they represent.

3. Conclusion

E-learning is a far-reaching pedagogical transformation for a traditionally inclined curriculum. Digital inequality and inequitable access to online curriculum could pose a great challenge when designing a new VLE model. An easy-to-use and flexible e-learning tool could act as a strong foundation for effective distribution of school curriculum.

Many authors and academicians believe that VLEs in a primary school setting are out of context and not useful. Many schools, both in developed and developing countries, are finding deployment of VLEs very challenging. One of the biggest concerns that most of these schools have is whether the e-learning tools used are transforming the pedagogical values of teaching in an adverse manner.

Of late, there is an intense debate on the efficacy of commercial e-learning tools in primary classroom teaching. One of the concerns expressed by [4] is that the learning-management system vendors are trying to maneuver their tools as the focal point for e-learning in organizations, eventually to remove control from the tools' end-users; here, the school management, instructors and learners. Merrienboer et al [6] believes that:
"It is not the media, but the instructional methods and the delivery systems that can enhance the overall quality of education and learning experience."

Most of the VLE systems are excellent at delivering course materials. However, this cannot be the sole criteria for the success of e-learning. In his article on Learning Management Systems - The wrong place to start learning, Siemens [4] opines that:

"A good learning tool will help enhance informal learning, support superior performance and manage skills and knowledge development. In fact, as an e-learning tool becomes more sophisticated and feature-rich, it tends to lose its practical utility and basic functionality."

Significant progress has been made in designing and engineering aspects of many e-learning platforms, tools and models. Although feature-rich and good in delivery, they may lack from inherent deficiencies and user-side weaknesses. In addition, such tools have been working with a common principle as noted by Siemens [4] – tool selection first and instructor requirements second. Simplified, it means that the tool itself tends to override the user and user's preferences or options.

Gagne's theory of learning espouses the ideas of signal learning, stimulus-response learning and verbal association [9]. Although, Novak and Taylor's Stimulus-Response Theory (1977) forms the basis of Gagne's theory, the latter includes a number of cognitive factors in the learning model that eventually leads to an all-around development of young learners' mental and brain development.

References

[1] A. Angehrn, T. Nabeth and C. Roda. (2001, Nov 2001) Towards personalised, socially aware and active e-learning systems: Illustrated with the agent-based system k-inca [Online]. Retrieved April, 7 2011. http://www.calt.insead.edu/eis/documents/K-InCA-whitepaper.pdf.

[2] British Educational Communications and Technology Agency (Becta). A review of the research literature on the use of managed learning environments and virtual learning environments in education, and a consideration of the implications for schools in the united kingdom, Coventry, Becta, 2004.

[3] C.H. Orrill. (2000) Learning objects to support inquiry-based online e-learning. In D. A. Wiley (Ed.), The Instructional Use of Learning Objects: Online Version. Retrieved April 6,2011, from the World Wide Web: http://reusability.org/read/chapters/orrill.doc.

[4] G. Siemens. (2004, November 22) Learning Management Systems: The wrong place to start learning, [Online] Available. http://www.elearnspace.org/Articles/lms.htm

[5] J. Novak and R. Tyler. (1977) A Theory of Education. Cambridge University Press, New York.

[6] J. V. Merrienboer, T. Bastiaens and A. Hoogveld. (2004) Instructional design for integrated e-learning. In, Intergrated e-learning; implications for pedagogy,technology and organisation (Ed): Jochems,W; Merrienboer, JV & Koper, R.London: Routledge Falmer

[7] L.S. Vygotsky. (1978) Mind and society: The development of higher mental processes. Cambridge, MA: Harvard University Press.

[8] M.A.C. Gonzalez, D. Suthers and J.G. Escamilla De Los Santos. (2003) Coaching web-based collaborative e-learning based on problem solution differences and participation. International Journal of Artificial Intelligence in Education 13(1).

[9] R.M. Gagne. (1985) The Conditions of Learning and Theory of Instruction. New York: CBS College Publishing.

[10] S.K.H. Chu, S.K. Tse and K. Chow. (2011) Using collaborative teaching and inquiry project-based learning to help primary school students develop information literacy and information skills, Library & Information Science Research, Volume 33, Issue 2, April 2011, Pages 132-143

Development of STEAM program Math centered for Middle School Students

EunJung Kim*, SunHoi Kim & DongSoo Nam &TaeWuk Lee
Dept. Computer Education, Korea National University of Education, Korea
*mathrose@hanmail.net

Abstract: Now as a society of knowledge and information , it is required the develoment of autonomous and creative human resources that can create new intellectual value using the knowledge and information given by rather than simply memorizing the knowledge. In terms of development of creativity, mathematics is a very important subject. However, many students do not feel the need and the value of mathematics and they just study without any interests. So looking at the buzzword of these days, STEAM . STEAM is the fusion of five fields of studies which are science, technology, engineering, art, mathematics. In the early advanced study it can be seen that using STEAM , education has a positive impact on attitudes and interests . And computer programming enhances students' algorithmic thinking in mathematics. Therefore, this research wants to develop STEAM program centered with mathematics using scratch to improve the mathematical interests of students.

Keywords: STEAM, Scratch , Mathematical interest, Learning program.

Introduction

We are living in an information society away from the industrial economy such that knowledge and information is the driving force of economic development. In such a society, development of autonomous and creative human resources is required because personal and national competitiveness is determined by the value of new intellectual value-added production capacity based on the high level of information processing capabilities. As a seed of invention and discovery, creativity is the most high-level thinking skills and is the most needed ability in a information and knowledge, cutting-edge science and technology era. Thus, creativity is the main goal in all stages of school education[1].

Mathematics has begun with the history of mankind and can be found as a common language of humanity in the worldwide culture and is an essential tool in everyday life .
In particular, mathematics plays a very important role to understand and use a lot of information that is invaluable in a rapidly changing and today's information society.
It is one of the aspects of creativity in the way it makes us to think outside the fixed frame to find the solution[2].

Most students recognize mathematics as an important subject but very few students properly recognize the value of mathematics. Studying for tests, but it's rare for the fun of math, STEAM is an integrated learning approached in the aspects of science, technology, engineering, mathematics, art to one subject. This increases the efficiency of learning and motivation as well as has a positive impact on the development of learners' interest.

In addition, the linkages between various disciplines through STEAM education is consistent with the buzzword "convergence" emerged in the modern world. STEAM is a great way of education is the opinion of experts that are common. In recent years, United

"This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (NRF-2012-S1A5A2A01016950)

States, Britain, Australia, Canada, and in developed countries have an attracting attention that the core of educational reform is STEAM . In earlier studies conducted, the results that STEAM has a positive effect on improvement of the level of learning attitude in math and science are coming out[3].

According to the principle of learning in NCTM, it emphasizes that students should consist new information actively based on previous experiences and knowledges. It says learnig through understanding is effective. It means when students are in the process where they feel and contemplate by themselves, there they will experience the complete understanding and will have the possibility of creativeness developing [4].

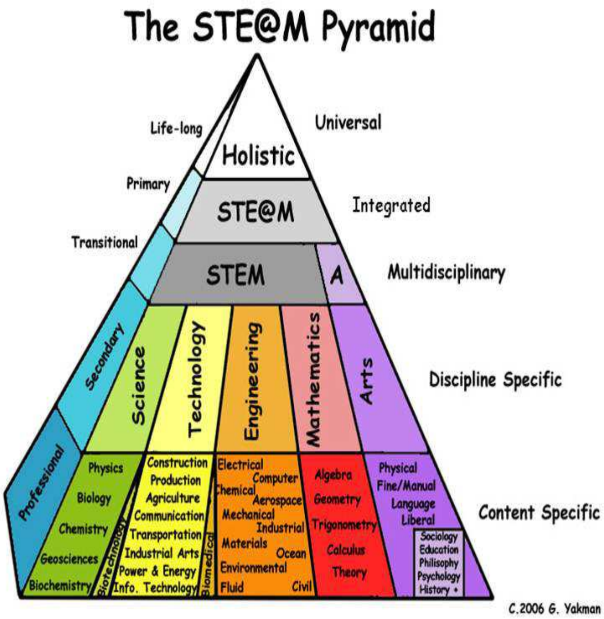
Learning math through real life is very similar in meaning with the concept of STEAM. By utilizing a computer program to solve the problem, students effort to think of the ideas and actually look the mathematics through the implement the program so it would increase the interests in mathematics.

Scratch is a programming language usable in all subjects. The command block works in a way of drag and drop. Compared to the existing engineering tools, it is more visual and dynamic, intuitive and easy to use for school students. Thus, I developed STEAM program centered with mathematics using the Scratch programming language in order to improve their math interests .

1. STEAM

STEM education was introduced in order to improve competitiveness in Science and Technology in the United States in 2003. STEM teaches science, technology, engineering, mathematics in an integrated way. In 2007 George yakman has announced STEAM in addition art to the STEM. [Figure 1] is the STEAM Pyramid. Yakman said by the STEAM education we can increase their relevance to real life and interests[4]. Many education scholars into a unified art in STEM education, said be STEAM.

Out from the dichotomous thinking such that science of conventional is logical and art is not logical, STEAM is to foster creative human resources by integrating STEM and art. Science provides a methodological tool in the art and art provides creative model in the development of science. Science uses imagination and emotion, thinking that the power of visualization principles of art and art uses scientific discoveries and principles of science[9]



2. Math based on STEAM

Principles and Standards for School Mathematics argues that mathematics should deal with topics related to the world we live in and it should be related with other fields of subjects. This is proposing mathematics should be integrated with other fields and be explored with real life[10]. Looking for examples of math related with real life, in ji-yeun Lee's paper, she finds the reason why students don't like math and says current math class teaches contents far away from real life. So she proposed contents realted with real life.

When instruct Mathematics, to generate mathematical interest and mathematical thinking and problem-solving kidneys, it is necessary to be away from a common textbook-centered learning methods but to introduce real-world problems associated with real life experience to ensure mathematical affinity[5]. When we see those products or technologies that we often use in everyday life we can see it is integrated with several studies. Thus, when teaching mathematics , it would be more interesting lessons if mathematics is integrated with other subjects and related to real life and based on STEAM .

3. Math and Computer Programming

In the bottom of Computer programming, it is situated mathematics there. In addition, in a course of mathematics, there is numerical analysis which uses computer to solve math problems and prove the algorithm. Learners improve logical thinking and can look back mathematical concepts such as functions through computer programming . Above all, using a computer. we can solve the problems which is difficult to solve by hand. The programming improves algorithmic thinking, problem-solving, particularly reflective thinking through the process of error correction[6].

Looking at the impact of computer programming on mathematical thinking, computer programming can help to develop individual's intellectual capacity by organizing the information given in a accurate and systematic way and giving the environment to arrange the information to operate logically. In the course of these programming learners are on the a position to build the system of knowledge rather than being audience[8]

4. SCRATCH

Scratch is an educational programming language made by the MIT Media Lab in 2007 for programming beginners and the teens. Programmers create programming so easily as building several blocks in eight block area. Scratch is so intuitive language that provides almost everything to make computer games, multimedia Presentations, interactive stories, illustrations, and animations [6]. If we list the features of Scratch,

1. It can be used in a variety of ages.
2. It was created based on Squeak.
3. It is programmed using the visual object.
4. It is the programming building blocks like Lego or puzzle pieces.
5. It is possible to parallel execution and stepping.
6. It is possible to utilize a variety of media.
7. Sharing and collaboration is possible.
8. The program is free and open source.

To study math based on STEAM, we need to realize it in some ways. While reviewing scratch, In the process making project using scratch, students learn how to choose and create and manage many objects while choosing many images, animations etc. And scratch makes students to respond to an unexpected situation with a creative solution. So it increases creative thinking skills [12]. So I think Scratch is a good way to realize it.

5. Methodology of the STEAM program centered with mathematics

5.1 Directivity and select content of Study program

- Choose the theme which learners can see mathematical principles and meet easily in real life
- Choose the theme which includes STEAM elements.
- Consider middle school student's interest, study level, learning ability.

5.2 Design the model for teaching

Table 1. The model for STEAM and content [7]

STEAM	content
Introduction	- Introducton Project based Learning program Motivation
Deciding Subject and planning	- Decide topic - Decide subtopic and make teams -Secure study sources.
Becoming one with project	-Investigate and research, sharing datas, Meeting and collaboration sharing and expanding thinkings - Express subject with art elements.
Presentation of the results	- Represent the results -Compare own results with others through representation
finish and evaluation	- Evaluate various methods - Finish and reconstruct

5.3 Instructional design of STEAM program centered with math.

Table 2. Example of Math Program <Making a Clock>

Table of content	Contents
Unit	Of the shape, Proportional expression
Learning goals	We can make the clock using proportional expression and the angle.
STEAM	S: The difference between analog and digital T: The development of technologies of watch E: Advanced clock A: Clock Design M: The concept of angle in geometry, proportional
Detailed activities	·Time can be displayed using the angle and proportional expression ·Understand the difference between analog and digital, and should design the clock with the advantage of this.

Table 3. Example of Math Program < Create geometric patterns >

Table of content	Contents
Unit	Function
Learning goals	Understands the graph of function and utilize

	Putting it into the props to create their own patterns.
STEAM	S: The principle of fabric dyeing T: The Development of dyeing technique E: The process of fabric dyeing A: design M: function and graph
Detailed activities	·Understand the function and can draw the graph ·Think about the principle of fabric dyeing and program it with scratch to create their own patterns and we can relate with real life.

6. Conclusion and Future Work

In this study, I developed STEAM program centered with mathematics using scratch to improve mathematical interests for middle school students. This program could increase learning interest. In the learning process, programming will help learner's mathematical thinking. To inspect the effects of this program, we have to apply this method to real class. Developed STEAM program will be taught to students in class and we'll observe the student's learning activities and analyze outcomes of learning.

STEAM-related learning programs have not been developed a lot. Learning programs based on STEAM should be developed a lot for school teachers field to utilize them. We need to discipline teachers to teach the program well and train many teachers. From now on, there should be many attempts of STEAM education. As the concept of STEAM, many teachers from various fields get together and share their knowledge to make better program.

References

[1] Dong-Hee Lee. (2009). The Effect of Climbing Learning Method on Mathematical Creativity and Inclination for Mathematical Creativity. Master Thesis. Busan national University.

[2] Byeng-Jo Kim. (2002). The Study on the effectiveness of the creativity training program by DESK model. Master thesis. Kyung Hee University.

[3] Hyo- nyung Lee. (2011). Investigation of STEM education of USA for the start of STEAM. Journal science creation .

[4] Woo- Jin Kim. (2012). STEAM Program Development and Application for Improving Creativity of the Gifted Elementary Student about Math Focused on 4D-Frame Teaching Aid Activity. . Master thesis. Korea National University of Education

[5] Ho-Soon Kim. (2002). The Effects of Problem Centered Learning based on Real Life Problems in Mathematical Beleifs and Attitudes. Master thesis. Korea National University of Education.

[6] Yun -A Kim. (2012). A study on the Function of first grade Middle School Using an Educatioal Programming Language "Scratch". Master thesis. Korea National University of Education.

[7] Soon- Beom Kwon. (2012). The Effects of Convergence Education based STEAM on Elementary School Students' Creative Personality. Master thesis. Korea National University of Education

[8] Byung-duk Ahn. (2004). Elementary computer science curriculum which based on Visual Basic programming in order to improve mathematical problem solving skills. Master thesis. Yonsei University.

[9] hyung-ju Park.(2012). A Study on analysis of Mathematical textbook based on STEAM Education. Master thesis. Ewha Womans University.

[10] Shin-deuk Lee. (2006). Development and application of ICT teaching and learning Materials driven real life and the changes of students' recognition after mathematics instruction. . Master thesis. Korea university.

[11] Ji-Yeon,Lee. (2011). Mathematics in (Real) Life. Master thesis. Chung nam National University.

[12] Natalie Rusk, Mitchel Resnick,& John Maloney. (2010) Learning with Scratch. http://info.scratch.mit.edu/sites/infoscratch.media.mit.edu/docs/scratch-21centuryskills.pdf

Digital Classrooms for Developing Subject Knowledge, Information Literacy and Twenty-first Century Skills

Siu Cheung KONG
The Hong Kong Institute of Education, Hong Kong
sckong@ied.edu.hk

Abstract: A study was conducted for creating digital classrooms to support students to learn subject knowledge and develop information literacy competency as well as critical thinking skills. A total of 144 students in four Secondary 1 classes in a secondary school in Hong Kong were invited to participate in the study. In the 12 trial lessons, every three students shared a Tablet PC for learning the topic "Air Pollution Problems in Hong Kong" in the Integrated Humanities subject. After the trial teaching, the students and teachers were interviewed about their perception of the support from the designed digital classrooms. The study found that the students and teachers positively perceived the support from the designed digital classrooms on developing subject knowledge, information literacy competency and critical thinking skills. Three recommendations are made for enhancing the pedagogical designs in digital classrooms that emphasize the use of mobile technology.

Keywords: Critical thinking skills, information literacy, Integrated Humanities, digital classroom, secondary school

1. Introduction

School education in the twenty-first century is expected to equip students with both subject knowledge and the twenty-first century skills in order to meet the requirements of a vigorously changing society [1]. Information literacy and critical thinking skills are two important parts of the twenty-first century skills [1, 2]. Information literacy (IL) refers to the mastery of necessary knowledge of gathering, synthesizing, analyzing, interpreting and evaluating information; and the proper attitudes for information processing with an understanding of the rationale behind using information [3]. Critical thinking skills are the capabilities to think reflectively and judge skillfully, so as to decide what information is reliable and what actions should be taken during reasoning and problem-solving [1, 4].

Future classrooms in the twenty-first century are digital classrooms wherein students have many chances to use digital technology to access digital resources that contain information in digital form for learning subject content [5]. Inside the digital classrooms supported by mobile technology, students use mobile devices that are wirelessly interconnected for completing learning tasks, especially accessing additional learning information from sources other than textbooks [6, 7]. For the successful learning in digital classrooms, students need to apply IL competency and critical thinking skills in daily subject learning [1, 2], in order to properly process information from different sources and then critically assimilate information for subject learning. In this regard, special pedagogical designs with appropriate learning tasks should be made for students to develop subject knowledge, IL competency and critical thinking skills in digital classrooms [1, 7].

For the effective development of subject knowledge in class, the class activities should be designed to progressively promote students' active learning, constructive learning and interactive learning [8]. In this regard, teachers should arrange class activities that engage students in the processes of resource access, knowledge construction and peer interaction. Examples of those class activities include information search, mind map drawing, and group discussion.

For the effective development of IL competency in class, there are three ways potential to prepare students to skillfully process learning information [2]. In way (i), teachers integrate IL elements into class activities, such as asking students to search information for completing learning tasks. In way (ii), teachers lead students to discuss IL elements involved in learning tasks after class activities. In way (iii), teachers guide students to discover IL elements tacit in learning tasks.

For the effective development of critical thinking skills in class, there are four ways potential to prepare students to critically solve learning problems [1]. In way (i), teachers give students sufficient time to think about the problem-solving questions. In way (ii), teachers give students sufficient time to discuss the problem-solving questions with group members. In way (iii), teachers guide different groups of students to make sharing with the whole class. In way (iv), teachers explain the answers for the problem-solving questions and then guide students to make reflection. The study reported in this paper created digital classrooms to support students to learn subject knowledge and develop IL competency as well as critical thinking skills. The class activities in the study were designed in line with the suggestions from [1], [2] and [8] as above.

2. The Study

This paper reports an initial experience from a study in Hong Kong that created digital classrooms to support students to learn Integrated Humanities (IH) knowledge and develop IL competency as well as critical thinking skills. By purposeful sampling, a total of 144 students from four Secondary 1 classes in a secondary school in Hong Kong were invited to participate in the study. Table 1 shows the profile of students participated in the study.

Table 1: Profile of students participated in the study

	Class A	**Class B**	**Class C**	**Class D**
Number of students	36	36	36	36
Ratio of boys to girls	8:28	10:26	20:16	18:18
Mean age in years	12.6	12.3	13.1	13.2

A trial teaching on the topic "Air Pollution Problems in Hong Kong" in the IH subject was conducted for each participating class. The trial teaching amounts 960 minutes, with 12 lessons of each lasted for 80 minutes. The two IH teachers of the four participating classes were responsible for the trial teaching. The years of teaching experience for the teacher of Class A and Class B was seven; while for the teacher of Class C and Class D was 13.

The students in each participating class were divided into groups of three for class activities. In a typical trial lesson, each group of students mainly used the Tablet PC provided for completing worksheets specially designed for the trial teaching. These worksheets aimed to lead students to properly process information from different sources and critically assimilate information on target topic in class. The students were first asked to answer worksheet questions by referring to the paper-based textbook and the selected e-learning resources on the school-based e-learning platform. The students then discussed worksheet questions with their group members. In the discussion activities, the students needed to apply critical thinking skills such as identifying keywords in the worksheet questions for problem-solving, and to apply IL competency such as searching information from reliable websites. After the discussion activities, the students were guided by their teachers to present and discuss their works with the whole class. In the trial teaching, the teachers integrated the three ways suggested by [2] and the four ways suggested by [1] for developing students' IL competency and critical thinking skills, respectively. The study focused on two research questions:

(1) How do the students perceive the support from the designed digital classrooms on developing IH knowledge, IL competency and critical thinking skills?

(2) How do the teachers perceive the support from the designed digital classrooms on developing IH knowledge, IL competency and critical thinking skills?

3. Methods

Semi-structured interviews were conducted with the students and teachers participating in the study for investigating their perception of the support from the designed digital classrooms on developing IH knowledge, IL competency and critical thinking skills. Two groups of students were randomly selected from the four participating classes for two group interviews. The two teachers participating in the study were invited for two individual interviews. The semi-structured interviews focused on three aspects, namely the school-based approach of developing IH knowledge in class; the three suggested ways of developing IL competency in class; and the four suggested ways of developing critical thinking skills in class. The invited students and teachers were asked about their experiences, comments and expectations of learning and teaching in the designed digital classrooms.

4. Results and Discussion

In general, the students and teachers perceived that the designed digital classrooms could support the development of IH knowledge, IL competency and critical thinking skills.

Table 2: Major interview feedback on the development of IH knowledge in the trial lessons

	Major interview feedback
Experiences	• The students and teachers indicated that the school-based e-learning platform was used in class to support the convenient access to additional resources and easy drawing of mind maps. • The students pointed out that the trial teaching emphasized much on the use of mobile devices and e-learning resources, but less on the use of school-based textbook. • The teachers reflected that they were unable to give their students sufficient time and guidance to interpret the key contents of the additional resources within class time.
Comments	• The students and teachers agreed that the use of additional resources and the drawing of mind maps were effective for supporting concept building and knowledge consolidation. • The students indicated that after trial teaching they were able to link the knowledge learned to other topics of the same subject, such as economic development and ecology system. • The teachers suggested that the additional resources selected should match students' learning stage, avoid lengthy and complicated documents that students cannot fully comprehend.

Table 2 shows the major interview feedback from students and teachers on the development of IH knowledge in the trial lessons. The results showed that the teachers were able to implement the school-based approach of developing IH knowledge in trial lessons, with a focus on class activities of accessing e-learning resources and drawing mind maps for subject learning. The students and teachers valued the many chances to access update learning information and exchange subject knowledge in the class activities, which promoted students to better understand the target topic and link knowledge among different topics in the target subject. The students and teachers expected for the use of more additional e-learning resources that match the level and need of students in learning within class and after school.

Table 3: Major interview feedback on the development of IL competency in the trial lessons

	Major interview feedback
Experiences	• The students and teachers indicated that all the three suggested ways were adopted in class, with way (ii) was adopted most frequently; way (i) followed; and then way (iii). • The students were impressed that the teachers often adopted way (ii) to lead the whole class to discuss effective ways of information search after completing every group task. • The teachers reflected that they did not have sufficient time to adopt way (iii) to guide students to discover IL elements tacit in the worksheets questions.
Comments	• The students and teachers thought that way (ii) was most effective for supporting junior secondary students to develop IL competency, with way (i) the second and way (iii) the least. • The students indicated that after trial teaching they often tended to search and compare information from various reliable websites for judging the suitability of target information. • The teachers suggested that junior secondary students should be provided with more guidance on the rationale behind IL.

Table 3 shows the major interview feedback from students and teachers on the development of IL competency in the trial lessons. The results showed that all the three suggested ways for developing IL competency were implemented in trial lessons, with different levels of frequency and effectiveness. The way of teacher-led discussions about IL elements after learning tasks was most frequently implemented and regarded as the most effective way to support junior secondary students to develop IL competency in class. The students and teachers generally perceived that these three suggested ways made students become more aware of the necessary knowledge and proper attitudes for processing information. The teachers hoped to provide more guidance for junior secondary students to explore and reflect on the IL elements tacit in the process of learning tasks in class.

Table 4: Major interview feedback on the development of critical thinking skills in trial lessons

	Major interview feedback
Experiences	• The students and teachers indicated that way (i) and way (iii) were adopted in almost every lesson; while way (ii) and way (iv) were often omitted because of time limit. • The students were impressed that the teachers often adopted way (iii) and provided special hints for question interpretation, as well as award incentives to peer discussions in class. • The teachers reflected that their students lacked sufficient time for discussions in problem-solving tasks, and lacked sufficient guidance in thinking exercises.
Comments	• The students and teachers thought that way (iii), which emphasized sharing and discussion among students, was very helpful for the development of critical thinking skills. • The students and teachers indicated that the trial teaching gradually promoted skills in critical thinking, such as distinguishing keywords “relieve” and “resolve” for problem-solving. • The students and teachers expected for the provision of clear guidelines and sufficient time for group discussion in class in future.

Table 4 shows the major interview feedback from students and teachers on the development of critical thinking skills in the trial lessons. The results showed that all the four suggested ways for developing critical thinking skills were implemented in trial lessons, with different levels of frequency and effectiveness. The way of teacher-guided sharing and discussion among students was most frequently implemented and regarded as the most effective way to support junior secondary students to develop critical thinking skills in class. The students and teachers generally perceived that these four ways made students become more aware of the process and skills in critical thinking, such as distinguishing different keywords “relieve” and “resolve” for problem-solving. The students and teachers expected for more time and guidance for the whole class to discuss and reflect on the process and products of critical thinking in subject lessons.

5. Conclusion

The study found that students and teachers in general positively perceived the support of the designed digital classrooms on facilitating the learning of IH knowledge and the development of IL competency as well as critical thinking skills. Based on the results of the study, three recommendations are made for the future development of digital classrooms using mobile technology in school education. First, the teachers in the interviews made a reflection that there was insufficient time for students’ discussions in trial lessons. Teachers are therefore recommended to increase time for interactive learning in digital classrooms. Second, the teachers in the interviews made a reflection that they did not make sufficient guidance in the use of e-learning platform for promoting students to develop IL competency and critical thinking skills. Teachers are therefore recommended to increase the use of e-learning platform and enhance their work on guiding students to develop IL competency and critical thinking skills before and after class time. Third, the teachers were observed to have difference in the mastery of pedagogies for promoting students to develop IL competency and critical thinking skills. Schools are therefore recommended to provide support of teacher professional development related to digital classrooms, so as to prepare teachers for the teaching work in the twenty-first century.

The second round of the study is conducting to collect quantitative data on the achievement and perception of students in learning the target topic in the designed digital classrooms. More results related to the impact of mobile technology supported classrooms on formal subject learning will be further reported.

References

[1] Gut, D. M. (2011). Integrating 21st century skills into the curriculum. In G. Wan & D. M. Gut (Eds.), *Bringing schools into the 21st Century* (pp.137-157). Dordrecht; New York: Springer.

[2] Kang, M., Heo, H., Jo, I. H., Shin, J., & Seo, J. (2010). Developing an educational performance indicator for new millennium learners. *Journal of Research on Technology in Education, 43*(2), 157-170.

[3] Kong, S. C. (2008). A curriculum framework for implementing information technology in school education for fostering information literacy. *Computers and Education, 51*(1), 129-141.

[4] Ennis, R. H. (2002). Goals for a critical thinking curriculum and its assessment. In A. L. Costa (Ed.), *Developing minds* (3rd Edition) (pp. 44-46). Alexandria, VA: ASCD.

[5] John, P. D., & Wheeler, S. (2008). *The digital classroom: harnessing technology for the future of learning and teaching*. London: Routledge.

[6] Wong, L. H., & Looi, C. K. (2011).What seams do we remove in mobile assisted seamless learning? A critical review of the literature. *Computers and Education, 57*(4), 2364-2381.

[7] Chang, S. B., Lin, C. J., Ching, E., Cheng, H. N. H., Chang, B., Chen, F. C., Wu, D., & Chan, T. W. (2009). EduBingo: Developing a content sample for the one-to-one classroom by the content-first design approach. *Educational Technology and Society, 12*(3), 343-353.

[8] Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73-105.

An Analysis of using Social Bookmarking for Acquiring Web Resources for e-Learning

Tak-Lam WONG*, Siu Cheung KONG
Department of Mathematics and Information Technology,
The Hong Kong Institute of Education, Hong Kong
*tlwong@ied.edu.hk

Abstract: With respect to the increasing trend that students acquire information from the Internet for their formal or informal learning, we aim at analyzing the use of social bookmarking to help students collaboratively access useful and creditable Web resources, facilitating collaborative learning. We also investigate the use of social bookmarking for after-school learning. Social bookmarking allows users to search for, tag, and share Web resources, and hence recommend students Web resources that are potentially useful and relevant. Experiments were conducted to study how social bookmarking can improve the quality of the acquired Web resource. In particular, students using social bookmarking could acquire a more focused and high quality set of Web resources compared with students not using social bookmarking.

Keywords: social bookmarking, Web resources, collaborative learning

Introduction

With the rapid growth of Internet technology, students are readily to access different kinds of information in the forms of online articles, Web sites, Wikis, etc., from the World Wide Web (the Web) for learning. Students increasingly use the Web to obtain informative Web resources to replace visiting libraries [1] for learning or completing their assignments. In light of this, this raises the need of cultivating the information literacy of students and developing their ability to evaluate the credibility and relevance of the Web resources such as Web pages, documents, journal articles, etc. [2]. On the other hand, finding accurate, trustful and reliable Web resources becomes essential. Tools are needed to assist students to acquire the correct information, so as to avoid them from obtaining incorrect information and learning the “wrong knowledge” [3,4].

Though a number of researches have been conducted and showed that social media can help collaborative learning [5], the use of social media to help students collaboratively obtain useful Web resources has not been well investigated. Social bookmarking, which is a particular form of social media, can help students collaboratively access useful and share Web resources. Students can supply key terms to the social bookmarking system to search for relevant Web resources. Instructors and students can provide additional information to their acquired Web resources by making use of descriptive tags if the resources are useful. Moreover, each user of the social bookmarking system can see the tags given by other users to any Web resources and have a rough understanding of the Web resources. Figure 1 illustrates the relationship between the Web resources, users, and tags in a social bookmarking system. The three entities are inter-dependent and correlated. For example, suppose user ***u*1** tags a Web page ***d*3** in the system tags ***t*2** and ***t*4**. Users ***u*2** and ***u*3** may read and give tags to the same Web page ***d*3** because of tags given may raise their interest.

Similarly, the acquired Web resources and the tags given by users ***u*2** or ***u*3** will influence ***u*1**. This leads to the situation that multiple users may access the same set of resources via the system and users have influence on each other via the acquired resources and the tags given, leading to collaborative learning. Therefore, social bookmarking is particularly useful for sharing Web resources. However, from our best knowledge, no effective pedagogical approaches have been developed to exploit social bookmarking to help learners acquire useful and quality Web resources. In this paper, we aim at analyzing the use of social bookmarking to help learners search for quality Web resources and investigate the possibility to use social bookmarking for e-learning and after-school learning. Durlak & Weissberg showed that after-school programs could improve students' self-perceptions and bonding to school, their positive social behaviors, and their school grades and level of academic achievement [6]. In particular, our research questions are as follows:

1. Can social bookmarking improve the quality of acquired web resources for e-learning?
2. Is there any potential for adopting social bookmarking for after-school learning?

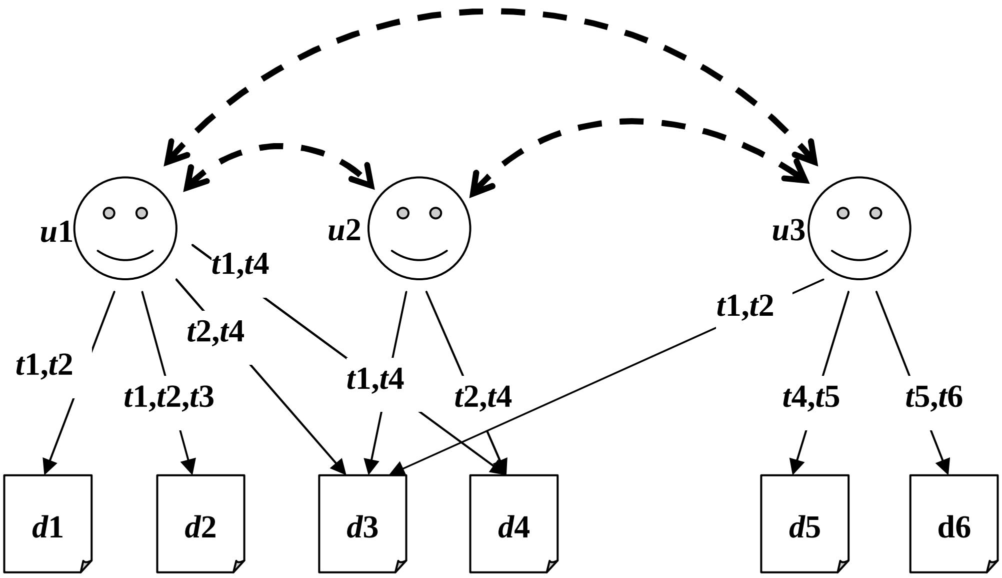


Figure 1. An illustration to show the relationship between users, tags, and Web resources

1. Literature Review

The vast amount of information on the Web triggers the need for students to cultivate their information literacy [7,8]. One important component of information literacy is to evaluate the usefulness, trustfulness, and creditability of the information acquired. Brandt has discussed several aspects to evaluate information on the Internet [9]. A study has been conducted to investigate the practices of students to verify the credibility of the Web information [1]. Another approach has been proposed to evaluate the credibility of the information in Wikipedia, which can be authored by any Internet user and may contain inaccurate information [2]. However, this method mainly focuses on Wikipedia articles instead of generic Web resources. On the other hand, several methods have been proposed to help access useful information. For example, Lee proposes a method to help students retrieve Web sites based on the searching patterns of the peers in a course [3]. The main idea of the method is to recommend a student potentially useful Web sites by employing association rule data mining algorithm to analyze the searching patterns from students in the same course. One limitation of this method is that it lacks interaction between students and hence not having sufficient training of collaboration to students. Mangina and Kilbride propose another approach to making personalized recommendation of documents based on user modeling and information retrieval techniques [4]. However, little prior research on such recommendation systems has focused on the quality of the accessed Web resources and assisting students to obtain useful and creditable Web information.

Different forms for computer-mediated communication applications and social media have been utilized in E-learning [10]. For example, using asynchronous online discussion forums has been shown to be effective in constructing deep knowledge among students [11]. Forms of social media such as Wikis, blogs, microblogs, etc, have been applied to facilitate formal or informal learning of students [5,12]. However, some researches show that it is doubtful in the effectiveness for using social media in learning [13]. Besides, the success of using social media greatly relies on the participation of students [14].

2. Method

20 students, who were in-service teachers of primary or secondary schools, from a course were invited to conduct the experiments. The students had general knowledge of computer and Internet usage. We setup a social bookmarking system for our classes[1]. A 30-minute training session about the usage of the social bookmarking system was given to all students to familiarize how to search for, tag, and share Web resources. In addition, the criteria of a good Web resource and examples of good Web resources were illustrated.

An experiment with experimental group and a control group was used as the overall research design. Students were divided into two groups. The first group consisted of 11 individual students who searched for Web resources using their own method. This was the control group of the experiment. The second group consisted of 9 individual students who used the social bookmarking system to search for and share Web resources. Students from both groups were asked to spend 30 minutes to search for relevant and useful Web resources in a topic. During this period, they were not allowed to communicate, so as to demonstrate the collaboration effect of social bookmarking. In the experiment, they were asked to search for resources about "information literacy". The obtained Web resources were then evaluated quantitatively and qualitatively. Quantitatively, we measured the number of Web resources found by each student, and the number of Web resource tagged by each subject. Qualitatively, we invited human experts to assess different aspects of each of the Web resource found by the students. Note that the human experts did not know the students from whom the Web resources originate. Each human expert is required to complete the following survey for each Web resource. Some questions of the survey are referenced to [2,9]. For each of the questions, a 5-point rating is used to score the resource, where point 1 refers to the most disagree and 5 refers to the most agree.

Q1: The resource is relevant to the topic.
Q2: The resource is trustworthy (e.g., the resource is published in an authority Web site, journals)
Q3: The resource is authoritative (e.g., the resource is published with names of authors, contain list of references)
Q4: The resource is accurate (e.g., the resource does not contain incorrect information)
Q5: The resource is unbiased (e.g., the resource does not have related advertisement)
Q6: The resource is informative (e.g., contain advanced knowledge, analysis)

[1] We employ the open-source social bookmarking system *scuttle*, which can be obtained in *http://sourceforge.net/projects/scuttle/*.

Q7: The resource is well-organized (e.g., the resource is easy to follow.)
The questions of the survey were provided to student as assessment rubrics to the Web resources acquired.

3. Results and Discussion

Table I shows the statistics obtained from the experiments. We have conducted statistically hypothesis testing, specifically, paired t-test, to evaluate the following hypothesis.

Hypothesis: There is no improvement in the quality of the acquired Web resources via using social bookmarking.

For Q1 of our survey, the average points of the experimental group and the control group were 4.46 and 4.25 respectively. It could be observed both experimental group and control group could acquire relevant Web resources for a particular topic. However, the points of the experimental group were significantly higher than those of control group from Q2 to Q7. It showed that using social bookmarking systems could improve the quality of the Web resources. The major reason was that students could collaboratively search for good Web resources, and share them via tagging. Students could give meaningful descriptive tags the Web resources, which were considered to be in high quality by the students. On the contrary, the students in control group could only search for Web resources individually without any collaboration and communication, and hence good resources could not be shared.

Table I. Statistics of the experimental group and control group

	Experiment Group		Control Group	
Question	Mean	S.D.	Mean	S.D.
Q1	4.46	0.65	4.25	0.64
Q2	3.86*	0.99	3.47	0.94
Q3	3.80*	0.95	3.40	0.88
Q4	4.09**	0.77	3.66	0.73
Q5	4.17*	0.65	3.77	0.90
Q6	4.14*	0.83	3.66	0.82
Q7	4.09**	0.81	3.66	0.87

*$p < .05$, **$p < .01$

The experimental group and the control group obtained a total of 26 and 39 different Web resources respectively. The average numbers of different Web resources acquired by each student in the experimental group and the control group were 2.88 and 3.54 respectively. Experimental group obtained less but better quality set of different Web resources because students could read the shared resources via the social bookmarking system. For the control group, students obtained a more diversified Web resources, but the average quality was less satisfactory. This illustrated that students using social bookmarking obtained a more focused set of good resources. This is particularly useful if students are asked to have group discussion during classes after studying these resources because they have a better common ground for discussion. In light of this, social bookmarking can be effective for after-school learning. Students can access the social bookmarking system via the Internet and collaboratively search for Web resources. After-school learning activities can be carefully designed so that students can leverage social bookmarking to receive useful information and knowledge. For examples, tasks of searching for resources of a particular topic can be designed as after-school learning activities for students. Such tasks can help students obtain additional information and deep knowledge outside classes.

4. Conclusions

This study conducted an experiment to investigate the use of social bookmarking to acquire quality Web resources. The results indicated that the students can obtained a more focused and better set of Web resources in terms of different dimensions such as relevance, trustworthiness, etc. The findings suggest that social bookmarking can be effective for after-school learning. Particularly, social bookmarking can be integrated into project-based learning or inquiry-based learning. Project-based learning aims at engaging students investigating and tackling challenging problems. The objective of inquiry-based learning is to construct knowledge via generating, investigating and answering questions. Social bookmarking can help students obtain good resources in a collaboratively to analyze the problems and make decision, facilitating project-based learning and inquiry-based learning.

Acknowledgements

The work described in this paper is substantially by the Internal Research Grant from the Hong Kong Institute of Education, Hong Kong (Reference Number: RG 76/2010-2011).

References

[1] Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student Web use, perceptions of information credibility, and verification behavior. *Computers & Education*, 41 (3), 271-290.

[2] Kubiszewski, I., Noordewier, T., & Costanza, R. (2011). Perceived credibility of Internet encyclopedias. *Computers & Education*, 56 (3), 659-667.

[3] Lee, Y.-J. (2005). VisSearch: A collaborative Web searching environment. *Computers & Education*, 44 (4), 423–439.

[4] Mangina, E., & Kilbride, J. (2008). Evaluation of keyphrase extraction algorithm and tiling process for a document/resource recommender within E-learning environments. *Computers & Education*, 50 (3), 807-820.

[5] Huang, S.-L., & Yang, C.-W. (2009). Designing a semantic bliki system to support different types of knowledge and adaptive learning. *Computers & Education*, 53 (3), 701-712.

[6] Durlak, J.A., & Weissberg, R.P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicao IL: CASEL.

[7] Kong, S. C. (2008). A curriculum framework for implementing information technology in school education to foster information literacy. *Computers & Education*, 51 (1), 129-141.

[8] Kong, S. C., & Li, K. M. (2009). Collaboration between school and parents to foster information literacy: Learning in the information society. *Computers & Education*, 52 (2), 275-282.

[9] Brandt, S. D. (1996). Evaluating information on the Internet. *Computers in Libraries*, 16 (5), 44-46.

[10] Barnes, S. B. (2003). *Computer-mediated communication: Human-to-human communication across the Internet*. Boston, MA: Allyn and Bacon.

[11] Berge, Z., & Collins, M. (1995). Computer-mediated scholarly discussion groups. *Computers & Education*, 24 (3), 183-189.

[12] Cavus, N., Bicen, H., & Uzunboylu, H. (2011). The efficient virtual learning environment: A case study of web 2.0 tools and Windows live spaces. *Computers & Education*, 56 (3), 720-726.

[13] Cole, M. (2009). Using Wiki technology to support student engagement: Lessons from the trenches. *Computers & Education*, 52 (1), 141-146.

[14] Hrastinski, S. (2009). A theory of online learning as online participation. *Computers & Education*, 52 (1), 78-82.

The Role of Teacher-Teaching Experience in VAE Teachers' Integration of ICT

Mohd Khairezan RAHMAT[a*] & Wing K. AU[b]
[a]*School of Education, University of South Australia, Australia*
[b]*School of Education, University of South Australia, Australia*
rahmk001@mymail.unisa.edu.au

Abstract: The introduction of Information and Communication Technology (ICT) and its integration has made a significant impact on the Malaysian education system. Its rapid development has become a major priority in achieving the nations' Vision 2020 of becoming a fully developed country by the year 2020. Despite its huge potential advantages, the Visual Art Education (VAE) teachers were reportedly not utilizing the full potential of ICT in the Visual Art classroom. Based on the literature, it appeared that only a limited published research had tried to explain the influence of teachers' experiences on their ICT integration (ICTInt). The purpose of this study **is** to explore the VAE teacher**-** teaching experience as moderator variables in their decision to integrate ICT. The Technology Acceptance Model (TAM) was used as a theoretical grounding of the study. Based on TAM, perceived usefulness (PU) and perceived ease of use (PEoU) were two specific beliefs that determine individuals' ICTInt. Self-reported data was gathered through an on-line survey from 296 VAE teachers in the state of Selangor, Malaysia. The findings revealed that both PU and PEoU were able to explain 65 percent of the variance toward their ICTInt. The findings also show that teacher-teaching experience was found to influence the relationship between their PU and PEoU toward ICTInt. It is expected that findings derived from this study will contribute toward assisting the school administrator and the Ministry of Education in ensuring all teachers would pose and instill a degree of ICT integration.

Keywords: Perceived usefulness, perceived ease of use, ICT integration

Introduction

Over the past two decades, the Malaysian government has been proactive in fulfilling the demands of a technologically literate, creative and innovative workforce for the Information Age. As an important agent of change, teachers are required to make a complex adjustment and substantial rethinking of new possibilities in teaching. However, local researches have indicated ICT integration is far from reaching its target. Many senior teachers are also reported reluctant to change their conventional way of teaching [1]. The notion of users' experience having an impact on their decision to integrate a proposed technology has been highlighted by previous researchers [2]. The embracement of ICT in VAE classrooms will provide opportunities for students to construct their own knowledge, meaning and solution. The potential usages of ICT in conducting successful art instruction have also been noted through its simulation, manipulation and creative expression activities [3]. Despite those advantages, many VAE teachers were reported to be reluctant to embrace ICT into the VAE classroom [4]. Given the importance to accelerate the integration of ICT, the study attempts to explain the relationship between VAE teachers' perceived usefulness (PU) and perceived ease of use (PEoU) toward ICT integration (ICTInt). It is envisaged that the findings of this study will act as a guide and reference for school administrator and the Malaysian government toward establishing a standard of successful integration of ICT.

1. Theoretical Framework and Hypotheses Development

Adopted from the Technology Acceptance Model (TAM), users' PU and PEoU were two specific beliefs that determine their decision to integrate a proposed technology. PU is defined as individual degree of beliefs that using particular technology would enhance job performance; while PEoU is defined as individual's degree of beliefs that using particular technology would free them of any effort [5]. Although users might perceive a proposed technology as being advanced, they will not adopt it if they think the technology cannot facilitates them [6]. The first hypothesis was formulated as:

H_1: The VAE teachers' PU and PEoU will positively influence their ICTInt.

In determining a potential moderator that might influence teachers' intention to integrate ICT, their teaching experience was also considered. Previous research on students' acceptance of Web-based courseware by Stoel and Kyu [7] found that as the experience with technology increases, users' perceive it to be easier and more useful, which in turn leads to more usage. In similar vein, Ramayah [1] research found that prior experience has moderated the relationship between PU, PEoU and computer usage. Based on the following justifications, the second hypothesis is stated as below:

H_2: The relationship between the VAE teachers' PU and PEoU toward ICTInt is moderated by their teaching experience.

2. Research Methodology

This study employed a quantitative research method through 20 items on-line survey. Both PU and PEoU questions were adopted from the TAM questionnaire [5]. The items used to measure users' ICTInt were adopted from the UTAUT questionnaire [8]. A five point Likert-type scale was employed, starting with SD as strongly disagree, D = disagree, U = undecided, A = agree and SA = strongly agree. The VAE teachers in the state of Selangor (n=887), Malaysia were identified as a target sample of the study. Out of that number only 296 (33.4%) teachers responded and thus become the respondent of this study. This response rate was acceptable for an on-line survey [14].

3. Research Findings and Results

3.1 Reliability and Validity of the Instrument

The Cronbach's alpha coefficient was used to test the reliability of items of the questionnaire. As can be seen from Table 1, all the measurement items were considered to be good (exceed 0.80) [15]. The result also indicates a large correlation between items (exceed 0.50), thus confirming the reliability of the instrument.

Table 1: Cronbach's Alpha and Item-to-total Correlation Values

Measurement Items	No. of Items	Cronbach's Alpha	Item-to-total Correlation
Perceived Usefulness (PU)	6	0.882	0.659 – 0.735
Perceived Ease of Use (PEoU)	6	0.881	0.620 – 0.742
ICT Integration (ICTInt)	8	0.806	0.500 – 0.858

A factor analysis was used to test the construct validity of the questionnaire. The result suggested that all items were represented by three factors. The KMO value was 0.860 and the significant level was .000. These results thus indicate that all items are tapping and collapsing into the same constructs [15].

3.2 Descriptive Analyses

This sub-section presents the descriptive analyses of teachers' teaching experience. Teachers' maximum years of teaching experience were divided into three groups, which represent the low, moderate and high experience teachers. As can be seen from Table 2, a majority (67.9%) of the respondents were categorized as low experience, where their teaching experience was less than eight years on average.

Table 2: Summary of Characteristics of Respondents of the Study

Characteristics	**Group**	**Cases (n)**	**Percentage (%)**
Teaching Experience	Low Experience	201	67.9
	Moderate Experience	60	20.3
	High Experience	35	11.8

3.3 Hypothesis Testing

H_1: The VAE teachers' PU and PEoU will positively influence their ICTInt.

The result reported from Table 3 shows that the significant F-Change value was highly significant ($p<0.01$). The coefficient determination (R-square) value was found to be 0.651; which indicates that 65 percent of the variance of the VAE teachers' ICTInt can be explained by their PU ($B = 0.77$, $p<0.05$) and PEoU ($B = 0.79$, $p<0.01$). This result indicates that research hypothesis 1 was accepted.

Table 3: Regression Analysis between the Teachers' PU and PEoU toward ICTInt

Model	**Unstandardized Coefficients**		**Standardized Coefficient**	**t**
	B	**Std. Error**	**Beta**	
PU	0.112	0.050	.077	2.227*
PEoU	0.733	0.32	.797	23.025**
R Square	0.651			
R Square Change	0.651			
F Change	273.583			
Sig.-F	0.000			

*p<0.05; **p<0.01

H_2: The relationship between the VAE teachers' PU and PEoU toward ICTInt is moderated by their teaching experience.

It is apparent from Table 4, the significant F-Change value in the third block was significant ($p<0.01$) when the moderator was entered in the second block. The R-square value of 0.507 in the third block also indicates that 50.7 percent of the variance between teachers' PU and PEoU toward ICTInt was moderated by their teaching experience. This result indicates that hypothesis 2 was accepted.

Table 4: Moderating Effect of Teachers' Teaching Experience on the Relationship between Teachers' PU and PEoU toward their ICTInt

Variables	Standardized Beta Step 1	Standardized Beta Step 2	Standardized Beta Step 3
Predictors			
PU	.106*	.097*	.075
PEoU	.633**	.597**	.853**
Moderator			
Teaching Experience		-.251**	.213
Interaction Terms			
PU*Teaching Experience			-.680**
PEoU*Teaching Experience			-.196
R Square	.429	.491	.507
Sig. F-Change	.000	.000	.010

*p<0.05; **p<0.01

Two graphs were drawn in showing the moderating effect of teachers' teaching experience. Firstly, teachers' PU and PEoU were categorized into two levels (low and high) based on its median score. From Figure 1, it is apparent that ICTInt among the VAE teachers who gained low teaching experience increase dramatically when ever their level of PU and PEoU was high. This finding also confirms that VAE teachers' teaching experience was moderate the relationship between their PU and PEoU toward ICTInt.

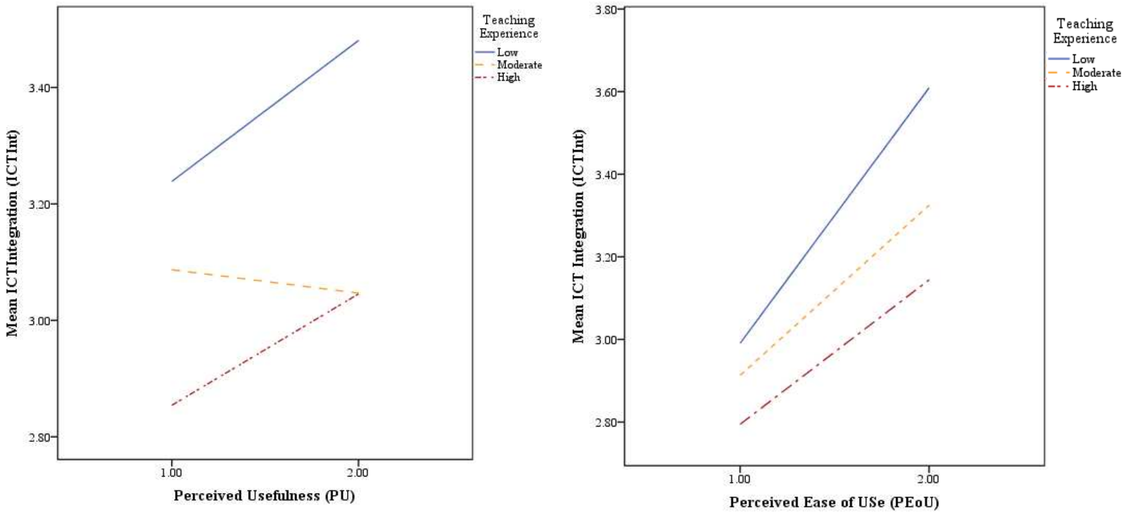


Figure 1: The Moderating Effect of Teachers' Teaching Experience on the Relationship between Teachers' PU and PEoU toward ICTInt

4. Discussion

The aim of the study is to test the relationship between the VAE teachers' PU and PEoU towards their ICTInt in the VAE classroom. Result of this study found that both PU and PEoU have positive significance effect on the VAE teachers' ICTInt. This finding is consistent with beliefs postulated in the TAM and UTAUT. It is confirmed that individuals' integration of technology is increased whenever the technology is perceived to be useful and easy to be used [1]; [6]. In ensuring better acceptance to the proposed technology, it is also possible to conclude that the technology should be perceived as better than the existing ones

[3]. The result from this study also shows that teaching experience does moderate the relationship of both the teachers' PU and PEoU toward ICTInt. This result support previous research that individuals' experience should be considered in determining their intention to integrate proposed technology [2]. The result indicates that when a teacher becomes comfortable and confident with their teaching, they are less likely to be attracted to integrate ICT into their instruction.

5. Implication and Conclusion

The result of this study demonstrates that the VAE teachers' PEoU was the more influential driver than PU in determining their ICT integration. Therefore, considering teachers' expectation toward ICT functions are required in proposing its usage in classrooms. The authorities (viz. the Ministry of Education) need to reflect on suggestions to better meet teachers' performance expectations. Basically, when ICT is perceived easy to use, teachers feel it is more useful; therefore, they will integrate it. Further, the result of the present study also suggests that teachers who have experience and confident with their subject area, are more likely to perceive the usefulness and ease of use of ICT, thus affecting their decision to integrate it [7]. Finally, the study suggests the needs of ongoing professional development courses, sufficient facilities and continues supports from authorities in ensuring teachers are ready to adopt ICT effectively into their instruction. Teachers' training colleges and universities must also take full responsibility in producing knowledgeable, skilful and confident teachers.

References

[1] Ramayah, T. (2010). The role of voluntariness in distance education students' usage of a course website. *The Turkish Online Journal of Educational Technology, 9*(3), 96-105.

[2] Baennett, R., Hartel, C. E. J., & McColl-Kennedy, J. R. (2005). Experience as a moderator of involvement and satisfaction on brand loyalty in a business-to-business setting. *Industrial Marketing Management, 34*, 97-107.

[3] Gregory, D. C. (2009). Boxes with Fire: Wisely Integrating Learning Technologies in the Art Classroom. *Art Education, 62*(3), 47-54.

[4] Roland, C. (2007). *The Art Teacher's Guide to the Internet*. MA: Davis Publications.

[5] Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340.

[6] Lai, C., & Gu, M. Y. (2011). Self-regulated out-of-class language learning with technology. *Computer Assisted Language Leaning, 24*(4), 317-335.

[7] Stoel, L., & Kyu, H. L. (2003). Modelling the effect of experience on student acceptance of Web-based courseware. *International Research: Electronic Networking Applications and Policy, 13*(5), 364-374.

[8] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward A Unified View. *MIS Quarterly, 27*(3), 425-478.

[14] Sekaran, U., & Bougie, R. (2010). *Research Methods for Business: A skill building approach*. New Jersey: John Wiley & Sons.

[15] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective* (7th ed.). Upper Saddle River, New Jersey: Englewood Cliffs, Prentice Hall.

Designing and implementing e-learning classrooms to improve students' writing

Kevin Kai-Wing CHAN*, Siu Cheung KONG
Department of Mathematics and Information Technology,
The Hong Kong Institute of Education, Hong Kong
* kevin@mysmartedu.com

Abstract: This paper aims to critically review how we can incorporate technology in school education to provide students with the critical skills needed for the 21st century. This paper will discuss how the knowledge economy demands a new set of survival skills from our students, and how appropriate use of pedagogical use of technology can enhance the quality of school education in the twenty-first century. Technological Pedagogical Content Knowledge (TPCK) framework and Active-Constructive-Interactive (ICAP) framework have been reviewed to examine the important balance of content, technology and pedagogy in designing and implementing effective e-learning classrooms for our students. The design of an online automated writing evaluation system has been examined to evaluate how it can improve students' English writing and their 21st century skills. This system will be put into practice in the next stage of this study.

Keywords: Process writing, 21st century skills, English writing skills, evaluation of CAL systems, automated writing evaluation

Introduction – The world is changing FAST!

IBM conducted a survey with 1,500 Chief Executive Officers from 60 countries in 33 industries in 2010 reveals alarming results [11]: 1) *Fewer than HALF* of the CEOs believe that their companies are equipped to deal with complexities and velocity of a world that is changing on a massively interconnected system. 2) The *impact of technology* on organizations has jumped from the 6th to the 2nd place in importance, and the executives believe that more technology-based solutions will be needed in the 21st century. 3) *Creativity* has been selected as the most crucial factor for future success in an increasingly complex world. *Are our students equipped with these skills?*

Thomas Friedman [8] also makes his case on his book, *The World is Flat*, that the world is shifting from an industrial economy to a highly complex knowledge economy. This global and Web-enabled platform allows any individual, any group, any school and any organization in the world to use new tools to communicate and collaborate. Wagner argues that the current school systems only prepare students for the tests, and even the best schools in the US do not teach the must-have skills students need to have to survive in the 21st century [17]. He describes such issue as the "*Global Achievement Gap*" and proposes that students must acquire subject content knowledge as well as the seven survival skills for the twenty-first century: critical thinking and problem solving, collaboration across networks and leading by influence, agility and adaptability, initiative and entrepreneurialism, effective oral and written communication, accessing and analyzing information, and curiosity and imagination.

1. The 21st Century Skills (21C)

1.1 What are the 21st Century Skills that our students need to develop?

Knight describes that education systems in the 20th century prepared people for work related to manufacturing, and supported a vocational training mentality [13]. The knowledge economy in the 21st century focuses on the trade in knowledge through the medium of communication technology. Knight indicates that there is a need to transform the traditional models of education with one that reflects the knowledge economy and the need for lifelong learning. Kong [12] further summarizes the 21st century skills (21C skills) as the skills needed to achieve the desired learning outcomes/educational goals in the 21st century: inquiry, critical thinking, communication and collaboration.

2. E-learning classrooms for the 21st century

2.1 Educational goals for the e-learning classrooms

When we design e-learning classrooms for the 21st century, it is important for us to make sure that we are providing students and teachers with an environment to support to achieve the 21st century educational goals [12]: use digital technology to facilitate learning and teaching, increase students' autonomy through the usage of ICT (on determining the educational goals and learning strategies, shift towards a more learner-centric model), as well as to provide students with more authentic learning opportunities through simulations, inquiry and collaborative learning (to develop students' inquiry, reflection, communication and collaboration skills)

2.2 Designing the e-learning classrooms

Many researches indicate that computer-assisted instruction in class had a positive effect on teaching and learning [6]. Mishra & Koehler [15] make the case that merely introducing technology to the educational process is not enough. They proposed a conceptual framework which describes three main components of the learning environments: content, pedagogy and technology. They emphasize the connections, interactions, affordances between and among these three components. The Technological Pedagogical Content Knowledge (TPCK) model suggests that knowledge about content (C), pedagogy (P) and technology (T) is crucial for effective teaching and learning. However, these three elements should not be viewed independently, and we should emphasize how these three elements relate to and complement each other.

Another important area to consider when we design e-learning classrooms relates to the design of different learning activities for our students. Chi [4] presents a conceptual Active-Constructive-Interactive framework for differentiating passive, active, constructive and interactive learning activities. Chi also proposes a hypothesis that interactive activities ("dialoguing") are most likely better than constructive ones ("generating") ; constructive activities will likely be better than active activities ("manipulating") ; while active activities will likely be better than being passive ("receiving").

3. Process writing / ETS Criterion – an online automated writing evaluation service for secondary school students

3.1 Process writing to improve students' writing skills

Process writing has proven to be a long-lasting and innovative teaching approach since the 1980s [2005]. Traditional approaches to the teaching of writing focus on a teacher-centric model and the written product [13]. As a result, students' writing pieces are rather mechanical, and they also lack the skills needed to do free writing [13]. In addition, struggling students tend to produce writing pieces that are shorter, more poorly organized and weaker in overall quality [9]. They lack the motivation and confidence in writing [3].

Flower and Hayes [7] are among the first group of researchers who studied writing as a cognitive and problem-solving process. They identify that writing can be viewed as a set of unique thinking processes, and these processes have a hierarchical and non-linear structure. In addition, Flower and Hayes indicate that writing is goal-oriented, and that writers explore and refine their goals through the process of writing (planning, translating, reviewing and the monitor). Teachers in Hong Kong were introduced to the process approach to the teaching of writing in the 1990s [3] & [16]. Research studies on teaching process writing to students confirm that this approach can be a workable and effective approach in enhancing students' writing skills [2], [3] & [10]. It also increases students' confidence in writing and a greater awareness of the different stages in writing. Despite the long history of its benefits, the process approach to writing is still not being widely adopted by schools in Hong Kong [5] & [13]. That could have been caused by the lack of teacher training and time constraints on both the students and the teachers. Teachers need to allocate extra lessons to go through the different writing stages and students need the time to do the actual writing.

3.2 Will an online automated writing evaluation service promote process writing?

We are implementing a process writing project with three secondary schools utilizing ETS Criterion. Criterion is a web-based automated writing evaluation ("AWE") service developed by Educational Testing Service (ETS) in the US to evaluate students' writing skills and provide instant score reporting and diagnostic feedback [1]:

Educational goals:	Improve students' subject content, essay writing skills (e.g. cause-and-effect and persuasive essays), collaboration and critical thinking skills
Pedagogy:	Process writing & Peer Reviews
Content:	Language across the curriculum (existing subject content being taught)
Technology:	Criterion online writing evaluation service (with Artificial Intelligence & Natural Language Processing techniques)

3.3 Technological-Pedagogical-Content Model

Teachers create their own writing assignments or select from a list of predefined topics from the online library. Once a student submits his/her essay online, the system provides *instant scoring and feedback on errors* in grammar, usage, mechanics, styles, as well as organization & development. Criterion uses an artificial intelligence system with natural language processing to extract distinct features (over 50 language features) from essays, and to predict (with statistical techniques) what human raters will score a particular writing assignment [1]. *Students will then be able to quickly revise, edit and re-submit/re-publish their essays. There are no preset limits on the number of resubmissions, and the ability for* *the system to provide instant feedback to the students upon their submission will motivate them to write more and learn from the trait feedback analysis and comments from teachers.*

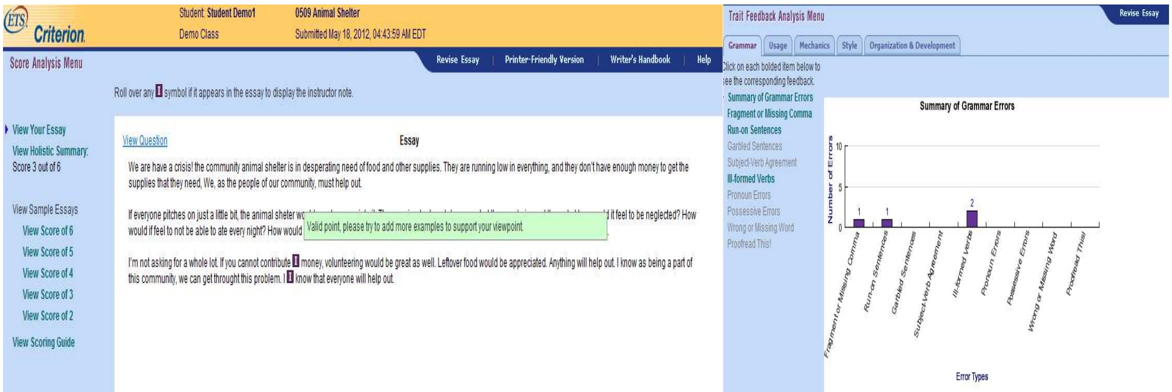


Figure 1: Students receive instant feedback upon submission of their writing

Teachers will also be able to give students feedback online, and the system also supports peer-editing. Teachers will also be able to access different analytic reports online to monitor students' writing progress.

Because the Criterion Service is available online, students and teachers **can access it seamlessly from anywhere with an internet connection at anytime** – school, home, libary or office. The instant feedback and easy-to-use features encourage frequent writing practice, one of the keys to improving writing. Students can also choose from **writing templates/scaffolds** to plan and organize their essays, as well as communicate with their teachers and fellow classmates on peer editig The writing service also enables teachers to spend more time on teaching and focus their instruction in critical areas by reducing time spent assigning, reading and correct student writings.

Attali [1] and Chen & Cheng [3] reported that students benefited from automated writing evaluation systems, for example students had higher writing scores [1]] fewer errors in their resubmissions and increased writing practice. Students could also use the system effectively during the drafting and revising stages of process writing [3].

That is consistent with Wagner's view regarding how this generation is motivated to learn: instant gratification and use of the web for self-directed learning and peer interactions. That would allow teachers to better understand their students' writing ability and thinking process. While it supports all stages of English-language writing, the Criterion service does not grade essay content and cannot take the place of instruction and feedback in a blended learning environment. We plan to study this area in detail (e.g. how we can most effectively incorporate this automated writing evaluation system into the Process Writing curriculum) in the next stage of our study.

4. Conclusion

This paper critically reviews how we incorporate technology in school education to provide students with the critical skills needed for the 21st century. Technological Pedagogical Content Knowledge (TPCK) framework and Active-Constructive-Interactive (ICAP) framework have been reviewed to illustrate the important balance of content, technology and pedagogy in designing and implementing effective e-learning classrooms for our students. The design of an online automated writing evaluation system has been examined to evaluate how it can improve students' English writing and their 21st century skills. This system will be put into practice in the next stage of this study.

References

[1] Attali, Y. (2004, April). Exploring the feedback and revision features of Criterion. Paper presented at the National Council on Measurement in Education (NCME), San Diego, CA.
[2] Chan, K.K.W., & Kong, S.C. (2011, December). Process writing: An online collaborative writing environment for primary school students. Proceedings of the 19th International Conference on Computers in Education: ICCE 2011, Thailand.
[3] Cheung, M. and Chan, A. 1994. Teaching Writing as a Process. Hong Kong: Education Department.
[4] Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science, 1(1), 73-105. doi:10.1111/j.1756-8765.2008.01005.x
[5] Curtis, A., Heron, A. (1998). On being less innovative: Peer groups and process writing in Hong Kong.1(1), Asia Pacific Journal of Language in Education, 99-118.
[6] Debevec, K., Shih, M.Y., & Kashyap, V. (2006). Learning strategies and performance in a technology integrated classroom. Journal of Research on Technology in Education, 38(3), 293-307.
[7] Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communication, 32(4), pp. 365-387.
[8] Friedman, Thomas (2006): The World is Flat: The Globalized World in the Twenty-First Century. London: Penguin.
[9] Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. American Educational Research Journal, 43(2), 295-340.
[10] Ho, B. (2006). Using the process approach to teach writing in 6 Hong Kong primary classrooms. New Horizons in Education, 53, 22.
[11] IBM (2010): IBM 2010 Global CEO Study: Creativity Selected as Most Crucial Factor for Future Success. Available: http://www-03.ibm.com/press/us/en/pressrelease/31670.wss
[12] Kong, S. C. (2007). The development and validation of an information literacy model for hong kong students: Key issues in the professional development of teachers for capacity building. Technology, Pedagogy and Education, 16(1), 57-75. doi:10.1080/14759390601168031
[13] Knight, C., Knight, B., & Teghe, D. (2006, May 31). Releasing the pedagogical power of information and communication technology for learners: A case study. International Journal of Education and Development using ICT [Online], 2(2). Available: http://ijedict.dec.uwi.edu/viewarticle.php?id=167.
[14] Lim, C. P., & Chai, C. S. (2008). Teachers' pedagogical beliefs and their planning and conduct of computer-mediated classroom lessons. British Journal of Educational Technology, 39(5), 807-828. doi:10.1111/j.1467-8535.2007.00774.x
[15] Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A new framework for teacher knowledge. Teachers College Record. 108(6), 1017-1054.
[16] Pennington, M. C., Brock, M. N., & Yue, F. (1996). Explaining Hong Kong students' response to process writing: An exploration of causes and outcomes. Journal of Second Language Writing, 5(3), 227-252.
[17] Wagner, T. (2008) The global achievement gap: Why even our best schools don't teach the new survival skills our children need--and what we can do about it. New York: Basic Books.

The Use of ICT in the Chinese Classroom: A Singapore Perspective

Meng HUANG*, Dongbo ZHANG & Shouhui ZHAO
National Institute of Education, Nanyang Technological University, Singapore
*meng.huang@nie.edu.sg

Abstract: The Chinese language (CL) curriculum for Singapore primary schools proposed in 2004 highlights two pedagogical changes: the modular approach and the integration of ICT. The modular approach puts students into three different groups (or modules, including bridging/reinforcement, core and enrichment) with an aim at catering to students' diverse learning needs. To respond to differentiated learning as initiated in the new curriculum, differences in the type of ICT and the frequency of ICT use in actual teaching practice are expected. This paper endeavors to provide a detailed picture of ICT integration in the Chinese lessons in Singapore primary schools. 207 Primary two Chinese lessons in 20 schools were observed, and teachers' teaching practices were coded by applying the Singapore Chinese Pedagogy Coding Scheme. The preliminary results showed that no significant differences of ICT usage were found across the three modules, which implies that Chinese teachers' use of ICT has not met the needs of differentiated learning as designated in the Modular Curriculum.

Keywords: Primary education, Chinese language teaching, use of ICT

1. Introduction

1.1 Research Purpose

Singapore is a multiethnic and multilingual country with its population comprised of three major ethnic groups, namely, Chinese, Malay and Indian. English, Mandarin, Malay and Tamil are designated as the country's four official languages. Ever since 1969, Singapore has been adopting an English knowing bilingualism policy that requires every Singaporean to master English as well as the language of their own ethnicities, viz., three designated Mother Tongue languages (MTLs): Mandarin for Chinese, Malay for Malay, and Tamil for Indian (Pakir, 1993). To be more specific, English is taught as the students' first language and the main medium of instruction in all schools for all subjects except for the MTL courses and a couple of humanity subjects such as civil and moral education.

As an effect of such bilingual policy, in the Chinese community, an inter-generational language shift from Chinese languages (including Mandarin and other dialects) to English has occurred in familial settings. According to the latest Singapore Census (2010), nearly half of the children entering primary school in that year were reported to use English dominantly at home, as compared to 9.3% in the 1980s. This home language shift has led to greater difficulty in CL learning, and a reform in the CL curriculum has been initialized to help students to adapt and thrive in this globalized world.

In 2004, the Singapore Ministry of Education (MOE) proposed a nation-wide curriculum reform in CL education in primary schools. The new curriculum reform proposed a pedagogical change featuring in differentiated learning. It also recommended the use of information and communication technology (ICT) to enhance students' interest in CL learning and promote the use of CL. In view of such a curriculum reform, this study aims to investigate how ICT is applied in CL teaching in Singapore's primary schools, with special interest to examine whether classroom teachers' use of ICT meets the requirement of differentiated learning.

1.2 The Chinese Curriculum Reform in 2004

The expansion of global markets and the development of new ICTs have brought a series of dramatic changes to education in today's world (Darling-Hammond, 2010). To catch up with such a trend, Singapore also initiated curriculum reforms in various subjects and fields, including the CL curriculum reform that was piloted and launched in 2006. Before 2006, students in primary schools were taught with a relatively standardized curriculum. Within each unit of the textbook, only one text was provided, and the learning focuses were equally emphasized on the four language skills, namely, listening, speaking, reading and writing. The pedagogical focuses were mainly on memorization and exam-oriented repetition (CLCPRC, 2004). The core pedagogical focus of the new curriculum (hereafter, the Modular Curriculum) was to reorient the teaching practice toward a more individually customized and communicative mode. It highlights a number of changes in pedagogies, among which the modular approach and the use of ICT are two main focal points. The flexible "modular approach" is designed to cater to students with varying levels of Mandarin proficiency. Meanwhile, ICTs are promoted so as to enhance students' communication skills (especially, listening and speaking) as well as to increase students' learning interest.

1.2.1 The Modular Approach

The Modular Approach is a pedagogy based on the concept of differentiated learning. It intends to provide customized content for students from various backgrounds and with diverse abilities. For example, different from previous textbooks, each unit in the new textbook series consists of three different modules.

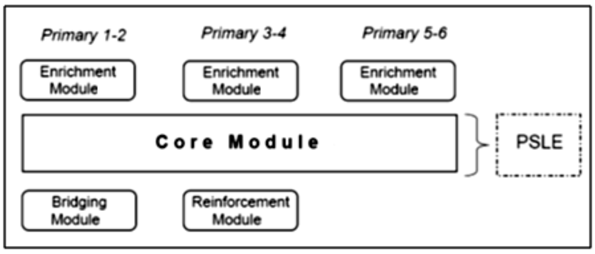


Figure 1. Structure of Chinese Language Modules for Different Learner Profiles (CLCPRC, 2004, p. 9)

Figure 1 shows the structure of module arrangement from Primary 1 (P1) to Primary 6 (P6). All students take the "core module" in each unit which focuses on the knowledge needed for the Primary School Leaving Examination (PSLE). Students who enter schools with little exposure to CL, typical of children from English dominant families, take the "bridging module", which emphasizes on the building of listening and speaking skills that would assist them with picking up the core module so as to catch up with their peers. The "bridging module" in the unit is replaced by "reinforcement module" during Primary 3 (P3) and Primary

4 (P4). For P5 and P6, the Basic Chinese course continue to serve the needs of linguistically weak students. The bridging module and reinforcement module differ in the length of the texts. The bridging module mainly introduces words useful for learning in the core module. The reinforcement module in P3 and P4 includes a passage of a topic similar to, but easier than those in the core module. Additionally, an "enrichment module" in the unit is added on top of the core module for those who have the ability and interest to go beyond the core syllabus.

The curriculum developers believe that through this modular approach, students with different language capacities and home language backgrounds could receive "customized" learning contents. Those who are weak in Mandarin can catch up with the average batch at their own pace, while the stronger ones will further improve with an enrichment module.

1.2.2 The Use of ICT

The use of ICT was greatly promoted in the CL curriculum reform with the hope to enhance students' Chinese learning. More specifically, ICT application in CL pedagogy was to achieve four goals, namely, to ignite students' interest, to facilitate Chinese character writing, to assist in independent learning and assessment, and to increase opportunities to use CL (CLCPRC, 2004). However, how the integration of ICT use can promote differentiated learning was not stated clearly in the Report of the Chinese Language Curriculum and Pedagogy Review Committee (CLCPRC, 2004), the major official document that spearheaded the curriculum renovation. In other words, there were only general concepts and visions without detailed guidelines and procedures for teachers to use ICT to enhance their teaching.

1.3 ICT and Chinese Language Education

Buang (2011) introduced the "10'C Programme", a Web 2.0 based programme that enabled students' independent learning and peer interaction in learning MTLs. The preliminary evaluation shows a "definite advantage of the new approach (ICT approach)" and the "new ICT-based pedagogy for the learning and teaching of MTLs is contributing to the fostering of learner autonomy in language learning in primary schools in Singapore" (Buang, 2011, p. 239). Lim and Tay (2003) reported how different types of ICT tools (informative, situating, constructive, and communicative tools) were used to engage Singaporean students in higher-order thinking. They pointed out that different types of ICT tools were often used to complement one another to achieve the teaching objectives. Moreover, the classification of an ICT tool is based more on how it is used than its characteristics, and effective management of digital instructional resources tends to ensure seamless and easy retrieval and supports the integration of ICT into the curriculum.

Since the use of ICT and differentiated learning are two major components of the modular curriculum, it is our focus to see how the use of ICT could be integrated into differentiated learning, and whether CL teachers' actual teaching practices would show any module-wise variation with respect of the use of ICT. To be more specific, since the aim of the modular curriculum is to provide learning contents tailored up for students with diverse proficiency in CL, we expected that teachers would adopt customized pedagogies for students in different modules. Thus, the use of ICT in different modules was expected to be different in terms of both type of ICT tools and frequency of their use.

2. Research Methods

2.1 Participants

This study adopted the method of class observation to investigate the use of ICT in different modules in Singapore primary schools. The video recordings were obtained from a large-scale ongoing research project at the Centre of Research in Pedagogy and Practice, National Institute of Education, Singapore. In that project, three classes of different modules (bridging, core, and enrichment) in twenty primary schools' P2 cohort were randomly sampled to be observed. For each class, one week's worth of lessons were systematically observed and recorded because one unit should be taught within one week according to MOE's curriculum. Each lesson lasted from thirty minutes to 1.5 hours. In total, 207 lessons were observed, including 56 bridging lessons, 68 core lessons, and 74 enrichment lessons.

During the observations, the researchers used the Singapore Chinese Pedagogy Coding Scheme (SCPCS) to code each lesson. This coding scheme was developed on the Singapore Pedagogy Coding Scheme (Luke, Freebody, Shun, & Gopinathan, 2005). The SCPCS was modified to facilitate classroom observation, help capture the features of teachers' pedagogical practices, and examine the similarities and differences among different modules. Each lesson observed was divided into several phases according to teachers' teaching activities. Each phase lasted at least three minutes, and within each phase, the pedagogical focus, the teaching tools, code switching, the learning tools students used, and students' engagement were coded according to the coding manual. In the present study, we focused only on the data of teachers' teaching tools. We recorded exactly whether the teacher had used any tools and what kind of tools the teacher adopted.

Since lessons were observed across three modules, it was expected that teachers' use of tools, especially ICT tools, would vary from module to module. To be more specific, students in the three modules would be taught with diverse pedagogical approaches since they possessed different CL capacities, and miscellaneous tools would be applied so that the aim of differentiated learning could be achieved. In order to investigate the differences of the use of ICT among the three modules, one-way between subjects ANOVA was conducted to compare the effect of module type on various conditions.

3. Preliminary Results

Since the study mainly focused on teachers' use of ICT, this paper only provides the results of the following categories, namely, Teacher's Tool_Nil (no ICT usage), Teacher's Tool _ICT_total (including Teacher's Tool_Traditional ICT and Teacher's Tool_New ICT), and Teacher's Tool_Traditional (teacher used traditional tools such as textbooks, whiteboard, worksheets, etc.). The use of traditional ICT such as PowerPoint, Audios and Videos was coded within the sub-category of Teacher's Tool_Traditional ICT, whereas the use of new ICT such as one-to-one computer, interactive whiteboard and multi-media involving interaction was coded under the sub-category of Teacher's Tool_New ICT. This paper examines the use of ICT across the three different modules (bridging, core, and enrichment).

One-way ANOVA results showed that for Teacher's Tool_NiL, there was a significant difference between Core and Bridging modules and between Enrichment and Bridging modules, $F(2, 195) = 6.577$, $p = .002$. However, no significant differences were found across modules, $F(2, 195) = 1.057$, $p = .350$ for Teacher's Tool _Traditional ICT, $F(2, 195) = 2.796$,

p = .064 for Teacher's Tool_New ICT, $F(2, 195) = 1.057$, p = .349 for Teacher's Tool_ICT_total, and $F(2, 195) = .085$, p = .918 Teacher's Tool_Traditional.

4. Conclusion and Discussion

The above result indicated a very intriguing finding, that is, how to integrate the use of ICT into differentiated learning. Although the adoption of the modular approach and ICT attempted to band the weak with the average through differentiated pedagogies, the expectations did not seem to be well met when the intention of the curriculum reform was translated and carried out into practices. The result implies that teachers' use of ICT in Singapore CL classrooms has not met the needs of differentiated learning. CL teachers' pedagogical practices in general, and ICT use in particular, remained the same even though they were conducting lessons to students with different CL proficiency. Therefore, in order to fulfill the purpose of differentiated learning, teacher training should consider reinforcing the integration of ICT in classroom pedagogies to accommodate learners' diverse learning needs.

Acknowledgements

The data used in this paper is derived from the project "Effectiveness of the Chinese Modular Curriculum in Singapore Primary Schools: An Evaluative Study" funded to the third author by the Office of Education Research, National Institute of Education, Singapore. Much appreciation goes to Mr. Wang Chu Meng and Miss Janice Lim Xiaoyuan.

References

[1] Buang, A. B. (2011). ICT-Based teacher-facilitated and self-directed learning for Mother Tongue Languages. In D. Gardner (Ed.), *Fostering autonomy in language learning* (pp. 233-239). Gaziantep: Zirve University.

[2] CLCPRC. (2004). *Report of the Chinese Language Curriculum and Pedagogy Review Committee*. Singapore.

[3] Darling-Hammond, L. (2010). *The flat world and education*. New York: Teachers College Press.

[4] Department of Statistics. (2010). *Singapore census of population 2010: Advance census release*. Retrieved from http://www.singstat.gov.sg/pubn/census2010.html.

[5] Lim, C. P., & Tay, L. Y. (2003). Information and communication technologies (ICT) in an elementary school: Students' engagement in higher order thinking. *Journal of Educational Multimedia and Hypermedia, 12*(4), 425-451.

[6] Luke, A., Freebody, P., Shun, L., & Gopinathan, S. (2005). Towards research-based innovation and reform: Singapore schooling in transition. *Asia Pacific Journal of Education, 25*(1), 5-28.

[7] Pakir, A. (1993). Two tongue tied: bilingualism in Singapore. *Journal of Multilingual and Multicultural Development, 14*(1/2), 73-90.

Analyzing ICT literacy of German Teachers: Focusing age and gender issues as well as identifying knowledge levels

Johannes ZYLKA
Center for Research on Educational Quality and Evaluation,
Centre for International Student Assessment (ZIB),
German Institute for International Educational Research (DIPF),
Frankfurt, Germany
zylka@dipf.de

Abstract: Owing to the increasingly fast development of media and especially the advancement of computer technology in recent decades, the topic of ICT integration in classrooms is being intensively discussed on different levels and media educators argue about the missing link between school education and the use of ICT in schools. In this respect, ICT-related competencies of teachers are supposed to play an important role, but approaches assessing ICT-related competencies – in general as well as in terms of the teacher profession – rarely exist and, thus, extensive data on ICT literacy of teachers is missing. The presented article focuses on identifying ICT knowledge of German teachers and teacher trainees, taken as a central aspect of media and ICT literacy, and in this respect especially addresses age-related and gender issues as well as different levels of ICT knowledge.

Keywords: ICT literacy, Media literacy, ICT knowledge, teaching profession, gender issues

Introduction

Schools worldwide are confronted with a variety of requirements related to digital media, for instance when looking at current developments such as mobile devices, cloud computing, collaborative working, or Game-based Learning [1]. Taking into account these developments, it is evident that the lacking integration of media literacy education in formal education has reached a new level of significance and will continue to gain importance [2].Thus, for teacher professionals in the 21st century, the ability to handle digital media in different forms and scenarios is a central part of their every-day life [3].

Searching for studies focusing media or ICT literacy, it gets clear that this is an area of research which has not been extensively analyzed so far. There are at least some approaches discussing media use or media related attitudes of teacher professions [4], analyzing computer availability of teachers at home [5], or discussing the necessity of media literacy education on a theoretical level [6], but a minor amount of studies focuses ICT literacy of teacher professionals [7], especially accounting for different groups of teacher professionals: student teachers (initial training), teacher trainees (advanced training) as well as experienced teachers. Therefore, the presented article will analyze teachers' knowledge on ICT as predictor of ICT competence and ICT literacy, enabling conclusions on the readiness of teachers for the use of ICT in their profession. This article discusses ICT-related competencies of teachers in context of ICT integration in schools.

1. Integration of ICT in schools

In terms of school contexts, desktop or laptop computers still represent the most frequently used devices. The PISA studies showed that nowadays most students are quite well equipped with computers at home and that the school-related computer-per-student ratio has significantly improved comparing the situations between 2000 and 2009 [8]. Besides ICT availability at school, it is also important that teachers frequently use the available media, especially in classroom contexts. Thus, as important as data on the school-related availability of ICT is, it is not sufficient to allow statements on the implementation of ICT in schools. Studies have shown that the use of ICT in schools by teachers as well as by students differs a lot from the given possibilities. Based on these data, it can be shown that teachers in those countries where schools are equipped with computers are not necessarily using them in classroom contexts or other school-related activities. While in some ICT-equipped countries there is a relatively intensive use of computers in school contexts, the situation in other countries such as Germany differs: The level of computer availability is at least sufficient, but the use of ICT in school contexts differs strongly from the given possibilities. This is – amongst others – leading to a missing integration of ICT in learning scenarios as well as a lack of media literacy education of students [9].

In this context, discussions lately noticed that ICT-related competencies of teachers are quite important in terms of this gap and, therefore, in context of using ICT potentials in school [10]. As mentioned above, researchers have mostly focused on the availability and use of ICT in schools or on the curricular grounding of media literacy education in the last decade. Searching for literature on teachers' media and ICT competencies, it seems clear that this topic has not yet been an area of major research interest [11].

2. Discussing Media Literacy of teachers

Extensive data on media and especially ICT-related competencies in terms of teacher profession are rare. There is a variety of reasons for the lack of research on this topic, for instance the resistance of formal education against change and especially digital media or the task of finding teachers who permit studies on their media literacy. Even if data is provided, it often does not meet the requirements of objectivity, reliability and validity. In the past, studies have mostly concluded experience-based that teachers' lack media education and media-related competencies [12] or have focused on ICT availability and use. Hence, the most important challenge might be that the measurement of media competencies or media literacy has only found little attention in the past. As a consequence, there are – for the best of the author's knowledge – no instruments for adequately measuring media literacy in a quantitative manner, especially talking about the teacher profession.

2.1 Accessing ICT Competencies of Teachers

Speaking of competencies in general, there is quite an extensive discourse on how these can be measured. For instance, large scale studies such as PISA or TIMSS assume that knowledge is the most essential basis of competencies. In this respect, declarative as well as procedural knowledge can be seen as a predictor of competency. While the level of knowledge is quite important from a theoretical perspective, empirical approaches to media competencies of teachers in the European area have so far not focused on the level of knowledge [13]. Therefore, a decision was made to develop a new instrument targeting ICT knowledge, applicable in the given context. A first version consisted of nearly 400 multiple choice items, comprehending extensive knowledge on most fields of ICT. Because the instrument was developed on the basis of the ICT standard work of [14], it can be seen as theoretically approved. After development, the questions were presented to experts (originating from the teacher profession as well as from the field of media research) rating the importance and the adequacy of the questions on a 5-point Likert scale. Finally, the experts agreed on the high importance and good adequacy of 39 items.

2.2 Description of the study

Based on this newly developed questionnaire, which is named ITK.basic in its German version, we conducted a study in the German federal state of Baden-Wuerttemberg, which focused on teachers and teacher trainees, in total including N=393 test persons from the different levels of teacher education. The smallest study group is the group of teachers (*T*) with exactly N=100 test persons. Teacher trainees are differentiated by student teachers (*ST*), who are studying teaching at university (initial training), and teacher trainees (*TT*) in the second phase of their vocational training, as is common in German teacher education. The study's results are by no means representative for the federal state of Baden-Wuerttemberg or the country of Germany. The study was paper-based as well as computer-based with an internet questionnaire, the assumption being that teachers as well as teacher students of the second phase would hardly participate in the survey if their questionnaires were not delivered paper-based. Thus, inquiry of *TS* was executed web-based, of *TT* both web- and paper-based, and of *T* just using a paper-based version of the ITK.basic. Since participation in the questionnaire (for the groups *TS* and partially *TT*) was voluntary, it is supposed that students with – at least by trend – higher ICT knowledge participated in the study. Thus, the empirical level of ICT knowledge is probably even overestimated in the presented results.

However, the overall reliability of the questionnaire resulted quite well with Cronbachs' α=0.914. A good validity can be assumed because of the expert rating, and objectivity was achieved due to the study method which was based on a questionnaire. Overall, no test person was able to solve all the questions of the questionnaire. The one student with the lowest percentage of correct answers scored just about 5% correct answers, those with the highest knowledge levels around 95% correct answers. The most difficult items were solved by only 13 persons (3%), while the easiest question was solved by all of the probands.

2.3 Central findings from the study

One of the recently discussed aspects related to ICT competencies – not only speaking of the teacher profession – is age-dependence. It is mostly assumed that older persons do not know very well how to use computers and especially the latest ICT devices, which is a matter of concern because of the given age ratio in terms of the school systems in most developed countries. Thus, as first focal point of this study it is supposed in the presented data that older subjects score significantly less than younger ones. In addition, some authors argue that students are just partially media literate and that students choosing to become teachers mostly are not media literate. As teachers are expected to teach students ICT literacy as well as use ICT in their teaching activities, their own ICT knowledge becomes an essential predictor not only for their ICT literacy but also for their teaching competencies. Coming to the presented study: While there is no significant correlation of the subjects' age with their ICT knowledge (r=-.099, p=.07, N=334), the heterogeneity from 20 years to beyond 60 years is immense. There were students who solved over 95% of the questions as well as students solving below 20%. Therefore, data suggests that groups of teachers with significant gaps in media knowledge can be found at all ages. In terms of data analyses, there are four groups that can be differentiated amongst the test persons. The following differentiation of groups should be understood as approximating in terms of practicable visualization of data – barriers between different groups do not represent fixed levels.

First, there is the group of ***expert users*** with profound ICT knowledge that solved more than 80% of the items in the questionnaire. The members of this group can be characterized by extensive knowledge in most areas of ICT. It therefore is assumed that subjects belonging to this group are well equipped with knowledge to use ICT in a self-determined way and are also able to use ICT in different pedagogical scenarios. It is also supposed that these test persons are somehow engaged and interested in ICT. 59 out of 393 test persons belong to this group, 56% of which are student teachers, 17% teacher trainees and 27% teachers.

Members of the second group, which is referred to as ***standard users***, scoring between 50% and 80 %, are supposed to have some ICT-related knowledge in basic ICT topics necessary in every-day ICT use. Subjects belonging to this group are supposed to be mostly standard users, being equipped with a minimum of ICT-related knowledge. This group might be seen as target audience for on-the-job trainings in media education, because they have a good basis for further media-related skills and abilities.171 test persons (43%) scored between 50% and 80% correct answers, 61 (35%) of which are student teachers, 59 (35%) teacher trainees, and 51 (30%) teachers.

Reaching between 25% and 50% of correct answers, the third group is labeled ***low affinity users:*** Members use ICT as a tool if necessary. They generally have little motivation to deal with ICT and minor expertise on standard tasks, but might be a good audience for basic professional training. 134 subjects (34%) – and therefore, the second largest group in this test – supposedly belong to the group of low affinity users. 56 (42%) of these are teacher students, 54 (40%) teacher trainees and 24 (18%) teachers.

Test persons of the fourth and last group solved below a quarter of items and, therefore, demonstrated a significant lack of substantial ICT knowledge in all areas of ICT. They should be strongly advised to attend extensive media-related training. Training on-the-job will mostly not suffice for these test persons, because they lack a basic understanding of ICT. 28 probands (9%) can be defined as ***no affinity users***, 25% of whom are teacher students, 43% teacher trainees, and 32% teachers. The fact that this group does not only consist of older teachers but also of younger teachers aged 20 to 30 shows that today's students are not necessarily ICT literate.

Analyzing the group with quite low ICT knowledge it is clear that mostly female students are included here: While there are19 (82%) female students in this group, just 4 (17%) male students possess minor ICT knowledge. Amongst the group of 59 students which scored above 80%, there are 70% males and 30% females. These aspects lead to the discussion of general gender-specific issues in terms of this article, which – in context of discussing media and ICT literacy – is an aspect of major importance. Gender issues are generally considered to be important in terms of the formal educational sector because of the given gender ratio in primary and secondary education. Thus, presented data showed highly significant differences between male and female probands in all of the groups: For student teachers ($r=.550$, $p=.000$, $N=157$), teacher trainees ($r=.382$, $p=.000$, $N=133$) as well as teachers ($r=.395$, $p=.000$, $N=98$). Gender differences are also significant in all included topics, the strongest differences are recognizable in the knowledge field of *data bases* ($M_m=.66$, $SD_m=.33$, $N_m=98$; $M_f=.36$, $SD_f=.28$, $N_f=240$) as well as *Software* ($M_m=.76$, $SD_m=.32$, $N_m=98$; $M_f=.48$, $SD_f=.29$, $N_f=240$). The least – but still highly significant differences – between males and females are extracted from the field of *basic knowledge* ($M_m=.79$, $SD_m=.20$, $N_m=98$; $M_f=.67$, $SD_f=.21$, $N_f=240$) as well as from *internet and web* ($M_m=.70$, $SD_m=.22$, $N_m=98$; $M_f=.52$, $SD_f=.21$, $N_f=240$).

3. Conclusion and Future Work

Bearing in mind that the ITK.basic was rated by experts to just ask for basic knowledge on ICT, the results mentioned here show that most teachers, teacher trainees as well as student teachers do not possess extensive knowledge in ICT. Considering the statement that presently there are major lacks in terms of measuring media literacy of teachers, the presented study adds to international educational research. It gives additional information on the addressed topics of gender issues. The focal point that older teachers generally are – at least compared to younger colleagues – not that media and ICT literate could not be proven in terms of this study. There are teachers older than 60 years who know lot more about ICT than some students in their early 20s. The instrument that was developed in terms of this study presents one step towards the development of an instrument which can empirically measure ICT competencies (not only in terms of teacher education) in a quantitative manner. Nevertheless, tasks such as the development of instruments focusing other areas of ICT literacy, remain.

References

[1] Johnson, L.; Levine, A.; Smith, R. and Stone, S. (2011). The 2011 Horizon Report. Austin, Texas: The New Media Consortium

[2] Frau-Meigs, D. & Torrent, J. (2009). Mapping Media Education Policies in the World. New York: United Nations.

[3] Hague, C. & Payton, S. (2010). Digital Literacy across the Curriculum. A Futurelab Handbook. Bristol: Futurelab; Johnson, L.; Levine, A.; Smith, R. und Stone, S. (2011). *The 2011 Horizon Report*. Austin, Texas: The New Media Consortium;

[4] Vrasidas, C.; Pattis, I.; Panaou, P.; Antonaki, M.; Aravi, C.; Avraamidou, L.; Theodoridou, K. & Zembylas, M. (2010). Teacher Use of ICT: Challenges and Opportunities. Proc. 7th International Conference on Networked Learning 2010.

[5] MPFS (Medienpädagogischer Forschungsverbund Südwest)(2003). Lehrer/-innen und Medien 2003. Nutzung, Einstellungen, Perspektiven (*Teachers and Media 2003: Use, Attitudes and Perspectives*). Retrieved from http://www.mpfs.de/fileadmin/Einzelstudien/Lehrerbefragung.pdf on June 3rd, 2012.

[6] Frau-Meigs, D. & Torrent, J. (2009). Mapping Media Education Policies in the World. New York: United Nations.

[7] Petko, D. (2012). Hemmende und förderliche Faktoren des Einsatzes digitaler Medien *[Positive and negative aspects of the use of digital media]*. In: Schulz-Zander, R.; Eickelmann, B.; Moser, H.; Niesyto, H. and Grell, P. (Hrsg.). Jahrbuch Medienpädagogik 9. Wiesbaden: Springer VS. pp.38.

[8] OECD (2011). PISA 2009 Results: Students On Line. Digital Technologies and Performance (Volume VI), p.146. Online: http://dx.doi.org/10.1787/9789264112995-en

[9] Hague, C. & Payton, S. (2010). Digital Literacy across the Curriculum. A Futurelab Handbook. Bristol: Futurelab.

[10] Buckingham, D. & Domaille, K. (2009). Making Media Education Happen: A Global View. In: Cheung, Ch.-K. (ed.). Media Education in Asia. Heidelberg u.a.: Springer. pp.19; Camps, V. (2009). Media Education beyond School. In: Frau-Meigs, Torrent (Hg.): Mapping media education policies in the world. Vision, Programmes and Challenges. New York: United Nations, pp.189;

[11] Petko, D. (2012). Hemmende und förderliche Faktoren des Einsatzes digitaler Medien *[Positive and negative aspects of the use of digital media]*. In: Schulz-Zander, R.; Eickelmann, B.; Moser, H.; Niesyto, H. und Grell, P. (Hrsg.). Jahrbuch Medienpädagogik 9. Wiesbaden: Springer VS. pp.38.

[12] Petko, D. (2012). Hemmende und förderliche Faktoren des Einsatzes digitaler Medien. *[Positive and negative aspects of the use of digital media]*. In: Schulz-Zander, R.; Eickelmann, B.; Moser, H.; Niesyto, H. und Grell, P. (Hrsg.). Jahrbuch Medienpädagogik 9. Wiesbaden: Springer VS. pp.38.

[13] Zylka, J. (2012). Medienkompetenzen und Instrumente zu ihrer Messung. Entwicklung eines Wissenstests zu informationstechnischem Wissen von Lehrkräften, Lehramtsanwärtern und Lehramtsstudierenden. *[Media competencies and Instruments for their measurement: Development of a knowledge test on IT knowledge of student teachers, teacher trainees and teachers]*Unpublished PhD-Thesis: University of Education Weingarten.

[14] Parsons, J.J. & Oja, D. (2010). New Perspectives on Computer Concepts 2011. Boston: Course Technology/Cengage Learning.

Investigating how to design interactive learning environments to support students' learning of upper secondary and university math

Sanela MEHANOVIC[a*], Daniel SPIKOL[b]
[a]*Department of Mathematical Sciences, University of Gothenburg, Sweden*
[b]*Department of Computer Science, Malmö University, Sweden*
*sanela.mehanovic@chalmers.se

Abstract: Students' difficulties in learning of mathematics have for a long time, been investigated by researchers in different fields. Within educational research there are claims that technological tools appropriately integrated in students' mathematical work can support their understanding of a wide range of concepts in mathematics. This paper reports on the initial investigation for the design of Interactive Learning Environments (ILE) to support students' learning of mathematics. The project is guided by the notion of Design Based Research (DBR) and aims to explore how to design ILE that support students' understanding of integrals in particular. The initial study was conducted at a Swedish university with 10 students in 4 groups. The study confirmed difficulties in students' understanding of integrals as reported in educational literature and provides a set of design aims for the next iteration of the ILE to support the learning.

Keywords: mathematics, education, students, interactive learning environments

1. Introduction

Society needs a well-educated population, who not only actively contributes to the shaping of the society itself, but who also, as a broadly qualified work force is able to activate and transfer school content knowledge, insights, and skills to a variety of situations and contexts. Mathematics, from a societal perspective is recognized as one of the key components in this process, has lately met considerable difficulties. Schools and universities across the world meet with an increasing problem with young people having difficulties in dealing with mathematical content. The use of computers in mathematics education has often been an underlying goal of presenting mathematical concepts to students in a new and dynamic way compared to previous learning environments. Some mathematical concepts are difficult for students to understand when presented in the paper/pencil based teaching lend themselves to computer representations as in the case with the integral concept [1]. Integrals have visual aspects that can be displayed on a computer screen along the other representations such as algorithmic, symbolic, numerical, or natural language representations.

With the use of mathematical software for visualization, the notion of integrals is more easily adopted by students [2]. On the other hand, it also makes the didactical situation more complex [3]. A technological tool that becomes a mathematic work tool in the hands of the students is a process that has turned up unexpectedly complex [4]. The process causes differentiation in students' work with technological tools, meaning that different students have different experiences and work differently with the same tool and within the same environment. Furthermore, the work of Guin and Trouche [5] argues the more complex environment the larger the differentiation of students' work methods with these applications can result in more diverse learning trajectories.

1.1 Research Aim

This paper reports on the initial investigation in the design of an interactive learning environment (ILE) to support students' understanding of integrals. The project is guided by the notions of design-based research (DBR). In education DBR is used to develop and investigate (content oriented) theories though iterative cycles of intervention and refinement. DBR aims to combine the intentional design of interactive learning environments with the empirical exploration of our understanding of these environments and how they interact with the individuals [6].

The research aim is to explore how to design ILE that support understanding integrals in particular. In our opinion, a design of ILE should consider following two aspects: Firstly, it should attempt to minimize issues related to students' difficulties to optimally use technological tools in their mathematical work. Secondly, the design should, in parallel, deal with difficulties in students' understanding of a particular mathematical content.

2. Background

In the upper secondary education, the integral is generally defined in the following way (see figure 1): Let f(x) be a continuous function in a closed interval [a, b] divided in subintervals with equal length □x. Then, for n subintervals we have the area $S \sim \sum_{i=1}^{n} f(xi) \vartriangle xi$. If we let $n \to \infty$ then □$xi \to 0$ and it can be shown that $\sum_{i=1}^{n} f(xi) \vartriangle xi$ approaches a limit called the integral of f from a to b, which is denoted $\int_a^b f(x)dx$.

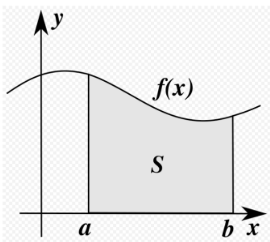


Figure 1: Integral defined as an area under the function f(x).

This definition based on a Riemann sum is difficult for students to understand. Students' difficulties with integrals are not a new behavior in the mathematics classroom and had, for the last several decades have been a subject for educational research [1,7,8]. In the early eighties, Orton observed that students had difficulties while solving tasks related to the understanding of integration as limit of sums [9]. In this study, students were able to apply, with some facility, the basic techniques of integration but further probing indicated fundamental misunderstanding about the underlying concepts. Students interpreted the integral as a procedure that transforms an input into some output. The same study revealed that students' technical ability could be quite strong, despite showing minimal conceptual understanding. Apart from showing strong procedural skills the students were found to demonstrate a strong reluctance to using geometric interpretations to complete an algebraic process, and when possible, were more inclined to move to an algebraic context [8].

Another study from Orton (1980) revealed that students had problems with the integral $\int_a^b f(x)dx$ if f(x) is negative or b is less than a [9].

More recent studies specialized in mathematics education show that this concept is still difficult for students' to grasp; they are not able to write meaningfully about the definition of a definite integral nor can they without difficulties interpret problems calculating areas and definite integrals in wider contexts [10]. The students also intend to identify the definite integral as an area [11].

2.1 Theoretical Framing

Students' understanding of integrals can be discussed from the perspective of the cognitive structure in their mind that is associated with the concept of integrals. Tall and Vinner [12] formulate a distinction between the mathematical concept as formally defined and the cognitive processes by which they are conceived by the students (p.1). The total cognitive structure that is associated with the concept, including all the mental pictures and associated processes, they name a concept image, and mean that a student's image of a mathematical concept may not be globally coherent and may have aspects which are quite different from its formal mathematical definition" (p.1). At different times, seemingly different conflicting images may be activated. The conflicting aspects, that are a part of a student's concept image and/or a concept definition, are called cognitive conflict factors (p.3). As a student does not necessarily see a conflict while using different methods in their mathematical work, the student simply utilizes the method he or she considers appropriate on each occasion [12]. The conflicting aspects that are a part of a student's concept image and/or a concept definition are called cognitive conflict factors (p.3). Only when conflicting aspects are evoked simultaneously need there be any actual sense of conflict or confusion (p.2).

3. Methods

Our initial study was conducted to investigate students' concept image of integrals, here in terms of the definite integral, in a way the concept is usually introduced to students (see figure 1). The study was conducted at a Swedish university and considered an introductory course in mathematics with 10 students. Four groups were self-created with 2 to 3 students for the intervention. The participating students, who were just about to finish their introductory course already had an image of the integral concept, were asked to take a test containing integral tasks based on a previous research conducted by Rolka & Rösken [13]. This test was developed in order to investigate students' understanding of the formal mathematical definition of the definite integral, and focused on aspects in integrals known to be difficult for students to grasp.

In the intervention, the students were supposed to within an hour, solve a test with eight integral tasks. They were asked to solve tasks in the test as a group, and to write their solution on a whiteboard while discussing a particular task. Their work was videotaped, and, once they agreed on a solution to a particular task, we took a photograph of their whiteboard notices. We are currently working on the analysis the video data and the solutions gathered from observing the groups of students.

4. Results

The initial study confirmed previous research within mathematics education. While solving the integral tasks, students have not always been aware of their conflicting images of the definite integral in relation to its definition. For instance, one of the images that has been shown to be highly present in the experiment group was the perception of the definite integral as an area, see figure 2.

In the task shown in Figure 2 (left image) students' were asked to calculate the value of the integral. The task considers the oriented area aspect of the integral and the results confirmed Rolka & Rösken's finding in which many students just equals the concepts of integral and area [13]. None of the four groups in our initial study came with a correct answer to this task. Three groups have chosen the option A as the final answer to the task. Only students in the remaining group considered that the correct answer might be another option than A. Indeed, they did suggest the correct option (B) although it has been suggested as a second alternative (even this group had A as their first choice).

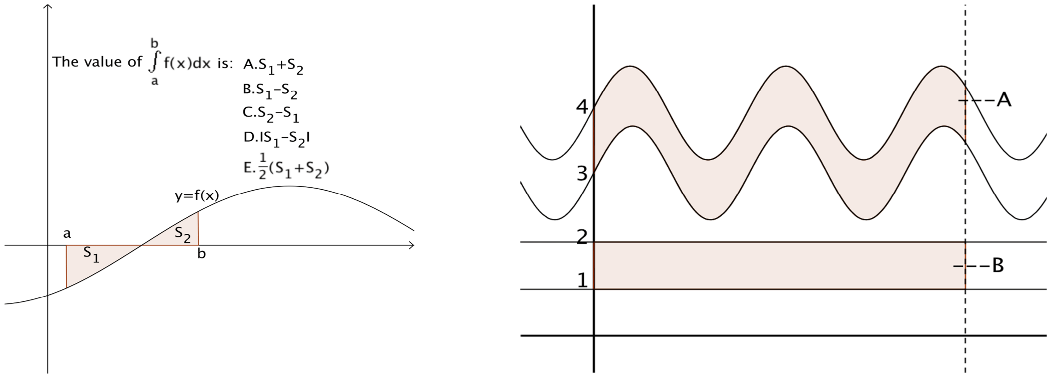


Figure 2: Examples of the students' perceptions of integrals.

Another task that all participating groups had difficulty with, was to deal with the problem illustrated in figure 2 (right side). The picture shows two areas A and B. What do you think is correct for the relation between the areas?

- The area of A is bigger than the one of B.
- The area of A is smaller than the one of B.
- Both areas are equal.
- Without any function given explicitly, it is not possible to answer this question

None of the groups had answers that they were certain about, rather, they were discussing different options having difficulties in choosing between the first and the third option. What was of a particular interest for the study was the contrast between these two options that seemed to cause a cognitive conflict for some of the students. The discussed aspect was following: The first option feels true instinctively, if one thinks that the area of A can be stretched outside of the interval (still keeping the same height). On the other hand, in the definition of the definite integral as a Riemann, the area between two curves is calculated as a sum of areas of infinitely thin rectangles. Only one student started to discuss the Riemann sum which led the whole group to move their reasoning to what answer option could be appropriate from the formal definition's perspective. In our opinion, this example demonstrates how conflicting images evoked simultaneously in students' work with a mathematical concept can lead them to a deeper reasoning of the meaning of the formal concept definition.

5. Discussion

The results of our initial exploration point to design goals for the ILE that include further investigations of the role of the technological tools for students' mathematical work with integrals. Misconceptions about the mathematical problem observed in the student groups point to the need for support. From our investigation of students' understanding of integrals and from the supporting research literature we are able to identify some implications for the next iteration of our ILE. In order to support the students to expose their conflicting images of integrals and gain a deeper understanding of this mathematical concept the following design goals have been identified:

- The ILE needs to guide the students though a learning process that exposes their concept image of integrals and then supports its' development, while taking into consideration a diversion in students' individual perceptions.
- This guiding process needs to provide support for individual learning exploration for the student though some types of externalization like adaptive scaffolding and teachable agents [14].
- Minimize issues related to students' difficulties to optimally use technological tools in their mathematical work by acting as a support component for learning instead of providing a wealth of features that can be seen at times to distract lower performing students [4].

For the next steps, we are currently designing a low-fidelity prototype that will explore teachable agent like qualities that we can test during a summer school mathematics class for university students. These students will form a new test group for the next round of participatory workshops with the teachers and researchers.

References

[1] Tall, D. (1991). Visualizing Differentials in Integration to Picture the Fundamental Theorem of Calculus. *Mathematics Teaching*, 137, 29-32.

[2] Herceg, D. & Herceg D. (2009). The Definite Integral and Computers. *The Teaching of Mathematics*, 12 (1), 33-44.

[3] Blomhøj, M. (2005). Villkor för lärande i en datorbaserad matematikundervisning - elevernas användning av avancerade matematikprogram. In B. Grevholm (Ed), *Matematikdidaktik - ett nordiskt perspektiv*, (pp. 185-218). Lund: Studentlitteratur.

[4] Artique, M.(2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, 7(3), 245-274.

[5] Guin, D. & Trouche, L. (2002). Mastering by the teacher of the instrumental genesis in *CAS environments: necessity of instrumental orchestration*. ZDM 2002 34 (5).

[6] Hoadley, C. M. (2004). Methodological alignment in design-based research. *Educational Psychologist*, 39(4), 203–212. Routledge.

[7] Orton, A. (1980). An investigation into the understanding of elementary calculus in adolescents and young adults. *Cognitive Development Research in Science and Mathematics*, 201-215. University of Leads.

[8] Machin, M.C. & Rivero, R.D. (2003). Using Derive to understand the concept of definite integral. *International Journal for Mathematics Teaching and Learning*, 1-16. ISSN 1 473-0111.

[9] Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics*, 14, 1-18. Dordrecht, Holland, and Boston, U.S.A.

[10] Rasslan, S. & Tall, D. (2002). Definitions and images for the definite integral concept. In Cockburn, A. D. Cockburn Nardi, E. (Eds), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education*, (Norwich, UK), 4, 89-96.

[11] Attorps, I., Björk, K. & Radic, M. (2011). The use of mathematics software in university mathematics teaching. Paperpresented at CERME 7, T*he Seventh Congress of the European Society for Research in Mathematics Education*, 9-13th February 2011. Rzeszow, Poland.

[12] Tall, D. & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with particular reference to Limits and Continuity. *Educational Studies in Mathematics*, 12, 151-169.

[13] Rösken, B. & Rolka, K. (2007). Integrating Intuition: the Role of Concept Image and Concept Definition for Students' Learning of Integral Calculus. *The Montana Mathematics Enthusiast*, ISSN 1551-3440, Monograph 3, 181-204. The Montana Council of Teachers of Mathematics.

[14] Blair, K., Schwartz, D., Biswas, G. and Leelawong, K.(2011). Pedagogical Agents for Learning by Teaching: Teachable Agents. To appear in *Educational Technology*.

Exploring pre-service teachers' perceptions of and participation in online communities

Liping DENG[a*], Nicole TAVARES[b]
[a]*Hong Kong Baptist University, Hong Kong*
[b]*The University of Hong Kong, Hong Kong*
*liping@graduate.hku.hk

Abstract: The study focuses on the participation of pre-service teachers in a course website set up on Moodle and a website developed by students on their own initiative to support their Teaching Practice. It explores PSTs' experiences and perceptions of the two different types of websites and how their online participation linked with the sense of community. Utilizing questionnaire as the main source of data, the paper reveals that the student teachers had more positive perceptions of the self-initiated website which was regarded as a useful platform for exchanging teaching ideas, sharing resources, gaining support, and maintaining communication with their fellow coursemates. The frequency of reading the self-initiated website was also correlated with the sense of community.

Keywords: online community, perceptions, pre-service teacher

1. Introduction

The study focuses on the participation of pre-service teachers in a course website set up on Moodle and a website developed by students on their own initiative to support their Teaching Practice (TP). We seek to understand how students participated in and perceived the teacher-initiated and self-initiated website and how their perceptions and participation are associated with their sense of community. Adopting the case study methodology, the study investigates the experience of 31 final-year student-teachers in the Faculty of Education at an university in Hong Kong. The particular questions that guide the inquiry are: (1) How do student-teachers perceive the course website and the self-developed website for TP? (2) How are student-teachers' perceptions associated with their sense of community?

2. Literature review

The proliferation of web-based tools opens up new dimensions and brings new meaning to the notion of community. A community is no longer conceptualized merely in terms of physical proximity, but in terms of social networks [11]. [5] identified five dimensions of an online community: commonality, computer system, interaction, social infrastructure, and social relationships. Online communities could provide their members with multiple resources including information, social support and emotional support [8]. For some researchers, an online community is more closely associated with the participants' sense of community. According to [13], a sense of community is "a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith that members' needs will be met through their commitment to be together" (p. 9). [4]

defined community as a "general sense of connection, belonging, and comfort that develops over time among members of a group who share purpose or commitment to a common goal" (p. 2).

There is a type of online community labeled as a "blended community" or "hybrid community" which is enabled and supported by both online and offline methods of communication [6]. Such communities are grounded in the rationale that the relationship between online and offline communication can be supplementary (see [9]). In blended communities, the existing social relationships, too tight or too loose, might have negative effects on members' engagement and commitment [5]. On one hand, a close community with abundant opportunities for its members to interact through traditional ways might make the additional online interaction redundant [1]. On the other hand, if there is only a loose connection between community members, the incentive for extending communication to online space will also be low.

A community cannot survive without active and ongoing participation of its members [10]. Active participation takes the form of creating and consuming content [3]. [14] assert that the more people participate in online group activities, the more likely they are to develop online relationships. Likewise, [15] reports that the more time and effort people invest in the community activities, the greater the chances are for their sense of community to be ingrained. The reverse is also the case: a sense of community enhances participation [17]. However, the degree of participation varies significantly among the members of same community. [2] reported that both participation and achievement levels were uneven in an online learning community based on voluntary participation. Since the contribution will turn into public good, the temptation of enjoying the free-ride without contribution is pretty high [10]. This stems from a social phenomenon known as "social loafing" which refers to the tendency of exerting less effort in collective tasks than individual tasks [12].

3. Methods

3.1 Research setting

The study includes a group of 31 final-year pre-service teachers enrolled in an English educational program in a comprehensive university in Hong Kong. The course the study focuses on is a core course with an eight-week Teaching Practice (TP). The instructor was an enthusiastic and experienced user of educational technology. She set up a course website on Moodle and tried to engage students in voluntary online discussion as an extension of face-to-face teaching. She herself was active in facilitating online interaction among students and responding to students' questions. However, the overall level of online participation was rather sporadic. During the TP, the students themselves took the initiative to create a website for sharing teaching resources and lesson plans. This stimulated our interest in the study that examines the phenomenon of student online participation and their perceptions of the two types of websites.

3.2 Data collection and analysis

The study utilizes questionnaires as the main source of data. The purpose of the questionnaire was to capture the overall picture of students' online behavior and their perceptions and attitudes towards the two websites. The questionnaire items were developed based on the researchers' previous work on online participation [18] and [16]'s instrument for measuring participants' sense of community. The first section of the questionnaire allowed the researchers to gather data about the participants' general computer skills and comfort level with the use of technology. In the second section participants were invited to report on their online behaviors such as their frequency of reading messages. The third section consisted of Likert-type questions that tapped into participants' perceptions, in particular, their perceived usefulness of the websites and their sense of community (adapted from [16]). The questionnaire was administered on 31 students, yet one student did not complete it, which left the researchers with 30 valid responses.

4. Results

According to the results from the questionnaire, this group of students was quite comfortable with computer-mediated communication (Mean =3.17 on a scale of 4 with 1= very uncomfortable and 4= very comfortable). Their average level of self-rated computer skills was close to average (Mean = 1.9 on a scale of 3 with 1= weak, 2 = average, 3=high).

4.1 Descriptive data on online activities

Under "Learning Forums" on Moodle, there were 10 forums created. The first forum (Forum 1) had the highest number of posts (18) followed by Forum 7 (10). The number of students who took part in the online discussion as writers was also quite limited. There are five students who posted in Forums 1, 2, and 7, four in Forum 3 and one in Forum 9. To provide support and foster interaction among students during TP, the instructor set up a forum on course Moodle (Forum 10). However, there was no activity in this space at all. We also asked the students to report on their frequency of reading the content on course Moodle on a scale of 5 ranging from "never" to "always". 4 students (13%) "often" and, 16 (53%) "sometimes" read the content on course Moodle, but 10 (33%) "rarely" or "almost never" (M=2.77) did so.

"TP no worries" is set up on Google Sites as a publicly accessible website. In addition to the homepage, there are 12 pages created which are organized under two major categories: "Teaching Materials" and "Other reference". During TP, the student-teachers shared quite a number of teaching resources among one another, including their lesson plans, worksheets, audio and video files. The other section –"Other reference" included four pages that contained references such as bookmarks, video links, and useful evaluation forms contributed by students as well. Through the questionnaire, they reported on the frequency of reading the content on the "TP No Worries" website. 7 claimed that they had never accessed the site, 10 (33%) "rarely" or "almost never" visited the site, 11 "sometimes" (37%), and 2 (6.7%) "often" did so (M=1.93).

4.2 Perceptions of usefulness

In the questionnaire, students were asked about their perceptions of the usefulness of the two websites. For the course website on Moodle, the students basically thought of it as a storage place for course materials. Less than one quarter of the students felt the course website was useful for exchanging ideas and gaining insights into teaching. Students' responses on their perceptions of TPNW were more positive. 87% of them agreed to the statement that TPNW fostered the exchange of teaching ideas and sharing of teaching resources, and helped them get peer support from one another during TP. 83% also acknowledged that the platform supported their communication among peers. 61% felt that TPNW prompted them to reflect on their own teaching and 52% agreed the website helped keep them connected despite being in different physical locations for the TP.

We also explored the relationships among the students' perceived sense of community, their online participation, and their perceptions through running the Pearson correlation. The sense of community (SOC) score was obtained by computing the mean value of the four items that aimed to gauge students' sense of community. The results show that SOC is correlated with the frequency of reading TPNW (r =.43, p< . 05), but not with the frequency of reading on Moodle (r =-.1, p> . 05). No association was detected between SOC and students' perceptions of the course website on Moodle. Yet SOC was strongly correlated with the perceived value of TPNW for exchanging ideas (r =.6, p< .001), sharing teaching resources (r=.65, p< .001), enabling peer support (r = .55, p <.01), prompting reflection on teaching practice (r=.46, p<.05), and fostering peer communication (r= .47, p<.01).

5. Discussion and conclusion

The two websites in the study were developed and used quite differently. The course Moodle site was set up by the instructor as a place for sharing course materials and extending in-class discussions. The "TP No Worries" (TPNW) website was purely designed, developed, and maintained by the students themselves to promote a stronger sharing culture during TP. As to the locus of control, the course Moodle site was perceived by students to be implemented in a top-down fashion with the instructor taking the leadership and a supervisory role; the TPNW website was a student- initiated website emerging from the students' genuine personal needs for sharing and connection during their TP. This echoes the findings in [6] about grass-root-initiated online communities. It was thus interesting to note that the students were found to be less active readers on TPNW than they were on the course website. This may be interpreted by the fact that the students all needed to access the course materials on the website at some point in time during the course. However, their frequency of accessing the course website was not associated with their sense of belonging to the group. That implies that when CMS is used mainly for storing information, it did not give students a sense of ownership and hence it fails to generate a community spirit among them. In line with previous studies (e.g. [7]), student-teachers in the present study did not perceive the Moodle CMS as a platform for exchanging ideas and having insightful discussions, thus contributing to their limited interest in online participation.

On the contrary, those who used TPNW responded very positively to its impact in fostering peer sharing, communication and support. TPNW was regarded as a useful platform for student-teachers to exchange teaching ideas, share resources, gain support, and maintain communication with their fellow coursemates during their TP. It is worth highlighting that although the results show a lower frequency of reading the content in TPNW as compared to that in their course Moodle website, the frequency was correlated with their felt sense of community. Those who were active in TPNW tended to feel a stronger sense of community. When students visited such the website with lesson plans, teaching materials and resources all shared by their peers, their sense of belonging was strengthened and their inclination to access the website heightened. This may have implications for enhancing students' skills in developing and maintaining websites for their own learning.

Acknowledgement
The study is funded by the seed funding provided by Hong Kong Baptist University.

References

[1] Ardichvili, A., Page, V., & Wentling, T. (2003). Motivation and barriers to participation in virtual knowledge-sharing communities of practice. *Journal of Knowledge Management, 7*(1), 64–77.
[2] Bruckman, A. (2004). Co-evolution of technological design and pedagogy in an online learning community. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 239–255). Cambridge, UK; New York: Cambridge University Press.
[3] Butler, B., Sproull, L., Kiesler, S., & Kraut, R. (2002). Community effort in online groups: Who does the work and why? *Leadership at a Distance*. Mahwah, NJ: Erlbaum.
[4] Conrad, D. (2005). Building and maintaining community in cohort-based online learning. *Journal of Distance Education, 20*(1), 1–20.
[5] Deng, L., & Yuen, H. K. (2007). Connecting adult learners with an online community: Challenges and opportunities. *Research and Practice in Technology Enhanced Learning, 2*(3), 1-18.
[6] Gaved, M., & Mulholland, P. (2005*). Grassroots initiated networked communities: A study of hybrid physical/virtual communities.* Presented at the 38th Hawaii international Conference on System Sciences, Hawaii: IEEE.
[7] Hamuy, E., & Galaz, M. (2010). Information versus communication in Course Management System participation. *Computers & Education*, *54*(1), 169–177.
[8] Haythornthwaite, C., Kazmer, M. M., Robins, J., & Shoemaker, S. (2000). Community development among distance learners: Temporal and technological dimensions. *Journal of Computer-mediated Communication*, *6*(1). Retrieved from http://jcmc.indiana.edu/vol6/issue1/haythornthwaite.html
[9] Koku, E., Nazer, N., & Wellman, B. (2001). Netting scholars: Online and offline. *The American Behavioral Scientist*, *44*(10), 1752–1774.
[10] Kollock, P., & Smith, M. (1996). Managing the virtual commons: Cooperation and conflict in computer communities. In S. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 109–128). Amsterdam: John Benjamins.
[11] Kollock, P., & Smith, M. A. (1999). Communities in cyberspace. In M. A. Smith & P. Kollock (Eds.), *Communities in cyberspace* (pp. 3–25). London: Routledge.
[12] Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Li, X., Cosley, D., et al. (2005). Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication, 10*(4). Retrieved from http://www.blackwell-synergy.com/doi/abs/10.1111/j.1083-6101.2005.tb00273.x
[13] McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology*, *14*(1), 6–23.
[14] Parks, M. R., & Floyd, K. (1996). Making friends in cyberspace. *Journal of Communication*, *46*(1), 80–97.
[15] Roberts, T. L. (1998). Are newsgroups virtual communities? *CHI'98 Conference on Human factors in Computing Systems* (pp. 360–367). Los Angeles, CA.
[16] Rovai, A. P. (2002). Development of an instrument to measure classroom community. *The Internet and Higher Education*, *5*(3), 197–211.
[17] Yoo, W. S., Suh, K. S., & Lee, M. B. (2002). Exploring the factors enhancing member participation in virtual communities. *Journal of Global Information Management*, *10*(3), 55–71.
[18] Yuen, H. K., Deng, L., Fox, R., & Tavares, N. J. (2009). Engaging students with online discussion in a blended learning context: Issues and implications. In F. L. Wang, J. Fong, L. Zhang, & V. S. K. Lee (Eds.), *Hybrid Learning and Education*, Lecture Notes in Computer Science (LNCS 5685) (pp. 150–162). Berlin Heidelberg: Springer-Verlag.

Supporting Creativity Learning in Digital Storytelling with Tablet Computers: A Peer Assessment Approach

Chen-Chung Liu[a], Yu-Feng Wu[a*], Kuan-Hsien Lu[a], Chia-Ching Lin[a]
[a]Graduate Institute of Network Learning Technology, National Central University, Taiwan
[*]yfwu@cl.ncu.edu.tw

Abstract: Peer assessment (PA) has been widely adopted in an educational context with ample evidence suggesting the potential to promote learning. We argue that although the process of asking learners to involve in assessing achievement of the others might be a beneficial manipulation for learning, the effectiveness is still unclear for young pupils when creative learning tasks such as storytelling are involved. Creativity self-efficacy might be interfered with peers' feedbacks. Thus, the presented study proposes an approach to support creativity learning through peer assessment aided by Tablet computers. This study sought a deeper understanding of the relationship between creative self-efficacy (CSE) and creativity product among 54 sixth graders on completing a digital storytelling project. Findings and discussions are included.

Key terms: Creative self-efficacy, Peer-assessment, Digital storytelling

1. Introduction

"Who does what to whom…" We tell stories in our entire lives for that it is one of the most familiar patterns to make sense of our surrounding world [5]. The process of telling a story requires reasoning and elaborating where the minds immerse in the scenario, interchange ideas with personal and cultural values [9], while negotiating and renegotiating meanings within experiences [3]. Therefore, educators and researchers consider storytelling as an important pathway to fulfilling instructional objectives. However, scholars did not recognize storytelling with its educational potential until later where constructivists started to approach instructional design with learner-centered instruction, situated learning, learning styles, and so forth [5]. Thus, investigations on practicing the ability to generate and tell stories should be seen as rehearsals of skills to effective learning. Robin [10] defined digital storytelling as a combination of multimedia with "computer-based graphics, recorded audio, computer-generated text, video clips, and music" that allows users to "become creative storytellers through the traditional process of selecting a topic, conducting some research, writing a script, and developing an interesting story" (p. 222). However, most of children seem to lack of the opportunity to learn how to create and tell stories. To tell stories digitally, storytellers need to understand the fundamental knowledge and evaluate what are good stories based on the knowledge. However, recent studies of creativity indicated that teaching students to learn fundamental knowledge may limit the levels of creativity self-efficacy. Thus, there is a need for a pedagogical approach to enhance students' knowledge about stories and self-efficacy for the creative activity.

Peer assessment may be a potential approach to address the above issue as literatures suggested that the practice of assessing one another's work can help form the ability of recognizing key performance that is associated to high quality work [12; 16]. Peer assessment has complicated influence on creativity learning as learners' cognitive, affective, and self-efficacy status may have interfered with the entire learning process [7; 14]. On one hand, by evaluating others' storytelling works, one may know the fundamental knowledge about high quality stories. On the other hand, previous studies might have underpinned the complexity of the forming and the effect of peer assessment [13]. Strijbos, Narciss, & Dünnebier [11] pointed out peer assessment from the more able peers may led to a negative effect, and it might negatively influence their creative tasks [13]. This study thus aims to develop a pedagogical approach with peer assessment enabled by Tablet computers. It is hoped that through the aid of Tablet computers and peers assessment, students not only can develop fundamental knowledge about storytelling but also sustain their self-efficacy toward the creative activity.

2. Method

2.1 iPad Storytelling Application

Portable and easy to share make a Tablet computer distinguishable from a desktop computer, so as to enable the progress of peer assessment and interaction among peers. Thus to support creativity learning in the storytelling, this study developed an iPad application. The system allows the storyteller to draw (e.g., using pencil, color picker, eraser, cleaner, and stamp), to tell (e.g., using voice recorder and background music), and to frame (e.g., using new frame), in order to produce a digital storytelling project (see Figure 1)

Figure 1. Interface of the iPad storytelling application.

2.2 Creativity Self-efficacy (CSE) Questionnaire

In order to better understand how peer assessment influence one's creative self-efficacy, this study adopted a CSE questionnaire that was originally developed by Hung [4] examining CSE with three dimensions: (1) self-efficacy on creative performance (CSE-performance), (2) self-efficacy on creative strategy (CSE-strategy), and (3) self-efficacy on the attitude toward significant negative feedbacks from others (CSE-Nfeedback). Additionally, the summation of these three factors represents an overall level of CSE. Cronbach's Alpha reported internal consistencies of the CSE questionnaire on the overall CSE (.825), CSE-strategy (.78), CSE-performance (.64), and CSE-Nfeedback (.74), indicating a moderate level reliability.

2.3 Peer Assessment (PA) Form

The current study aimed at a peer assessment process that would facilitate learning of creativity in the context of storytelling. Based on a previous study [6], we proposed a 5-items-criteria that focuses story grammar, which consisted: (1) general settings, (2) storyline, (3) creativity, (4) technical performance, and (5) overall presentation of the story. Such PA was designed for testing our assumption that applying PA could promote students' creativity performance.

2.4 Grading Rubrics and Process of Storytelling Product

Since the Draw&Tell application turns stories as slow motion frames (i.e., slideshows), the current study adopted a widely used checklist for movie production as the grading rubrics [15]. The rubrics examine the final production within nine emphases: (1) transitions & edits – 2 items, (2) planning & storyboarding – 4 items, (3) action & dialog – 2 items, (4) accuracy of information – 2 items, (5) originality & creativity – 2 items, (6) sound usage – 2 items, (7) drawing – 2 items, (8) camera picturing – 2 items, and (9) framing – 2 items. Grades were calculated per item from five (i.e., the highest points) to one (i.e., the lowest points). The grades from these nine categories with total 20 items were calculated together to indicate an overall performance. Additionally, we would look more closely at one emphasis of originality & creativity than the others. Two raters used this rubric to grade all of the 54 final products. The results were tested by correlational analyses and the result yielded to a range between .539 and .848 ($p < .01$) which indicated the grading process is highly reliable.

2.5 Participants & Procedures

Two classes consisted with 54 sixth graders from an elementary school in northern Taiwan were randomly selected to participate in the current study.

Eight consecutive classes with one hour weekly meeting were scheduled. At the 1st class, the pre-CSE was administered, and orientation for storytelling and the iPad Draw&Tell application. Students were then had the opportunity to practice a story for the 1st and 2nd week. From the 3rd to the 8th week, students were assigned a new topic, "Saving the ecological environment", as the formal project. The experimental group started to review their peers' works during the 6th and 7th weeks. Rather than grading on them, they review works with a PA handout. Students then submitted the handouts to the instructor without further process regarding PA. As a comparison, the controlled pupils did not know and process the peer assessment. After the eighth weeks, all students were asked to fill the post CSE questionnaire at home in the 9th week. In the 10th and 11th weeks, 12 pupils from both the groups were randomly selected to be interviewed. These data were treated rather as anecdotal data than qualitative data. Finally, the entire treatment was completed.

3. Results

T-test did not report any difference on the overall creative self-efficacy (CSE), CSE-strategy, CSE-performance, and CSE-Nfeedback between the 2 classes' pre-CSE questionnaire, suggesting an equality of homogeneity on the 2 groups' pre-CSEs.

In regards to the post-CSEs, T-test reported some significances between the 2 classes' post-CSE: overall CSE, $t(52) = 2.163$, $p = .035$; CSE-strategy, $t(52) = 2.230$, $p = .030$.

In regards to creative performance, T-test reported some significance on final scores of students' creative products between the two groups: overall score, $t(52) = 2.591$, $p = .012$; transitions & edits, $t(52) = 3.769$, $p < .001$; planning & storyboarding, $t(52) =$

3.495, p = .001; accuracy of information, t(52) = 3.230, p = .002; drawing, t(52) = 2.620, p = .012. Selected significant items are listed in Table 1 with descriptive statistics.

Table 1. CSE and Final scores

Group	**N**	**Creative Self-Efficacy**	**Mean**	**S.D.**	**Selected Final Scores**	**Mean**	**S.D.**
PA	28	Performance	3.36	.70	Originality & Creativity	2.87	1.09
		Strategy*	3.76	.61	Transitions & Edits*	3.60	.89
		Nfeedback	3.87	1.02	Planning & Storyboarding*	3.20	.94
		Overall*	3.64	.49	Accuracy of information*	3.35	1.12
					Overall*	2.98	.90
NonPA	26	Performance	3.10	.68	Originality & Creativity	2.36	.93
		Strategy*	3.43	.44	Transitions & Edits*	2.77	.71
		Nfeedback	3.65	.66	Planning & Storyboarding*	2.39	.74
		Overall*	3.37	.43	Accuracy of information*	2.42	.98
					Overall*	2.39	.74

Note. * indicates significant difference between the 2 groups

4. Discussion

Hypothesis 1. Peer assessment has a positive effect on performance of digital storytelling utilizing Draw&Tell Tablet computer application among elementary schoolers. – *PA has a positive effect on participants' overall score.* Hypothesis 1 was retained as the result indicated that PA had a positive on participants' final score. Such finding is consistent with the existing literature that PA promotes peer learning [12]. Furthermore, the finding supported our assumption that a simplified PA could eliminate a potential negative effect on creativity outcomes and promote creative outcomes. Bandura [1] suggested that instead of applying traditional types of PA (e.g., grading, commenting, feedbacks, etc.), a PA process to be more informing and less evaluating-orientated might produce a positive effect on creative outcome.

Hypothesis 2. Peer assessment has a positive effect on creative self-efficacy (CSE) of digital storytelling utilizing Draw&Tell Tablet computer application among elementary schoolers. – *PA has a positive effect on participants' levels of CSE.* Hypothesis 2 was retained as the result indicated that PA had a positive effect on participants' levels of CSE. Literature suggested a raise on the level of CSE might indicate a higher quality of creative outcome [8], as well as the potential on the academic success [2]. The finding echoes that a personal belief on creativity could be enhanced by training (i.e., PA could be seen as a reinforcement of informing personal belief on creativity) [8].

In order for a better understanding of participants' thoughts behind the data, we randomly interviewed 12 students from both the groups as anecdotal data. Unlike in the experimental group where students had the opportunity to learn from their peers, students in the controlled group mostly revealed that "*I don't feel my work is special*", or, "*I feel others might have done better than me.*" When we took one more step by asking "*Better on? Or worse on what?*" Students could not specify what they mean about the differences. This echoed our finding that a peer assessment process was helpful for students to be appreciative of ones' own works, and thus then increase the level of personal belief on creative performance (i.e., CSE). We heard similar reflections when asking what they would have felt if negative feedbacks were given. Most students from both groups revealed that they would not give up their ideas. This might explain a PA did not constrain self-efficacy on creativity, and why students' final score correlated positively with CSE-Nfeedback. It was conjectured that students who possessed a higher level of creative self-efficacy on dealing with significant feedbacks, there was a potential that they could perform a higher quality of the task.

5. Conclusion

The current study investigated a peer assessment effect on students' creative work (i.e., storytelling) and creative self-efficacy in a creativity learning context. The findings suggest implementations in elementary education that peer assessment may promote students' creativity performance and creative self-efficacy. Additionally, the current study re-tested the assumption that creative self-efficacy could be enhanced by reinforcing the beliefs about creative performance. The reinforcement was carried with a peer assessment process in the presented study. As a result, students performed better quality of their creativity works as well as reflected a higher level of creative self-efficacy. Such findings may inform an implementation for educational settings.

Acknowledgement

The presented study is supported by the National Science Council in Taiwan, NSC 100-2631-S-008-004.

References

[1] Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
[2] Beghetto, R. A. (2006). Creative self-efficacy: Correlates in middle and secondary students. *Creativity Research Journal, 18*(4), 447-457. doi: 10.1207/s15326934crj1804_4
[3] Bruner, J. (1990). *Acts of meaning*. Harvard University Press: Cambridge.
[4] Hung, S.-P. (2003). *A social cognitive model of significant other's feedback, creative self-efficacy, internal motivation and external motivation on creative behavior: A structure equation modeling study.* Master Thesis, National Chiao-Tung University, Taiwan.
[5] Jonassen, D. H., & Hernandez-Serrano, J. (2002). Case-based reasoning and instructional design: Using stories to support problem solving. *Educational Technology Research and Development, 50*(2), 65-77. doi: 10.1007/bf02504994
[6] Liu, C.-C., Chen, H. S. L., Shih, J.-L., Huang, G.-T., & Liu, B.-J. (2011). An enhanced concept map approach to improving children's storytelling ability. *Computers & Education, 56*(3), 873-884. doi: 10.1016/j.compedu.2010.10.029
[7] Lu, J., & Law, N. (2012). Online peer assessment: effects of cognitive and affective feedback. *Instructional Science, 40*(2), 257-275. doi: 10.1007/s11251-011-9177-2
[8] Mathisen, G. E., & Bronnick, K. S. (2009). Creative self-efficacy: An intervention study. *International Journal of Educational Research, 48*(1), 21-29. doi: 10.1016/j.ijer.2009.02.009
[9] Polkinghorne, D. (1988). *Narrative knowing and the human sciences*. Albany, NY: SUNY Press.
[10] Robin, B. R. (2008). Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory into Practice, 47*(3), 220-+. doi: 10.1080/00405840802153916
[11] Strijbos, J.-W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning & Instruction, 20*(4), 291-303. doi: 10.1016/j.learninstruc.2009.08.008
[12] Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249.
[13] Topping, K. (2010). Methodological quandaries in studying process and outcomes in peer assessment, Editorial, *Learning & Instruction,* pp. 339-343. Retrieved from http://0-search.ebscohost.com.source.unco.edu/login.aspx?direct=true&db=ehh&AN=48891870&site=ehost-live
[14] Tseng, S.-C., & Tsai, C.-C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education, 49*(4), 1161-1174. doi: 10.1016/j.compedu.2006.01.007
[15] Via, S. (2002). http://its.ksbe.edu/dst/PDFs/Rubrics/rubric.pdf.
[16] Yu, F. Y., & Wu, C. P. (2011). Different identity revelation modes in an online peer-assessment learning environment: Effects on perceptions toward assessors, classroom climate and learning activities. *Computers & Education, 57*(3), 2167-2177. doi: 10.1016/j.compedu.2011.05.012

An Online Platform for Problem-Based Learning with Operational Concepts Map

Chung-Cheng TSENG[a*], Po-Yao CHAO[b], Chih-Yuen CHOU[a] & K. Robert LAI[a]
[a]*Department of Computer Science & Engineering, Yuan Ze University, Taiwan.*
[b]*Department of Information Communication, Yuan Ze University, Taiwan.*
*cct0403@saturn.yzu.edu.tw, poyaochao@saturn.yzu.edu.tw, cychou@saturn.yzu.edu.tw, and krlai@cs.yzu.edu.tw

Abstract: Problem-based learning (PBL) encourages learners to acquire new knowledge through cognitive model development and hypotheses-deduction process. Graphical modeling tools, such as concept map, can help learners to exteriorize the process of knowledge construction and visualize learner's cognitive model to support meaningful learning. But few of those tools integrate hypothesis-reasoning mechanism into the process to support the discovery of knowledge deficiencies and foster in-depth reflection. To that end, this paper proposes a novel graphical tool, *operational concept map* (OCM), for supporting the PBL. In OCM, it combines the merits of conceptual modeling tool and computer-based simulation. The OCM not only inherits all basic functions of concept map, also facilitates the learning-by-doing approach for problem solving. Application of the proposed idea is illustrated in learning a well-known sort algorithm, *bubble sort*, in the sort problem.

Keywords: Problem-based learning, Concepts map, Operational concepts map, Simulation, Learning by doing.

Introduction

Problem-based learning (PBL) has been seen as an active learning approach to promote meaningful learning [1, 2]. This active approach encourages learners to acquire new knowledge through an iterative hypotheses-deduction process when solving a problem. During the process, a learner analyzes the problem by identifying key fact with the problem scenario, proposing hypotheses and solutions, and evaluating the hypothesis by monitoring outcomes of the solutions. Among this process, research shows that the development of cognitive model and the evaluation through experiments can help a learner to identify the key concepts and evaluate the hypothetical relationship among these concepts [3, 4].

Cognitive model [5, 6] is a representation of key elements with relationships among these elements in the specific phenomena or knowledge. When solving a problem, cognitive model can provide a learner with reference framework to facilitate the thinking of a problem[7]. Some graphical modeling tools, such as concepts map[6], have demonstrated their usefulness in construction of learners' cognitive models. With a concept-mapping tool, a cognitive model can be represented as a well-organized graphical map composed of verbal and symbolic elements. This useful tool provides scaffolds to assist a learner to develop and represent the structure of knowledge as personal cognitive model regarding a problem or topic [8].

While, in problem solving, it often requires a learner to reflect and then refine his/her own cognitive model by iteratively evaluating the model's effectiveness, the mechanism for evaluating hypothetical relationship among key concepts to facilitate learning by experimenting is seldom integrated with a concept-mapping tool. To this end, computer-based simulation[9] seems to provide an environment in which the learners can evaluate their hypotheses by experimenting. The simulation functions assist learners to develop critical skills including identify key concepts of a problem, generate hypotheses, and propose solutions for the problem[10]. These simulation environments not only facilitate the development of cognitive model by evaluating hypotheses, but also provide learning goals for learning by doing activities. Based on the necessity of evaluation mechanism of graphical modeling tool and the need of learning by experimenting in PBL, this paper proposes a novel graphical tool, namely *operational concept map* (OCM), which incorporates the basic functions of concept map with additional hypothesis evaluation mechanism. The intent is to foster effective problem solving through developing cognitive models and evaluating hypothesis in a computer-based simulation environment.

The remainder of this paper is organized as follows. Section 1 introduces a conceptual framework of OCM. In this framework, learners can construct a knowledge model with hypotheses-reasoning mechanism. Then, Section 2 describes how the OCM can enhance the process of PBL. In this section, we present an example to illustrate how the OCM can be applied for learning the concept of bubble sort in the sort game followed by a conclusion in Section 3.

1. Operational Concept Map

1.1 Overview

Operational concept map (OCM) is a graphical modeling tool for organizing and representing knowledge with deductive-reasoning mechanism. An OCM is composed of concepts, hypothesis links among these concepts, and propositions. Firstly, concepts, which are enclosed in circles as shown in Figure 1, are perceived regularity in events or objects, or records of events or objects, designate by label. Secondly, the hypothesis link, indicated by rectangle with round corners and several connecting lines linking concepts, is a possible causal relationship among these concepts. Each link contains a hypothetical statement that describes nominal, logical or mathematic formulations. Finally, a proposition contains one hypothesis link connecting two or more concepts to form a meaningful statement. To implement hypotheses-reasoning mechanism, each operational concept and corresponding hypothesis links in OCM are bound with specific set of problem instances. These problem instances are collected to provide real world problem data that helps learners to evaluate the effectiveness of an OCM. Then, learners apply different sets of problem instances to ensure the satisfaction of hypothetical statements in hypothesis links.

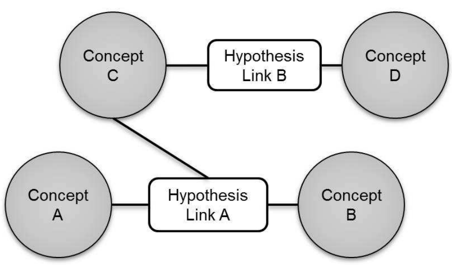


Figure 1. A simple operational concept map

1.2 Key Features

1.2.1 Development of cognitive model and hypotheses evaluation

Like concept map, the OCM assists learners to incorporate new information into the relevant concept framework of which they have already built when a learner generates his/her cognitive model for a problem. The OCM helps the learner to identify general concepts held by the learner prior to instruction and on more specific concepts, and anchor the new knowledge into the conceptual framework. Moreover, the relationship among these key concepts in OCM represents not only the semantic relationship, but also the rule of constraints. The rule can be represented as a mathematic formulation, such as “A=B+C, or logical representation, such as “IF A > B, then Swap(A,B)”. For hypothesis evaluation, hypothesis links represented as a constraint serve not only as semantic description of model to be simulated, but also as rules telling that certain conditions must be satisfied [11-13]. A learner maintains the satisfaction of their hypothesis constraints by iteratively revising their problem cognitive models. Moreover, in OCM, the just-in-time feedbacks of hypotheses evaluation are offered immediately as simulation outcomes. A learner refines his/her cognitive model according to the just-in-time feedbacks.

1.2.2 Construction process of OCM

The construction process of OCM is one kind of externalizing internal and hidden mental model. Some kinds of modeling process [14, 15] facilitate the learner to construct cognitive model with verbalizing and writing. But, in this process, OCM is an external memory tool to facilitate the cognitive process by visualizing. Because the thinking process of problem-solving is represented as organization of concepts and hypothesis links in OCM, the simulation then can not only facilitate the exploration of what-if situations and also support the hypothesis-reasoning process for identifying the knowledge deficiencies. Finally, the visualization with graphical cognitive model facilitates this constructive process with guidance of what to do and how to do.

2. Application of OCM to learn a concept of Bubble sort in the sort problem

To demonstrate the application of OCM in supporting PBL, a well-known learning issue in an introductory college computer science course, the sort problem, is adapted for this study. A learner will learn various sort algorithms to solve the sort problem. In this example of sort problem, the bubble sort algorithm is chosen because it is a simple algorithm that compares each pair of adjacent items and swaps them through the list to be sorted until no swaps are needed. Figure 2 displays the role of instructor and learner played in the process of applying the OCM in the sort problem.

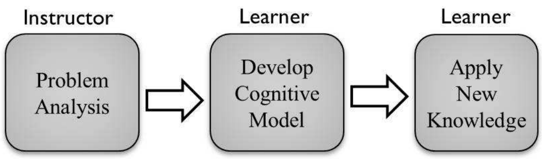


Figure 2. The process of applying the OCM in the sort game.

2.1 Analyzing problem scenario by Instructor

In order to apply OCM to sort problem, the instructor has to analyze a problem scenario at first. In the example of sort problem scenario, four values are randomized and placed individually in the four cells in sequence. A learner has to sort this set of values in ascending order. The instructor has to analyze this problem to identify the key concepts and possible relationships among these key concepts. In this case, the key concepts are the position of each cell. Therefore, the instructor uses four cells of array B to represent the key concept. So, in the figure 3, these key concepts (the four cells of array, B(1) to B(4)) are included in the Concepts Repository for learners to construct their own OCMs, and the Simulation Panel will represent the sort process and result of this array for evaluation. The instructor then identifies the relationships between two adjacent cells and constructs possible reference models to evaluate learner's cognitive model.

2.2 Learner develops his/her cognitive model

In this phase, the learner has to construct his/her personal cognitive model for the sort problem. But, at first, the learner has to read the description of bubble sort algorithm which is prepared by the instructor. Then, the learner has to realize the learning goal and key concepts of a sort problem. The learning goal is defined by the instructor to tell learners that some conditions they have to satisfy. The goal of sort problem is that the unsorted list has to be sorted in ascending order. The key concepts have been identified in the concepts repository.

After realizing the problem scenario, the learner starts to develop his/her cognitive model by OCM. In this example of sort problem, the learner will describe each step of the bubble sort algorithm in sequence. In this case, the learner places two adjacent elements, B(1) and B(2) in the OCM Construction Panel, to determine hypothesis link among them. The possible relationship between B(1) and B(2) is that if the value in B(1) is larger than the value in B(2), these two values have to be swap. Therefore, the learner places a hypothesis link, named "swap", and establish connection lines between two concepts and hypothesis link to form a proposition. Next, the rule of "swap" is established to represent a logical representation: "if B(1) > B(2), then swap(B(1), B(2))". Then, this hypothesis can be evaluated by simulation. In the simulation, the learner observer the result generated from the problem scenario for this hypothesis. If this hypothesis link is work, the learner can observer that value in B(1) will be swap for value in B(2) (if B(1) is larger than B(2)). Then, the learner will apply this hypothesis link to the next two adjacent concepts repeatedly. The cognitive model of bubble sort algorithm is represented as Figure 3.

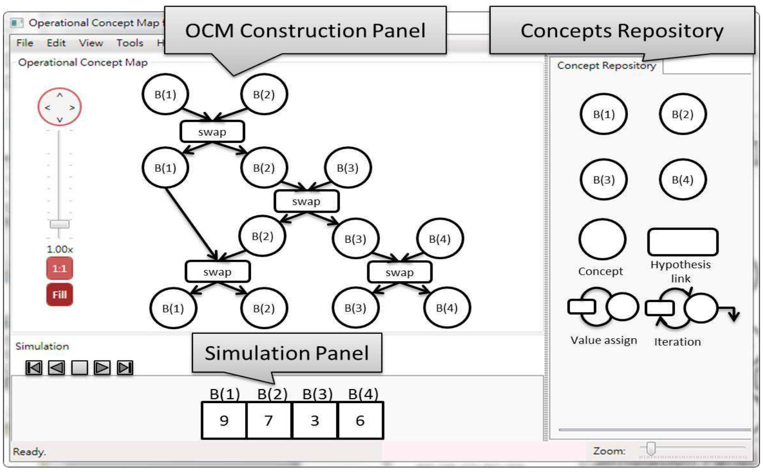


Figure 3. OCM Construction Environment

2.3 Learner acquires and applies new knowledge

When personal cognitive model is completed, another scenario of sort problem is applied to evaluate it. If this model cannot reach the learning goal of new scenario, it has to be revised.

For example, an unsorted list consist of 100 elements cannot be solved by the model for four elements of list. The learner may develop various models for different size of lists, or deduct a general model for different scenario by using the iteration construct. This is an adaptive process that personal cognitive model will be adapted by various problem scenario.

3. Conclusions

This paper has proposed an operational concept-mapping approach for supporting PBL. In OCM, it combines the merits of conceptual modeling tool and computer-based simulation. During the iterative construction process, OCM can foster in-depth learning by experimenting with the proposed game simulation environment. Learners can have better understanding of what they do and why they do in the problem-solving. While the preliminary results look interesting, more issues remain to be further investigated, including application of OCM for supporting various learning through problem solving, and development of better online environment for supporting PBL with OCM.

References

[1] Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, 16(3), 235-266.

[2] 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.

[3] Macklin, A. S. (2001). Integrating information literacy using problem-based learning. *Reference Services Review*, 29(4), 306-313.

[4] Torp, L., & Sage, S. (2002). *Problems as Possibilities: Problem-Based Learning for K-16 Education (2 ed.):* Association for Supervision & Curriculum Development.

[5] Novak, J. D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education*, 86(4), 548-571.

[6] Novak, J. D., & Cañas, A. J. (2008). The Theory Underlying Concept Maps and How to Construct and Use Them, *Technical Report IHMC CmapTools.* Florida Institute for Human and Machine Cognition: Florida Institute for Human and Machine Cognition.

[7] Ormrod, J. E. (2010). *Educational Psychology: Developing Learners (7 ed.):* Prentice Hall.

[8] Rendas, A. B., Fonseca, M., & Pinto, P. R. (2006). Toward meaningful learning in undergraduate medical education using concept maps in a PBL pathophysiology course. *Advances in Physiology Education*, 30(1), 23-29.

[9] Schank, R. C. (1995). What We Learn When We Learn by Doing, from http://cogprints.org/637/1/LearnbyDoing_Schank.htm

[10] Savery, J. R. (2006). Overview of Problem-based Learning: Definitions and Distinctions. *The Interdisciplinary Journal of Problem-based Learning*, 1(1), 9-20.

[11] Lai, K. R., & Lan, C. H. (2006). Modeling Peer Assessment as Agent Negotiation in a Computer Supported Collaborative Learning Environment. *Educational Technology & Society*, 9(3), 16-26.

[12] Lan, C. H., Graf, S., Lai, K. R., and Kinshuk (2011). Enrichment of Peer Assessment with Agent Negotiation, *IEEE Trans. on Learning Technologies, Vol. 4, No. 1, pp. 35-46*.

[13] Mitrovic, A. (2010). Modeling Domains and Students with Constraint-Based Modeling. In R. Nkambou, J. Bourdeau & R. Mizoguchi (Eds.), *Advances in Intelligent Tutoring Systems* (Vol. 308, pp. 63-80): Springer Berlin / Heidelberg.

[14] Collins, A., Brown, J. S., & NewMan, S. E. (1989). Cognitive Apprenticeship: Teaching the Crafts of Reading, Writing, and Mathematics. In R. Lauren B (Ed.), *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser* (pp. 453-494). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

[15] Pedersen, S., & Min, L. (2002). The Transfer of Problem-Solving Skills from a Problem-Based Learning Environment: The Effect of modeling an Expert's Cognitive Processes. *Journal of Research on Technology in Education*, 35(2), 303.

Small Group Learning with Digital Pens in High School Physics

Ming-Hong CHIU[a], Chi-Jen LIN[b*]
[a]*National Lo-Tung Senior High School, Taiwan*
[b]*Department of Learning and Digital Technology, Fo Guang University, Taiwan*
*cjlin@mail.fgu.edu.tw

Abstract: Physics learning was known to be hard for high school students and it was known that there was a gender difference in this perspective. One important method to promote student learning achievement in physics is to increase the opportunities to receive feedbacks during problem-solving practices, and small group learning is one of such a good method. Another method is to provide multiple viewpoints to assist student learning, and digital pens is a potential tool to achieve this goal. The experimental results showed that female students are more benefited from small group learning and male students may suffer in learning retention when digital pens were used in small group learning setting.

Keywords: small group learning, digital pen, physics concept

Introduction

Physics has been perceived as a hard subject by most high school students in Taiwan. One of the reasons that make physics hard to be mastered was that most students did not know how to solve problems with physics concepts properly, thus they tended to acquire superficial problem solving techniques [1]. Without interacting with physics teachers, students were unaware whether their learning was superficial or not. The authors believe that increasing student opportunities to receive feedback during their problem-solving practices at their early learning stage is crucial to improve their learning. However, due to limited time in the classroom, a physics teacher can spend time only with partial of the students in a class, and the rest of the students in the class are, unfortunately, on their own. The solution to this situation is the well-known two sigma problem [2], which is a search of group instruction methods that are as effective as one-to-one tutoring.

In a physics class, besides the teacher, more capable students are the potential helpers to enhance student learning. Peer instruction [3] is an instruction method that matches students pairwise in order that the interaction between the two students helps their learning. However, it was found that if the two students in a pair were both less capable students, the learning outcomes might be worse than they studied alone. With careful assignment of the group leaders, the discussion in the small groups may result in learning enhancement. Thus, it is interesting to investigate whether such a small group learning method will enhance physics learning. Another method to improve student learning is provide students with multiple viewpoints. We use digital pens to achieve this goal. The effects of introducing digital pens will also be investigated to see if there exists any further enhancement. As gender was also known to be significant factor in physics learning [4] [5] [6], we will also investigate how the gender factor influence student learning under relevant situation.

1. Characteristics of Digital Pens

The usage of digital pens is quite the same as those of ordinary pens, even their appearance is the same (as shown in Figure 1). The main differences between a digital pen and an ordinary pen are that a digital pen comes with an internal camera to capture what was written using the digital pen and a wireless transmitter to transmit what the camera captured to a whiteboard or a computer. The configuration of transmission target is up to the teacher. Thus, a teacher can monitor answers of all the student groups and react accordingly at one location when all the groups use digital pens (teacher time saving feature). The monitoring screen can also be displayed to all students to demonstrate various solutions of a problem. The benefit of introducing digital pens is that it does not require additional training for students to use them. However, to ensure the proper capturing function of the cameras, specially prepared paper, which comes with invisible printed dots, must be used for writing. In this study, we use digital pens as a tool for demonstrating alternative solutions in order to investigate whether students learning will be further improved in the small group learning format.

Figure 1. Digital pen(bottom) and ordinary pen(above).

2. Two-staged Experiment

Due to the nature of this study, a two-staged experiment was planned. The goal of the first stage experiment was to investigate whether the small group learning format outperformed the individual learning format, and the goal of the second stage experiment was to investigate whether the introduction of digital pens did further improve physics learning in the small group format.

2.1 Experimental settings

The learning topic of the first stage experiment is the unit of gravity on earth, while the learning topic of the second stage experiment is the unit of satellite movement. Each topic was taught for two contiguous weeks with a total of eight classes. The class delivery format was consisted of three parts: concept introduction, problem-solving demonstration, and student problem-solving practices. The learning format of the control group and experimental group differed only at the third part, that is, the student problem-solving practice part. In the individual learning setting, or the control group of the first stage experiment, each student solved problems by themselves, while the teacher walked through the classroom to assist those who need help. In the small group learning setting, which includes the experimental group of the first stage and the control group of the second stage, a class of students was divided into eight groups. Each group was required to work out a solution and their solutions were reported to the whole class after all the groups completing their solutions. The group leader was chosen by the teacher according to their previous physics learning performance. Eight top ranked students were divided into the eight groups and were assigned as the group leader. The group leader was in charge of coordinating the group solution among the group members and reporting the group solution to the whole class. The teacher played the role of commentator during their reports. In the digital pen learning settings, students also learn in the group format, except that group leaders use digital pens as their written tools and group solutions were shown on a common display that was visible to the whole class.

2.2 The participants

The subjects come from two classes of high school students of eleventh grade. One of the classes served as the control group and the other served as the experimental group. There are 44 students in the control group with 13 female students and 31 male students, and 37 students in the experimental group with 23 female students and 14 male students. Before the two-staged experiment, a formal learning performance assessment was performed for the two classes. The results showed that there were no significant statistical differences between the previous physics learning performance of the two classes, with the average of learning performance of the control group slightly better than that of the experimental group. The two classes were taught by the same teacher.

In each stage of the experiment, a pretest and a posttest of the learning unit were conducted. In the second stage, a delayed test was also performed to investigate the retain effects of the intervention of digital pens. The scores of these tests were the number of problems the students solved correctly. There were totally ten problems in each test.

3. Experimental Results

3.1 Small group may suffer from classroom order degeneration

Before the formal two-staged experiment, a prior experiment, which was arranged in the same setting as the first stage experiment, was conducted for the same subjects. To our surprise, the results showed that the learning performance of students in the individual learning setting was significantly better than that of in the small group learning setting. After reviewing the classroom process, it was found that the subjects chatted much more often in the small group learning setting than their previous classes. That is, classroom order degenerated during the experiment. For the rest of the classes the student attended, their format was individual learning. The discussion nature of the small group was novel to the students and they misused the opportunities of learning discussion by chatting. However, it was also noticed that one of the groups that was close to a video recorder chatted much less than the other groups. Consequently, in the experiment, we set up a video recorder for each group in order to prevent unnecessary chatting. This strategy worked well in the experiment.

3.2 Small group learning outperforms individual learning

Table 1 denotes the student performance in the first stage experiment. The numbers in the table denote how many problems the student correctly solved in the tests. The result of t-statistics of the pretests indicates that there is no significant difference between the student performance in the control group and the experimental group. On the other hand, the students in the experimental group outperform the students in the control group. That is, small group learning is better than individual learning. This result is a support that more capable students in the small group do help enhance the learning of the other group members.

Table 1. Student performance in the first stage experiment (n=81)

		N	Mean	SD	Learning gain
Control group	pretest	44	3.30	1.81	2.97
	posttest	44	6.23	1.66	
Experimental group	pretest	37	3.08	1.48	4.08
	posttest	37	7.16	1.48	

4. Conclusions

In this study, we investigated the effectiveness of small group physics learning by assigning more capable students as group leaders, whether the introduction of digital pens in the small group learning can further improve student learning, and whether gender difference exists in these investigations. The results showed that such a small group learning method did improve student learning and the learning gains of female students were greater than those of male students. However, introducing digital pens in the small group learning format, which was a means to provide students with multiple viewpoints, did not produce additional learning gains.

References

[1] Crouch, C.H. & Mazure, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970-977.

[2] Bloom, B. S. (1984).The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13, 4-16.

[3] Mazure, E. (1997) *Peer Instruction: a User's Manual*. Prentice Hall, New Jersey.

[4] Cavallo, A. M. L., Potter, W. H., & Rozman, M. (2004) Gender Differences in Learning Constructs, Shifts in Learning Constructs, and Their Relationship to Course Achievement in a Structured Inquiry, Yearlong College Physics Course for Life Science Majors. *School Science and Mathematics*, 104(6), 288-300.

[5] Hazari, Z., Tai, R. H., & Sadler P. M. (2007) Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors. *Science Education*, 91(6), 847-876.

[6] Kost, L. E., Pollock, S. J., & Finkelstein N. D (2009) Characterizing the gender gap in introductory physics. *Physical Review Special Topics - Physics Education Research*, 5(1), 010101.

Aiding Digital Natives Learn Positive Learning Behaviors through Reflection

Roberto LEGASPI*, Paul Salvador INVENTADO, Rafael CABREDO & Masayuki NUMAO
The Institute of Scientific and Industrial Research, Osaka University, Japan
*roberto@ai.sanken.osaka-u.ac.jp

Abstract: Commonly attributed to digital natives is the ability to quickly, yet effectively, shift from one task to another. However, several works have debunked this assumption by showing that multitasking even among digital natives led to poor learning performance and productivity. Our aim is to provide a tool to help digital natives be self-aware of desirable, while curbing undesirable, learning behaviors. Our tool is infused with self-annotation and feedback mechanisms that allow students to reflect upon their entire learning history. Our results indicate that the annotation process with the tool helped students understand their learning behaviors better and identify ways in which their behaviors can be improved.

Keywords: Digital natives, learning behaviors, productivity annotation tool, user modeling

Introduction

For a time, the notion that digital natives [17], also called "Gen M" (for multimedia) [19] or "Net Gen" [17,19,22], are capable of switching tasks [8,15,18] prevailed until other researchers started to level this assumption. Digital natives are technology-savvy and carry out many of their activities on the web [18]. They engage with and communicate socially through various multimedia [8,15]. Since the tools they use to study, get entertained and socialize are equally accessible at any time, digital natives tend to switch from learning (e.g., viewing tutorial videos) and non-learning (e.g., engage in game or social network websites) tasks. However, there is evidence to suggest that task-switching behavior can lead to poorer performance and learning outcomes [11] and result to time lost that increases with the complexity of, and unfamiliarity to, the task [20]. It is plausible that productivity deteriorates when alternating tasks.

Our objective is to help the digital native learner become aware of problems resulting from learning and non-learning task-switching and help him/her self-regulate towards increased learning productivity. This paper reports the results we obtained from subjects who experimented with our tool designed to raise self-awareness of effective learning behaviors - to discover routines that lead to productive learning and become cognizant of personal factors that support positive study by constantly aiding reflections on activity choices made that helped achieve (or otherwise) the desired learning outcomes.

1. Productivity Annotation Tool

Being aware of one's own learning behavior is metacognitive and students who are able to step back, reflect in retrospect, and adjust their thinking are able to adapt and monitor their behaviors to know what they need to study and how to do it [10]. We have designed a tool basically to this end. We infused in its design a self-annotation scheme that will allow students to review and reflect on their past learning sessions.

When using our tool, students identify their goals for the current learning session and then start the tool's data collection process to run in the background. Once the session is over, a timeline representing their entire learning session is shown to the students. When the mouse hovers over this timeline, screenshots of the desktop and webcam are shown to help students recall what happened at any point in time. Apart from using their mouse for annotation, an image player can also be used to select time partitions, much like how a video player is used, together with a selection button to indicate when to begin and end the selection. After selecting a time span, students can then provide their annotation.

Students annotate three aspects of their learning session, namely, *intention*, *activity*, and *affective state*. The intention is either goal-related, i.e., the activities relate to reaching any of the defined goals, or nongoal-related. Activities can be labeled as primary or secondary. A primary activity refers to what the student was focused on during the selected time span (e.g., using `chrome.exe`, reading a technical paper, or viewing a tutorial video). Secondary activities are those done together with the primary (e.g., listening to music, drinking coffee and eating, aside from others). The students also input the affective states they experienced as they performed the activities, i.e., delighted, engaged, confused, frustrated, surprised, afraid and neutral for goal-related activities [5], and angry, disgusted, sad, delighted, afraid and neutral [6] for nongoal-related ones.

After annotating the entire session, students assess their learning session in terms of the percentage of goal completion, activities that helped achieve their goals, and their perceived productivity level (in a scale of 1-5). The tool then provides statistics regarding the students' learning sessions in terms of the amount of time they spent in goal- and nongoal-related activities, together with the emotions they felt during each type of activity. The tool then produces a log file that contains timestamps, intentions, activities and affective states for a learning session. The productivity and goal completion ratings will be stored in a separate log file. All these will be used by the software for analysis of student learning patterns.

We designed our tool while taking cues from prior works. Goals enhance self-regulation, self-motivation, self-efficacy, and self-evaluation of progress [4,21]. When goals are defined, there is the tendency for students to do less the other tasks since they are aware of what they are supposed to do. Secondly, when students annotate their activities, they are actually performing self-monitoring. Seeing in retrospect how they spent their time learning, they get an idea of what comprises their study habits [10]. They can identify what behaviors they think are helpful to them as well as those that cause distractions. It will also help them see how their affective states actually influenced their learning process [5,16]. For example, emotions such as boredom and frustration can cause students to engage in non-learning tasks or stop learning altogether [12]. The tool can therefore help students identify and retain good study habits and mitigate unwanted learning behaviors.

2. Related Works

The importance of students setting their own goals, selecting the appropriate strategies, applying these strategies and monitoring their effects in order to modify them when necessary has spawned research in providing support tools for such metacognitive tasks while learning with computer-based learning environments, such as intelligent tutoring systems and hypermedia-based environments (e.g., [2,3,14]). Furthermore, some learning environments keep track of student actions to help maintain the student model (e.g., [1,13]).

However, we see the need to analyze student behavior outside a learning environment and the importance of providing feedback for these instances to support learning. In our previous work, we reported that when students study on their own, they do not only engage in goal-related tasks but also in nongoal-related activities. We reported that the digital natives who participated in our experiments in a span of two weeks spent 47.8% of their time learning and the rest not learning [9]. Furthermore, although most systems track student activities by looking at what they do on the computer, many of their activities are also done without the use of a computer [7].

Our motivation is to provide support to student metacognitive tasks and actions when learning using the computer and without it. Our students may not be constraint to a learning environment or particular domain knowledge. They may use our tool given their chosen topic of study and the tools they will use to study.

3. Experimentation

We are interested in observing learning behaviors in an environment with no guidance from teachers and where students have complete control over their learning tasks, hence, can pose as potential source of distractions. We identified foreign scholars performing their graduate research projects in Japan as good source of data since they are expected to retrieve, process, and analyze information on their own while managing their own time and effort. Since they have full control of their activities, playing games, watching videos, sending SMS and chatting can always pose distractions.

Four subjects participated in our experiments, specifically: two females and two males, who are Filipino, Indonesian, Thai and Vietnamese, respectively, with an average age of 25-years. Everyone fits as digital natives. We aimed for the sessions to be as natural as possible with the subjects deciding when to use the tool, for how many times and for how long. They were taught about the definition of goal and how to provide it to the system. For the annotation, they were taught about the definitions and specifics of intentions, activities and affect. Specific to affect, they were taught to indicate the strongest emotion in cases where they seem to have felt more than one emotion.

At the end of every session, the subjects were asked to fill up a questionnaire to help assess the influence of the tool to their reflections. We wanted to know if they discovered anything interesting about their learning behavior, whether positive or negative, during the annotation process, and based on the statistics provided by the system, whether they see the need to improve their learning behavior.

4. Results and Analyses

4.1 User Feedback on the Tool's Assistance

The subjects agreed that the tool helped point out what is both positive and negative in their learning behaviors. One subject, for example, felt good when after annotating, he realized that he actually got absorbed into work and got distracted less than he thought. He also realized that even though it was very easy for him to fall asleep when confused or frustrated, most of his time was actually spent on goal-related activities. In his last session, after previously having reflected on his past behaviors wherein he spent on the average a third of his time on nongoal-related activities, he still allowed himself to enter into nongoal-related states while making sure he spent still most of the time in learning. Another subject realized that it was a disadvantage for her to switch between tasks - that when she attempted reading three papers simultaneously, none of the goals she set for the session was completed and she gave herself a low productivity score. She also realized that when she got confused, her tendency was to switch to non-goal related tasks rather than stay engaged in learning. Yet another subject realized that he spent more time watching game video streaming and felt bad after. And one reported that standing up every now and then or drinking coffee would help stay focused. Lastly, one subject noticed that the urgency of his goals (i.e., submission deadline) "coerced" him to spend more time learning than usual, and that the lack of urgency or importance of the goal and the lack of desire to accomplish it resulted in less productive learning.

All of the above show positive effects of the tool because the students were able to reflect and monitor themselves through the annotation process. They became more aware of the activities they performed and actively tried to avoid moving away from being goal-engaged thereby learning to improve their learning behavior.

4.2 Self-assessed Productivity

As previously mentioned in the first section, students were asked to assess their learning sessions in terms of how productive they felt they were. It is interesting to note that even though the subjects performed nongoal-related activities only 24% of the time at the average, their average self-assessed productivity is only 71.3%, with the amount of time in goal-related activities not necessarily directly proportional to productivity. Table 1 illustrates this point using snapshots of all the sessions for all subjects, where $\%T_G$, $\%T_{NG}$ and *%Pr* refer to percentage of time spent in goal- and nongoal-related tasks and the percentage of self-productivity, respectively. The norm is that if a subject stayed most of the time in goal-related states, productivity assessment is supposed to be high. However, for the second sessions (S2) of subjects B and D, this did not hold. According to B, she felt that the time she spent in nongoal-related activities in S2 is higher than S1, and on that basis alone, she gave herself a lower productivity score. For D, he could have given a higher productivity score (resulting to %Pr > 80) but he felt that most of his time learning was spent thinking about the problem and not finding the solution. The implication here is that students rate themselves based on the standards they set for themselves and the tasks they set for the sessions. Our aim is that in their process of self-reflection, the students should be able to see the effects of what they impose on themselves in relation to their goals.

Table 1. Self-assessed productivity of each subject per session

	Subject A			Subject B					Subject C					Subject D		
	S1	S2	S3	S1	**S2**	S3	S4	S5	S1	S2	S3	S4	S5	S1	**S2**	S3
$\%T_G$	89	46	84	83	**76**	90	70	89	93	51	64	71	80	90	**83**	64
$\%T_{NG}$	11	54	16	17	**24**	10	30	11	7	49	36	29	20	10	**17**	36
%Pr	80	60	80	60	**40**	80	80	80	80	40	60	60	60	100	**80**	100

5. Conclusion

Although digital natives switch from one task to another as afforded by the advanced digital technologies they grew up with, it does not follow that they can be academically productive in it all the time. The tool we developed allowed our experiment subjects to realize both the positive and negative aspects of their learning behavior.

Our results are clearly preliminary. We aim to find possible generalizations of learning behaviors across learners, which would mean significantly increasing the population of our subjects as well as the amount of time we expose them to our tool. Secondly, we aim to lessen the cognitive load being imposed upon the user by the annotation process. This may mean automating to some effective degree the inference of intention and affective states.

Acknowledgements

This research work is supported in part by the Management Expenses Grants for National Universities Corporations from the Ministry of Education, Culture, Sports, Science and Technology of Japan and by the JSPS KAKENHI Grant Number 23300059.

References

[1] Aleven, V. (2010). Rule-based cognitive modeling for intelligent tutoring systems. *Advances in Intelligent Tutoring Systems*, pp. 33-62.

[2] Azevedo R. & Feyzi-Behnagh R. (2011). Dysregulated learning with advanced learning technologies. *Journal of e-Learning and Knowledge Society*, *English Edition*, *7*(2), 9-18.

[3] Azevedo, R., Johnson, A., Chauncey, A. & Burkett, C. (2010). Self-regulated learning with MetaTutor: Advancing the science of learning with MetaCognitive tools. *New Science of Learning: Computers, Cognition, and Collaboration in Education,* pp. 225–247.

[4] Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: Freeman.

[5] Craig, S. D., Graesser, A. C., Sullins, J., and 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.

[6] Ekman, P. (1999). Basic emotions. *The Handbook of Cognition and Emotion*, 45-60.

[7] Foehr, G. (2006). *Media multitasking among American youth: Prevalence, predictors and pairings*. Kaiser Family Foundation.

[8] Hartman, J., Moskal, P. & Dziuban, C. (2005). Preparing the academy of today for the learner of tomorrow. *Educating the Net Generation*, pp. 6.1-6.15.

[9] Inventado, P.S., Legaspi, R., Suarez, M. & Numao, M. (2011). Investigating transitions in affect and activities for online learning interventions. *Proceedings of 19th ICCE* (pp. 571-578). Chiang Mai, Thailand: Asia-Pacific Society for Computers in Education.

[10] Kerka, S. (2007). Study Skills. *What Works - Evidence-based strategies for youth practitioners*. Learning Work Connection, The Ohio State University.

[11] Kirschner, P.A. & Karpinski, A.C. (2010). Facebook® and academic performance. *Computers in Human Behavior*, *26*(6), 1237-1245.

[12] Luo, M. M., Nahl, D. & Chea, S. (2011). *Uncertainty, Affect, and Information Search*. Los Alamitos, CA, USA. IEEE Computer Society.

[13] Mitrovic, A. (2012). Fifteen years of constraint-based tutors: what we have achieved and where we are going. *User Modeling and User-Adapted Interaction*, *22*, 39-72.

[14] Moos, D.C. (2010). Self-regulated learning with hypermedia: Too much of a good thing? *Journal of Educational Multimedia and Hypermedia*, *19*(1), 59–77.

[15] Oblinger, D. & Oblinger, J. (2005). Is it age or IT: First steps toward understanding the Net Generation. *Educating the Net Generation*, pp. 2.1-2.20.

[16] Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D. & Strohecker, C. (2004). Affective learning - A manifesto. *BT Technology Journal*, *22*(4), 253-269.

[17] Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, *9*(5), 1-6.

[18] Roberts, G.R. (2005). Technology and learning expectations of the Net Generation. *Educating the Net Generation*, pp. 3.1-3.7.

[19] Roberts, D.F., Foehr, U.G. & Rideout, V. (2005). *Generation M: Media in the lives of 8-18 year olds*. Kaiser Family Foundation, Menlo Park, CA.

[20] Rubinstein, J.S., Meyer, D.E. & Evans, J.E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology*: *Human Perception and Performance*, *27*(4), 736-797.

[21] Schunk, D. H. (1995). Self-efficacy and education and instruction. *Self-efficacy, Adaptation, and Adjustment: Theory, Research, and Application*, 281-303.

[22] Tapscott, D. (2009). *Grown up digital: How the Net Generation is Changing Your World*. McGraw-Hill.

Variation Based Discovery Learning Design in 1 to 1 Mathematics Classroom

Euphony F. Y. YANG*, Hercy N. H. CHENG, Emily CHING, Tak-Wai CHAN
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*euphony@cl.ncu.edu.tw

Abstract: Learning retention is an important issue for instruction. However, traditional mathematics teachers tend to use direct instruction to tell students a formula, and ask them to keep in mind. But students may forget it very soon. This study adopts a variation based discovery learning strategy, which controls the similarity and difference among examples and problems to guide students to discover the critical features of mathematical concepts through personal observation and inference. A one-to-one technology-enhanced learning system is therefore designed to provide cognitive tools and scaffolding mechanisms for the guidance of discovery. To investigate the learning effect, a pre-test and a post test were conducted. The result of post test showed that the average score of the experiment group were significantly higher than that of the control group. The interview data revealed that the perspectives of some students showed higher self-confidence and learning motivation.

Keywords: Guided discovery learning, critical feature, one-to-one technology-enhanced learning, elementary mathematics

1. Introduction

The traditional mathematics education emphasizes on revealing and applying mathematical theorems. In other words, teachers tell students a formula and assign them exercises to make them keep in mind. But telling students a correct mathematical operator might hamper their ability of recall the operator [1], and also makes math a boring or daunting course.

Comparing with traditional direct instruction, education experts widely believed that discovery learning, which is different from providing complete explanation of concept in direct instruction, was an effective way of profound and lasting understanding for students' learning [9] [15] [10].Bruner [2] proposed four benefits of discovery learning better than direct instruction: (1) Growing intellectual. (2) Rewarding from the initiative discovered process and getting satisfaction. (3) Learning the "discovery" method, the capacity of analogy and independent learning. (4) Memorizing knowledge longer. McDaniel and Schlager [16] also pointed out that students in discovery learning were more able to utilize and expand knowledge. Students must discover and induct themselves.

However, only in discovery learning with proper guidance, students may learn better than in direct instruction [15]. The reason perhaps is that students tend to try aimlessly if discovery learning activities have no or insufficient guidance or students lack of precise objectives and discovery skills. Even if students are engaged in learning tasks, there are not much knowledge constructions. Furthermore, lacking clear instruction, students who have less prior knowledge are hard to get basic information and become frustration.

Besides, inappropriate guidance may force students followed the instruction one by one to complete the task; guiding too much also reduced the discovery ingredients. Therefore, guiding instance design and choices were extremely important. Related literature have discussed that discovery learning needed guidance or not (e.g. [10][15]), and what kinds of guidance were effective [5][19][7][17][18]. However, how to design and present learning materials in the discovery learning environment were limited, and the using of computer-aided discovery learning in mathematics was only sporadic [6].

This study proposes a discovery learning environment design which guide properly to promote students to find and organize the critical concepts through summarizing the mathematics text description. Therefore, this study has two research questions:

1) How is the learning effectiveness of variation-based-discovery learning?
2) Could the variation based discovery learning facilitate students' learning motivation?

2. Methods

2.1 Participants

This study held an experiment on a formal mathematics class for one year. Each week, the experiment was conducted three times, and each time lasted 40 minutes. The participants were the third grade primary students in north Taiwan. There were two groups. One was experiment group (n = 26), which used the guided discovery learning approach supported with one to one devices to learn mathematics [4]. The control group (n = 26) used traditional direct instruction.

2.2 Design Structure of Learning Activity

Many researches put into the development of guided tools in guided discovery learning [20]. Comparing with those guided tools, this study designs cognitive scaffolding tools to help students dealing with the task and go beyond their original extent [8] [11]. More specifically, the tools had the following characteristics:

1. Using “focus” to guide students discovering the critical features from examples for reducing the cognitive loading and establish the important attributes of concept [13].
2. Using learning content itself as the guidance to reduce the extra guidance and controlling the similarities and differences between examples to naturally highlight the critical features. In other words, “seeking common ground in diversity” and “seeking difference in similarities” may be easier noticed.

The design structure of learning material was divided into three parts: observation, identification, and generalization. In the observation step, the students focused on the relationship between examples and questions. The design of observation highlighted the critical point of concepts. In the identification step, students might check their possible assumption after completing the related questions. If students answer incorrect, they have to go back and observe the differences and similarities between examples and questions again; otherwise the next question would come out for students to make sure the possible manner and answer the next question on the correct base. In the generalization part, our design used short sentences and symbols as the algebra to help students describe the critical feature in a summarized statement for ensuring their mathematical concepts.

2.3 Theoretical Structure

This study based on the variation theory which was originated from Bruner’s [2] discovery learning and proposed by Marton, Runesson, & Tsui [14]. Considering the primary school mathematical concepts of relevant learning theory and the assistive role computer played in the classroom, we controlled the similarity and differences of examples and established theoretical structure to help students learn mathematical concepts. Variation theory highlighted the key point which students needed to notice and identify from the overall phenomena. It was only when students noticed and linked more than one thing change in appearance at the same time, they could experience the similarities and differences [12] [14]. As for how to change was based on four modes of variation theory: (1) Contrast: Provide positive and negative examples in the same type to compare. (2) Generalization: Present different concept facet for students extracting the general rule. (3) Separation: Change the results corresponding to the value of key property, but other properties held constant. (4) Fusion: Multiple properties changes at the same time.

2.4 Data Collection

This study considered two parts to evaluate the learning effect of variation based discovery learning: for the academic ability, learning effectiveness, pre-test in beginning of the first semester before first mathematics instruction and post test after one week of the midterms test were used and to prevent the practice effect, post-test using a parallel test and also changed the questions order and numbers; for the affection conservation, such as students' mathematics motivation. We interviewed teacher and six high, medium, and low performance students of experiment group about their learning perspective.

3. Result

3.1 Comparison of Learning Performance

Using variation based discovery learning approach to learn the concept comprehension of mathematics, students needed to observe examples, imitate examples and words summarized to explain to the learning concepts. From the result of a t test to compare the pre-test scores of this approach with the traditional teacher's direct instruction, $t_{(2, 51)}$ = 0.447, p =0.657>.05, there was no significant difference. However, after the experiment of half year, the average score of post test in direct instruction group was 63.5 points and that of variation based discovery learning approach group was 75.92 points. The average score difference of the two groups was 12.42 points. There was significant difference between two group, $t_{(2, 51)}$=-2.429, p=0.019<.05, so the learning performance of variation based discovery learning approach was better than the direct instruction.

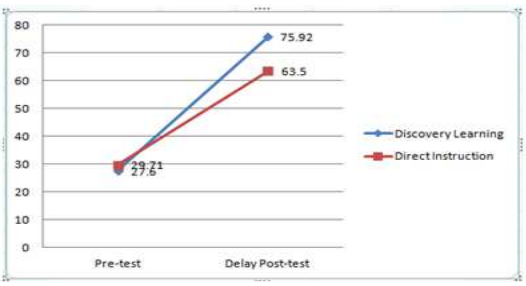


Figure 1. Pre and post-test scores of discovery learning and direct instruction

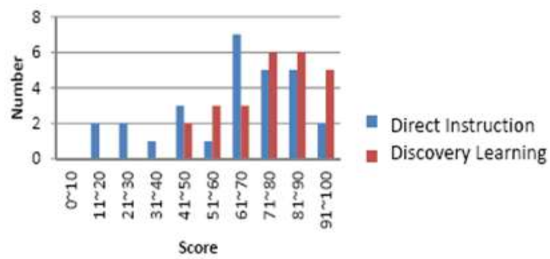


Figure 2. Post-test scores histogram of discovery learning and direct instruction

Furthermore, discovery group not only improved significantly, but their standard deviation also dropped from 18.24 points in the pre-test to 13.95 points in the post test. However, the standard deviation of direct instruction increased from 16.23 points to 21.90 points. It meant that using one to one computer supported discovery learning approach could lessen the students' learning difference.

3.2 Learning Interest

Compared to traditional teachers teaching, students of low and medium achievement preferred variation based discovery learning, and they also cultivated the concentrative and quiet habits of self-directed learning in mathematics. A medium-achieving S1 said: "*I prefer the discovery learning more. I want to think. The teacher always teaches one after another, I feel that it was too fast*". Low-achieving students generally preferred the discovery learning. A low-achieving S11 said: "*I like to see, discover and find out the key points. I do not like the teachers telling me about what is the key point, although sometimes it is difficult to discover on my own*". Students feel more solid and less pressure to control their learning speed and proceed to the next concept after they real understand the critical feature rather than pursuit the teachers' fast instruction, to hurry completing a lesson but did not real know the learning content. So discovery learning enhanced the self-confidence and accomplishment of medium and low achieving students.

But not everyone liked computer-guided discovery learning. For example, the high-achieving students had different opinions. S5 responded that he liked the teacher instruction. He said "*computer instruction was boring, there were only math problems. It always asked me to finish the math questions. Unlike teacher would give us practice after teaching*." High achieving students could follow the teacher's teaching pace and relatively had no learning difficulties in the learning process. However, the variation based discovery learning only guided through questions, which was unable to meet the students' emotional needs and wanted a real teacher to promote their learning motivation and enthusiasm.

Some students were frustrated because some questions they tried many times yet not found the critical features. Some students needed additional guidance in variation based discovery learning to successfully find the critical feature between the questions presented. The teacher said: "*1/3 children were very excited and want to surpass themselves, but 2/3 children encountered bottlenecks. In variation based discovery learning, students had no way when they couldn't pass in one stage. So they would be very frustrated and afraid of this course.*" Even if the variation based discovery learning had cognitive scaffolding and prompted students to solve problems, it still unlike real teacher who could find out students' individual problems and provide suitable instruction. The current system couldn't do so precise detections of the problem difficulties. Students couldn't pass some stage by using variation based discovery learning cause them to stay put and feel depression.

Overall, most students expected teachers' oral encourage. Some students remarked that they were used to the traditional instruction which teachers and textbooks directly told the concepts and answers, even if they could answer by observing examples from the computer and quite had achievability, but still hoped to acquire knowledge from the direct instruction.

4. Conclusion

This study focused on the design method, learning effect, and learners' interesting in variation based discovery learning in 1:1 mathematics. This study proposed a guide strategy based on variation theory to control the similarities and differences between examples to facilitate students to be aware of the critical features of concepts for further analysis, reasoning and inductive to learn knowledge. According to this strategy, this study design computer-aided function to help students follow a certain learning process and discovery activities to ensure in correct exploring direction, and provide scaffolding to facilitate students' discovery process and describe results in a summarized form. This guide strategy gives students effective guidance and is able to retain the opportunity of self-discovery for students. Therefore, this study may bring new thinking direction for the research community of discovery learning.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (NSC 101-2811-S-008-009, NSC 101-2631-S-008-003, NSC 100-2511-S-008-013-MY3, & NSC 99-2511-S-008-002-MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

[1] Bernstein, D. A., Clarke-Stewart, A., Roy, E. J., & Wickens, C. D. (2003). *Psychology*. Boston: Houghton-Mifflin.

[2] Bruner, J. S. (1961).The act of discovery. *Harvard Educational Review, 31*, 21–32.

[3] Brunstein, A., Betts, S. A., & Anderson, J. R. (2009). Practice enables successful learning under minimal guidance. *Journal of Educational Psychology, 101* (4), 790-802.

[4] 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 Journal, 1*, 3-29.

[5] de Jong, T. (2005).The guided discovery principle in multimedia learning. In R. E. Mayer (Eds.), *The cambridge handbook of multimedia learning* (pp.215-218). New York: Cambridge University Press.

[6] de Jong, T., Hendrikse, H. P., & van der Meij, H. (2010). Learning mathematics through inquiry: a large scale evaluation. In M. Jacobson, & P. Reimann (Eds.), *Designs for learning environments of the future: International perspectives from the learning sciences* (pp.189-205). New York: Springer Verlag.

[7] Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42* (2), 99-107.

[8] Jonassen, D.H. (1999). *Computers as mindtools for schools: Engaging critical thinking.* Upper Saddle River, NJ: Prentice-Hall.

[9] Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15* (10), 661-667.

[10] Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist, 41*(2), 75–86.

[11] Lajoie, S., & Derry, S. (1993). *Computers as cognitive tools.* NJ: Erlbaum: Hillsdale.

[12] Lo, M. L., Pong, W. Y., & Chik, P. P. (Eds.). (2005). *For each and everyone: Catering for individual differences through learning studies.* Hong Kong: Hong Kong University Press.

[13] Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associcated with discovery learning and worked examples. *Journal of Educational Psychology, 91* (2), 334-341.

[14] Marton, F., Runesson, U., & Tsui, B. M. (2004). The space of learning. In F. Marton, A. B. Tsui, P. P. Chik, Y. P. Ko, M. L. Lo, D. F. Ng, et al., *Classroom discourse and the space of learning*. Mahwa, NJ: Erlbaum.

[15] Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59* (1), 14-19.

[16] McDaniel, M., & Schlager, M. (1990). Discovery learning and transfer of problem-solving skills. *Cognition and Instruction, 7*, 129-159.

[17] Moreno, R. (2004). Decreasing cognitive load in novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*, 99-113.

[18] Prasad, K. S. (2011). Learning mathematics by discovery. *A Multidisciplinary Journal, 1* (1), 31-33.

[19] Van Joolingen, W. (1999). Cognitive tools for discovery learning. *International Journal of Artificial Intelligence in Education, 10*, 385-397.

[20] Varnhagen, C., Elio, R., Henry, A., Braga, J., Dyer, C., Haag, M., & Sundararaj U. T. (2009). Teaching, Research, and Discovery Learning. Retrieved from http://www.psych.ualberta.ca/~varn/Documents/TeachResearchCLEJan10.pdf.

Explaining My Solutions: An Integrated Model of Peer Tutoring for Facilitating Mathematical Communication Abilities

Hercy N. H. CHENG*, Euphony F. Y. YANG, Ying-Yen HUANG, & Tak-Wai CHAN
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*hercy@cl.ncu.edu.tw

Abstract: For elementary students, mathematical communication ability is a fundamental learning objective. Previous studies have showed that self-explanation could allow students to inspect what they have learned in the learning process, while peer tutoring could facilitate students to reorganize their learning knowledge as well as to express their idea. As a result, both could enhance students learning performance. Therefore, this study aims to design a system that may increase the mathematical communication abilities of elementary students. The study also conducts a learning activity which incorporates representation generation, self-explanation and peer-explanation in mathematical word problem solving. Preliminary evaluation shows that the integrated model may facilitate students' mathematical communication abilities.

Keywords: mathematical communication, self-explanation, peer-explanation, peer-tutoring

1. Mathematical Communication

Even if children have not learnt mathematics, they live in a world with numbers and shapes. Mathematics helps people understand the world by simplifying complex problems, solving them reasonably, and conveying the solution to other people persuasively. However, our primary education about mathematics focuses too much on problem solving and ignores the importance of mathematical communication. Mathematical communication involves adaptive reasoning (Kilpatrick, Swafford, & Findell, 2001, p. 170) and even argumentation (Andriessen, 2006).

In terms of adaptive reasoning, students have to acquire the ability to think logically, to explain a mathematical concept or procedure, and to justify their own or others' assertions. Adaptive reasoning also relates to the usage of representation (English, 1997). The ability to use appropriate representation can facilitate conceptual understanding, and problem solving. In terms of argumentation, students have to elaborate what they think, and to debate with sufficient evidences (Toulmin, 1958). When students attempt to build arguments, they aim to produce their mathematical ideas. For doing so, they may direct themselves to learn new concepts and procedures.

Self-explanation (or think aloud) is a domain-general learning strategy (Chi, de Leeuw, Chiu, & Lavancher, 1994), which emphasizes the linkage between prior knowledge and new one (Chi & van Lehn, 1991). Previous research has shown that successful problem solvers can generate more explanation (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). Compared with self-explanation, peer-explanation is an interactive explanation strategy, which can be applied in a natural and social learning environment. Among various peer-explanation pedagogies, peer instruction is a widely adopted and effective pedagogy, which allows students to explain their own ideas for reducing misconceptions (Mazur, 1997).

Furthermore, students may benefit from tutoring others (Cohen, Kulik, & Kulik, 1982; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003) as well as preparing teaching materials (Ching, Chen, Chou, Deng, & Chan, 2005). Additionally, peer teaching facilitates spontaneous and appropriate use of diagrams in order to solve mathematics word problems (Uesaka & Manalo, 2007; 2011). Therefore, this study aims to design a system to support a peer tutoring model, which integrates generating representations, self-explanation and peer-explanation for facilitating students' mathematical communication ability.

2. Activity Design

As shown in FIGURE 1, the model consists of three main phases: material preparation, peer teaching and public teaching.

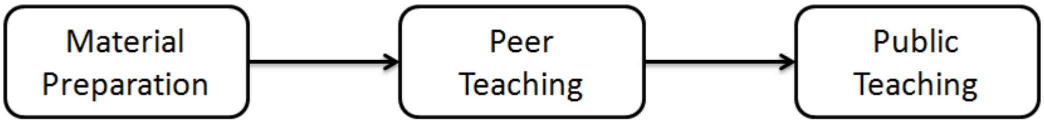


FIGURE 1. Activity Model

In the first phase, every two students are paired as a group and receive two similar but different mathematical word problems. They are told that they have to teach each other one of the two problems, and thus have to prepare their teaching materials. For doing so, they should solve their own word problems in a way to draw representations, to formulate expressions, and to calculate their answers on their own tablet PCs. They are also asked to practice their teaching by self explaining.

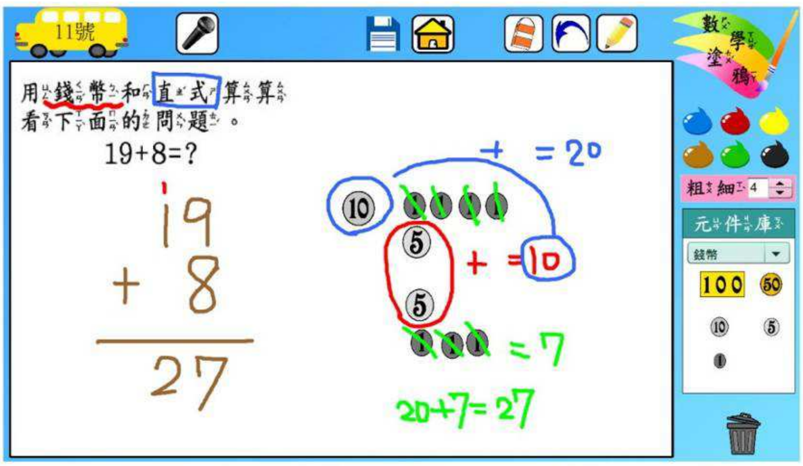


FIGURE 2. Interface

In the second phase, they reciprocally play as a tutor and a tutee. For this reason, the tutor has to teach his/her tutee how he/she solve the word problem. The tutor, more specifically, have to elaborate their representations and expressions to his/her tutee. After the elaboration, the tutee has to ask his/her tutor questions, and the tutor should answer the tutee.

In the third phase, the teacher conducts a session of public tutoring. The teacher may allow several tutor students to teach their own word problems. As the tutoring practices in the second phase, the tutor students have to elaborate their representations and expressions to the whole class. The other tutee students have to ask the tutor students questions related to the tutoring. The tutor students have to response these questions.

3. Preliminary Evaluation

3.1 Research questions

This study focuses on a main research question whether the peer tutoring model can facilitate students' mathematical communication ability. More specifically this research question can be divided into three sub-questions:

1. Can the peer tutoring model facilitate students' ability to explain self mathematical procedure?
2. Can the peer tutoring model facilitate students' ability to explain others' mathematical procedure?
3. Can the peer tutoring model facilitate students' ability to explain others' mathematical statements?

3.2 Settings

The participants were two second-year classes (N_1=25, N_2=26), in which students had similar mathematical communication abilities (see 3.3 for more details). One of the two classes was assigned as the experimental group, in which the integrated model of peer teaching was conducted for eight weeks. Another class was assigned as the control group, in which students received traditional courses of word problem solving.

The teacher in the experimental group could conduct the activity one or two sessions in a week and each session took 80 minutes. In this experiment, students were participated in the activity thirteen times in total. The materials were mathematical word problems, which involved addition, subtraction and multiplication. More specifically, in the first seven sessions, the word problems were about addition and/or subtraction, while they further involved multiplication in latter six sessions.

3.3 Measures

In this study, the dependent variable was the mathematical communication ability. For this purpose, a test on mathematical communication was conducted. This test, developed on the basis of a (Lin, & Lee, 2004), consisted of three sub-abilities: the ability to explain self procedures, the ability to explain others' procedure, and the ability to explain others' statements.

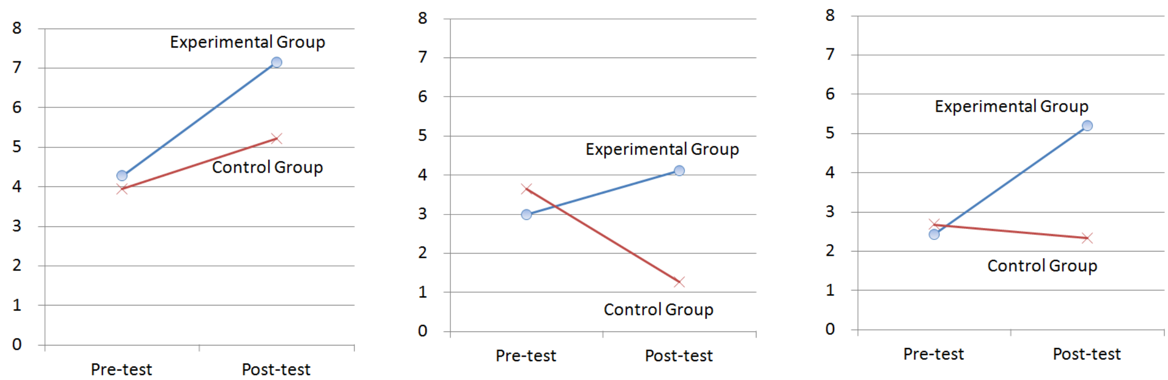
Independent t tests show that there are no significant differences between experimental and control groups in terms of explanation for self procedure ($t(49)=0.879$, $SE=0.362$, $p>0.05$), explanation for others' procedure ($t(49)=0.861$, $SE=0.759$, $p>0.05$), and explanation for others' statements ($t(49)=0.541$, $SE=0.466$, $p>0.05$). Therefore, the mathematical communication abilities of experimental and control groups are similar.

3.4 Results

FIGURE 3 illustrates the results of students' mathematical communication ability. First, in terms of the ability to explain self procedure, a two-way ANOVA reveals that there is a significant interaction between groups and time ($F(1, 49)=7.441$, $MSE=2.222$, $p<0.05$). As shown in FIGURE 3(a), both of the experimental and control group performed significantly better in the post-test than in the pre-test (the experimental group: $t(24)=7.955$, $SE=0.362$, $p<0.05$; the control group: $t(25)=2.744$, $SE=0.463$, $p<0.05$). An independent t test on the post-test further indicates that there is a significant difference between the experimental and control group ($t(49)=3.788$, $SE=0.509$, $p<0.05$). This result shows that the integrated model of peer tutoring can facilitate the ability to explain their own procedures. Although the ability of the control group increase as well, the students who have the experience of peer tutoring can improve more.

Second, in terms of the ability to explain others' procedure, a two-way ANOVA shows that there is a significant interaction between groups and time ($F(1, 49)=19.831$, $MSE=3.947$, $p<0.05$). Interestingly, while the experimental group improved significantly ($t(24)=2.206$, $SE=0.508$, $p<0.05$), the control group performed significantly worse in the post-test than in the pre-test ($t(25)=-3.990$, $SE=0.598$, $p<0.05$). The reason is probably that the pre-test asked students to distinguish and explain a wrong procedure, and the post-test asked student to distinguish and explain a correct procedure. The results may raise a further question 'can students explain more about a wrong procedure than about a correct one', which need further investigation.

Third, in terms of the ability to explain others' statements, a two-way ANOVA shows that there is a significant interaction between groups and time ($F(1, 49)=27.583$, $MSE=2.229$, $p<0.05$). Further analysis indicates that while the experimental group significantly improved their performance in the post-test ($t(24)=6.157$, $SE=0.448$, $p<0.05$), the performance of the control group did not change significantly ($t(25)=-0.892$, $SE=0.388$, $p>0.05$). The results suggest that the experience of peer tutoring can facilitate students to understand and explain others' statements.



(a) explanation for self procedure

(b) explanation for others' procedures

(c) explanation for others' statements

FIGURE 3. The effect on mathematical communication abilities

4. Concluding Remarks

This study aims to incorporate representation generation, self-explanation and peer-explanation into an integrated peer tutoring model in order to facilitate students' mathematical communication abilities. The results showed that the integrated model may significantly improve students' ability to explain self procedures, others' procedures, and others' statements. Furthermore, this study also revealed that students could explain procedures in a more complete and more contextualized way. Besides, students could generate more and more abstract representations, when they prepared teaching materials for teaching their classmates. Students were also found that they became more enthusiastic and confident about teaching in public. These findings suggested that the integrated model could help not only students' cognition, but also their affects.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (NSC 101-2811-S-008-009, NSC 101-2631-S-008-003, NSC 100-2511-S-008-013-MY3, & NSC 99-2511-S-008-002-MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

[1] Andriessen, J. (2006) Arguing to Learn. The Cambridge Handbook of the Learning Sciences, Chapter 26, Edited by Sawyer R. K., Published by Cambridge University Press.
[2] 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.
[3] Chi, M. T. H., de Leeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.
[4] Chi, M. T. H., & van Lehn, K. A. (1991). The content of physics self-explanations. The Journal of the Learning Sciences, 1(1), 69-105.
[5] Ching, E., Chen, C. T., Chou, C. Y., Deng, Y. C., & Chan, T. W. (2005). P^3T: a system to support preparing and performaing peer tutoring. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds), Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology. (pp. 768-770). Amsterdam, The Netherlands: IOS Press.
[6] Cohan, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: a meta-analysis of findings. American Educational Research Journal, 19(2), 237-248.
[7] English, L.D. (Ed.). (1997). Mathematical Reasoning: Analogies, Metaphors and Images. Lawrence Erlbaum & Associates.
[8] Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). Adding it up: Helping children learn mathematics. Washington, DC: National Academies Press.
[9] Lin, Y. H., & Lee, Q, Y. (2004). A study of ability on mathematics communication for pupils of elementary schools. Journal of Research on Measurement and Statistics, 12, 233-268. (in Chinese)
[10] Mazur, E. (1997). Peer instruction: a user's manual. NJ: Prentice Hall, Upper Saddle River.
[11] Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: a meta-analytic review. Journal of Educational Psychology, 95(2), 240-257.
[12] Toulmin, S. E. (1985). The uses of argument. Cambridge: Cambridge University Press.
[13] Uesaka, Y., & Manalo, E. (2007). Peer instruction as a way of promoting spontaneous use of diagrams when solving math word problems. In D. McNamara & J. Trafton (Eds.), Proceedings of the 29th Annual Cognitive Science Society (pp. 677-672). Austin, TX: Cognitive Science Society.
[14] Uesaka, Y., & Manalo, E. (2011). The effects of peer communication with Diagrams on students' math word problem solving processes and outcomes. In L. Carlson, C. Hoelscher, & T. Shipley (Eds.), Proceedings of the 33rd Annual Conference of the Cognitive Science Society (pp. 312–317). Austin, TX: Cognitive Science Society.

The Effects of STEAM by RST Instruction Using Writing on Elementary School Student's Problem-Solving Skills

SuRyun Jeon[a*], SunHwoi Kim[b] & DoongSoo Nam[c] &TaeWuk Lee[d]
[abcd]*Dept. Computer Education, Korea National University of Education, Korea*
*ynot1178@nate.com

Abstract: Problem- solving skills are mandatory for everyone in the 21st century. The goal of STEAM also is to increase the problem-solving skill based convergence thinking. So, in this study, I proposed a study program of the STEAM by RST(Reverse Science from Technology) for improving problem-solving skills. The topics are hi-tech products which are the convergence product of Science, technology, culture and art like Refrigerators and mobile phones. This program is starting from the exciting technology and conversely learning basic science, mathematics, art and etc. That learns technology in high-tech products which were subject of curiosity, and then learn mathematics and science, which are the basis of the technology. So, learners will learn how to use basic studies to their life and it increases understanding of course content as well as the interest in learning. Also, that provides learning opportunity of Language art by utilizing writing in the learning process and learners can recognize and solve problem on their own to found and modify their thinking.

Keywords: Problem-Solving Skill, STEAM, RST, Convergence Education, Writing

Introduction

The topic of education in today's information society is the problem-solving skills. Problem-solving skills are mandatory for everyone in the 21st century by Partnership for 21st century skills, DeSeCo(Definition and Selection of Key Competencies) report of OECD, SCANS(Secretary's Commission on Achieving Necessary Skills). The complex problems in 21st century's information and knowledge-based society require problem-solving skills based convergence thinking. STEAM is also aimed at improving problem-solving skills based convergence thinking. STEAM is convergence with Science, Technology, Engineering Mathematics and Art. This study proposed a study program of the STEAM by RST(Reverse Science from Technology) using writing for improving problem-solving skills. This program is starting from the exciting technology and conversely learning basic science, mathematics, art and etc. The topics are hi-tech products which are the convergence product of Science, technology, culture and art like refrigerators, mobile phone, etc. These products are used very frequently in everyday life of learners. Furthermore, the program provides the opportunity for reflective thought by using writing in learning process. So, learners can recognize and solve the problem by themselves

1. STEAM by RST

STEAM is one of convergence study, an advanced concept of STEM. STEM is a convergence with Science, Technology, Engineering, and Mathematics. STEAM is adding =to Art to STEM. This Art means Fine Art, Liberal Art, Language Art and so on. [Figure 1] is the STEAM Pyramid. RST is an abbreviation of the 'Reverse Science from Technology'. STEAM by RST teaches the principles of high-tech products such as mobile phones, TV, car, etc and integrated system. That increases learning interest and achievement by understanding how basic science is used in every life. Furthermore, learners can recognize the truth how to increase the value of the product by design and User Interface.

2. Problem- Solving Skills

General Problem-solving skills are procedural knowledge which is procedure and method. These can be generally applied to solve a widespread and various problems in the real world. Creative problem-solving skills refer to every kind of process that an individual or a group thinks creatively to solve a certain problem or to such efforts. These are concepts which emphasize process than outcome. So, the education for improving problem-solving skills should be focused on the process of solving the problem.

The common process of problem-solving by OECD 2004, Polya can be summarized as follows: First, understand and analyze the meaning of the problem. Second, solve the problem. Third, reflect and evaluate the ideas. So, this program was designed that learners proceed the process of problem-solving and recognize and solve the problem by themselves.

3. Writing

The process of writing is very similar to the process of solving the problem. Many subjects such as Mathematics, Science, Social Studies, Art, etc focus on these similarities and use actively the writing. Based on these previous studies, the educational effects of writing can be summarized as follows: First, writing is a tool for finding problems that is beginning of solve the problem. Second, writing provides the opportunity for reflective thinking. So, learners can be helped to solve the problem by themselves. Third, Learners are to participate actively in their learning. Fourth, Effective communication is possible by writing. So, learners can review the idea of solving problems with peers. Fifth, writing provides the opportunity to think deeply about the representation of the target. That can be useful STEAM education including art.

4. Methodology of the STEAM by RST using Writing Study Program

4.1 Directivity and select content of Study program

- Prepare standards and emphasis on STEAM by RST, writing, the process of problem-solving.
- Consider elementary school student's interest study level, learning ability.
- Choose the theme which learners can meet easily in real life

4.2 The concept model of STEAM by RST using writing

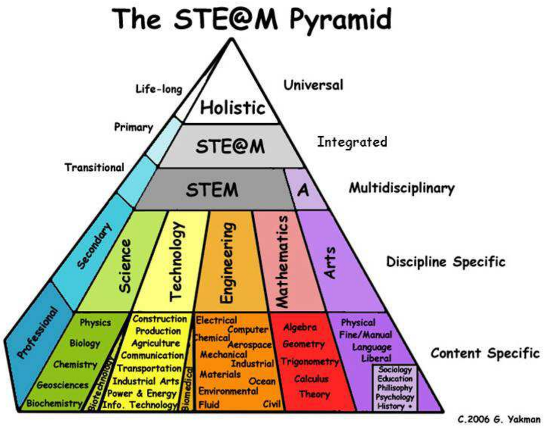


Figure 1. STEAM Pyramid

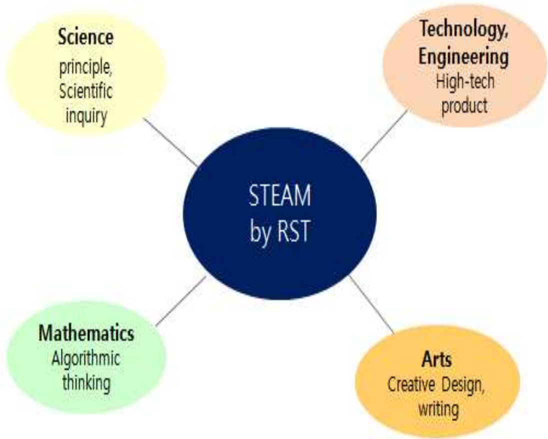


Figure 2. Concept model of STEAM by RST using writing

4.3 Design the model for teaching

Table 1. The model for STEAM by RST using writing and content

Polya's problem solving	STEAM by RST	Writing Content
Understanding the problem	Presenting the problem	- Question and answer for Motivation
	Analyzing the problem	- Draw a picture of the subject and write the role of the component - Compare with peers and modify
Devising a plan	Devising a plan	- Write experience about the topic - Explorer and write scientific principles on the topic - Write the uncomfortable points during using the topic and writing solution
Carrying out the plan	Carrying out the plan	- Write algorithm flowchart - Discuss the scientific principles on the topic - Creative Design
Looking back	Looking back	- Presentation - Review the idea

5. Conclusion and Future Work

In this study, we developed the STEAM by RST using Writing study program to improve problem-solving of elementary school. This program could increase learning interest and problem-solving by hi-tech products. In the learning process, writing will help learner's reflective thinking. To inspect the effects this program, we have to apply this method to real class. Developed STEAM by RST program will be taught to students in class and we'll observe the student's learning activities and analyze outcomes of learning.

References

[1] OECD. (2004). Problem solving for tomorrow's world : First measures of Cross-Curricular Competencies from PISA 2003.
[2] The Secretary's Commission on Achieving Necessary Skills. (1991). What work Requires of School.
[3] Gorgette Yakman. (2007) STEAM Education : an overview of creating a model of integrative education.
[4] Polya, G.(1957). How to solve it.
[5] L. Flower. (1998). Problem solving Strategies for Writing in College and the Community.
[6] Kwang-Hyung Lee. (2011). http://www.sciencetimes.co.kr/article.do?atidx=0000048348&todo=view

Facilitating Story Revision Through Peer Questioning on a Scribble Environment

Siou-Lan WANG*, Calvin C. Y. LIAO, Tak-Wai CHAN
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*shirley@cl.ncu.edu.tw

Abstract: For the past 1 1/2 years, eight second-grade Taiwanese classrooms (n = 207) have used *BookTalker* to draw pictures and write stories. Although, this allows for easy sharing of one's story, students still seem to lack an understanding of what constitutes a "quality" story or how to self-improve. The study proposed in this paper aims to extend the use of *BookTalker* to enhance the quality of students' stories by incorporating scaffolding through the process of asking questions regarding the content of peers' stories, critical analysis of asked questions, and story revision.

Keywords: story writing, peer questioning, drawing, writing revision, scaffolding

1. Introduction

In traditional elementary school Chinese writing classrooms, students passively execute writing tasks assigned by teachers with the goal of merely turning in written homework assignments containing bits of knowledge acquired in past lessons. To approach this problem, interactive peer questioning on a digital environment is considered as a strategy to increase students' analytical skills in order to improve their written Chinese. For the past 1 1/2 years, eight second-grade Taiwanese classrooms (n = 207) have used *BookTalker*, an application of "Crazy Brush" [3] developed to allow students to create stories through drawing and writing. The main purpose of the proposed study is to facilitate students' story revision through peer questioning on the scribble environment "BookTalker." In addition, this study attempts to observe specifically what about peer questioning enhances story quality.

1.1 Sharing and Revising Stories

In traditional Chinese classrooms, computers and the Internet provide potentials of making such stories meaningful and interactive. By drawing and writing on the computer and then sharing these creations online with classmates, more interactive opportunities can be created. In the present study, the researchers present a scenario in which students need to solve problems (questions) generated by peers.

1.3 Peer Questioning

Several studies have been carried out investigating question generation [1, 2, 6]. Saddler [5] reports that students with learning disabilities increased their story-writing ability through self-regulated strategies. Although much interaction took place in the previously reviewed studies, but do to the fact that this interaction was conducted face-to-face, this interaction was extremely limited. Hence, the study proposed in this paper provides students with an extra period of time for interaction through the online-based peer interaction afforded by *BookTalker*.

2. Peer questioning to enhance story quality

Based on previous research, we designed a three-session activity to stimulate students' peer questioning: (1) Writing a story; (2) Questioning peers' regarding their stories; and (3) Revising one's own story. In session one, students' literacy skills will be improved through writing practice during the story creation process. Besides, students will be able to gain the knowledge of the elements that constitute a good story. In session two, students are empowered to judge and give comments about peers' stories. In session three, revising stories enables students to gain ownership of their stories. At the same time, their literacy skills can be improved by revising errors.

The teacher will first introduce the system and all "BookTalker" functions including how to draw, write, give titles, submit a story, choose another's story to ask questions, submit questions, and revise one's own story (see Fig. 1). Then the teacher will illustrate examples of the seven elements that generally occur in a good story and how the peer-questioning task should be completed through the use of a sample story included in *BookTalker*. Afterwards, an entire class peer questioning story practice using a second sample story from *BookTalker* will be conducted to confirm if students can take story-specific questions into practice. In addition, an instructional video and an image laden textual description as references are available to students while interacting on *BookTalker*.

Fig. 1 BookTalker Drawing Interface

First, students choose and read two classmates' stories on *BookTalker*. The drawings and writing associated with a story is unavailable for preview. Instead students are given a list of story titles to choose from. Once a story is chosen, the story writer's name will be appended to the story title and the story will be removed from the list of available stories; the story writer's name is shown after selection to aid face-to-face interaction that takes place during the story revision process.

Second, students are able to propose questions below each picture or on the webpage that presents all four pictures of the story. Each student is asked to propose at least two questions for each story chosen. A five-minute countdown begins when a student starts the process of posing questions on a classmate's story; if two questions have not been generated after the five minutes are up the student will receive links to the instructional video and reference material as well as the question prompts. After submitting two questions for the first story *BookTalker* will then send the student to the second story where the process is repeated. After all students have proposed questions regarding two classmates' stories, each student revises one's own drawing or writing according to peers' questions. Before revision, students must determine if the questions contribute to the revision. Students decide which questions to take into consideration or to debate with the question givers face-to-face until a consensus is reached. After the story is revised, individuals resubmit through *BookTalker*.

3. Discussion

The purpose of the proposed study is to develop a three-session activity to guide students in generating questions regarding peers' stories in order to improve students' Chinese writing and story creation. Students are expected to improve from three different perspectives; they are: social, psychological, and cognitive.

Social perspective: Empowerment in the classroom. In this study, the role of reviewer shifts from teachers to peers. With this empowerment [8], students' performance in story writing is expected to improve allowing students to produce more authentic creative quality stories.

Psychological perspective: The Ownership of work. Creating and revising stories is not merely to meet teachers' requirements, but also to foster ownership [7]. During the processes of generating questions and revision, students become deeply involved in their own and peers' story creation to not only improve story quality but also analytical skills.

Cognitive perspective: The practice of literacy skills. Compared with explicitly teaching students the knowledge of literacy skills, a meaningful writing task of creating and revising stories enables students to take these skills into practice [4]. Peers' questioning plays the role of facilitating the process of story revision.

The progress of this proposed study so far has been the construction of *BookTalker*. The upcoming work is to run the experiment in a third-grade Chinese writing classroom as a pilot.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (NSC 101-2811-S-008-009, NSC 101-2631-S-008-003, & NSC 99-2511-S-008-002-MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

[1] André, M. E. D. A., &Anderson, T. H. (1978). The Development and Evaluation of a Self-Questioning Study Technique. *Reading Research Quarterly, 14*(4), 605-623.

[2] King, A. (1990).Enhancing Peer Interaction and Learning in the Classroom Through Reciprocal Questioning. *American Educational Research Journal, 27*(4), 664-687.

[3] Lee, Y. C., Liao, C. C. Y., & Chan, T. W. (2010). Crazy Brush: Designing the Scribbles Environment to Improve Children' Interest in Writing In S. L. Wong et al. (Eds.), Proceedings of the 18th International Conference on Computers in Education (pp.73-75).

[4] MacArthur, C. A., Graham, S., Schwartz, S. S., & Schafer, W. S.(1995). Evaluation of a Writing Instruction Model That Integrated a Process Approach, Strategy Instruction, and Word Processing. *Learning Disability Quarterly, 18*(4), 178-291.

[5] Saddler, B. (2006). Increasing Story-Writing Ability through Self-Regulated Strategy Development Effects on Young Writers with Learning Disabilities. *Learning Disability Quarterly, 29*(4), 291-305.

[6] Singer, H., & Donlan, D. (1982). Active Comprehension: Problem-Solving Schema with Question Generation for Comprehension of Complex Short Stories. *Reading Research Quarterly, 17*(2), 166-186.

[7] Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44*, 331-349.

[8] Zimmerma, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70.

Three Methods of Applying Digital Technologies to Enhance Teaching and Learning Traditional Chinese Paintings

Shih-Lun CHEN, Yuan-Hsun CHUANG
Faculty Director, China University of Technology, Taiwan
Assistant Professor, China University of Technology, Taiwan
ricead@gmail.com

Abstract: Digital technologies have enormous impacts on all aspects of art education. By applying powerful computer technologies in the classroom, various activities become possible for introducing traditional Chinese painting, which is generally considered dreary, pedantic and too far removed from modern life by young students today. In this research, three methods that were applied in a fine art class, especially in primary school and secondary school, are presented, to describe how digital technologies can be adopted for relevant pedagogic practices. The three methods include three-dimensional animation, digital storytelling and interactive modes. It is concluded that these three methods are powerful and meaningful ways to enhance teaching and learning traditional Chinese paintings for students of the digital generation.

Keywords: Digital technology, fine art education, pedagogic practice, Chinese painting

1. Previous Works

Digitalizing oriental fine art such as traditional Chinese painting and calligraphy has accomplished a remarkable evolution in the past two decades, and has been discussed intensively. From the pioneering projects of Saito & Takahashi, Haeberli, Hertzmann, and Freeman et al., systems of generating, rendering or translating painterly brush works were invented and introduced (Saito & Takahashi, 1990; Haeberli, 1990; Hertzmann, 1998; Freeman et al., 2003). Many computer devices such as tablets and styluses, and software, are able to simulate or generate vivid brush works (Xu et al., 2008). From translating, simulating, generating to animating, technology did digitalize traditional Chinese painting, and utilized it mostly in the fields of art creation/recreation, commerce or entertainment. Now people are surprised to see amazing animations with the style of Chinese painting and enjoy the oriental aesthetic brush works that often appear in daily life via web pages, MTV, TV commercials, various graphic designs and so on. However, digital applications, particularly for traditional Chinese painting, still lack attention and need further research in respect of education.

2. Research Approaches

In order to assist educators to apply appropriate digital technology for pedagogic practices, three methods are proposed in this section. Each method solves issues of the modern audience's difficulty with understanding traditional Chinese Painting. Additionally, many concrete examples are given, to introduce how the methods address the issue.

2.1 Method One: Three-dimensional animation

Traditionally, most Chinese painting is composed on two-dimensional hanging scrolls, both vertical and horizontal formats, and the spatial effects are more subtle and metaphorical, rather than the sharp and dramatic effects of Western painting. However, the students of the digital generation are more familiar with realistic or three-dimensional space and with moving images, due to the influence of digital media such as TV and video games.

Also, the Qi, atmospheric phenomena throughout the painting, seems too abstract for modern students to comprehend. Therefore, it is very difficult to invite students to place themselves in the picture and appreciate the scenes, stories and aesthetics that the art work conveys. To address this issue, the techniques of modeling, layering and animating were utilized to transform two-dimensional Chinese painting into three dimensions with moving actions. In three-dimensional images, it is easier for teachers to interest students in the picture and to explain the theme. Three-dimensional animations also allow viewers to go behind or around the objects to look closely, if the brush works are too abstract to understand. For the philosophical images of Qi and empty space in paintings, it is more possible for students to experience the mystical aesthetics if they can "play around inside the picture" just like they are used to doing in computer games

2.2 Method Two: Digital storytelling

Many traditional Chinese paintings illustrate historical stories or persons, and mark the content with calligraphic scripts in the painting. However, as time goes on, new generations are no longer familiar with those events and with handwriting; thus, they have difficulty in accessing them. To address this issue, multimedia technology was utilized to add story lines, conversations or role-plays into paintings. Students were able to understand the background, history and story behind the art more directly and effectively. Some traditional paintings lend themselves to plots about popular and current issues that are especially applicable to the daily life of modern times. Figures 1 shows a digital application that adds a story line about cosmetic skills to introduce a traditional portrait.

Figure 1

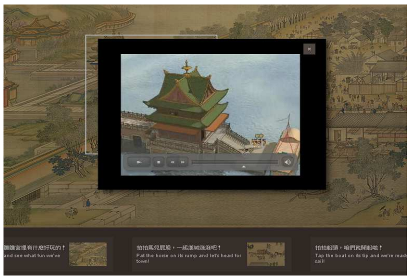


Figure 2

2.3 Method Three: Interactive mode

In traditional times, Chinese painting was created and appreciated by scholars and gentry, and presented elitist art forms that conveyed cultural, historical, political and philosophical values, as well as artistic skills. It is very challenging for modern people without such an elite background to fully and truly understand this art. It would not be too much of an exaggeration to say that to study each traditional painting requires a certain level of knowledge, a knowledge bank.

To address this issue, an interactive function is invented for traditional Chinese painting, that served as a powerful database. With the interactive function, students can actively browse the parts of painting that they are interested in, zoom in for details, react with the images, or link to more instructions for relevant information. These kinds of interactive painting were publicly exhibited recently, and received tremendous feedback. Audiences found it was great fun to play with the art instead of just looking at it (CNN, 2010). The interactive version of “Along the River During the Ching-ming Festival” *(*Figure 2) is a great example of digitalizing traditional painting. It is a portrayal of prosperity and affluence along the Bian River in Kaifeng, capital city of the Northern Sung dynasty. This collaborative piece painted in 1736 measures 35.6 cm in height and 1152.8 cm in length. The meticulously rendered painting features unique customs of the Ming and Ch’ing dynasties, providing a wonderful first-hand account of contemporary lifestyles. In the interactive version, viewers can browse the entire image and select a section for more detail, which comprises three-dimensional animations of the architecture, scenes and people.

3. Conclusion and future work

Digitalizing traditional Chinese painting has made tremendous progress in the past two decades and has been utilized for graphic design, game design and many other entertainment functions. Now, how to utilize it in classrooms, enrich pedagogic practices and consequently benefit the teaching and learning of Chinese paintings, is the main concern. Three methods of digital applications have been introduced and summarized in this research. All of these applications are amazing in visual effects and received very positive feedback from audiences in exhibitions. They are also relevant and appropriate pedagogy because each method addresses the issue of the lack of attractiveness of traditional Chinese painting for modern students. Furthermore, while it is proposed that these three methods of digital application be used to enhance the teaching and learning of traditional Chinese painting, further research is needed to quantify the effectiveness of digital applications for students in different age ranges and cultural backgrounds. Also, questionnaire of students can be used for further research to get detail feedbacks from students regard the effectiveness of the proposed methods.

References

[1] C. Chan, E. Akleman, and J. Chen, Two Methods for Creating Chinese Painting. *Proceedings of Pacific Graphics'02*, Beijing, P. R. China, 2002.

[2] J. Lambert, Digital Storytelling: Capturing lives, creating community. *Berkeley, CA: Center for Digital Storytelling*, 2002.

[3] N. J. Livo & S. A. Rietz. *Storytelling: Process and practice.* Littleton. CO: Libraries Unlimited, 1986.

[4] D. Meadows, *Digital storytelling: Research-based practice in new media.* Visual Communication, *2*(2), 189–193(2003)

[5] J. Silbergeld, *Chinese painting style: media, methods, and principles of form* (Seattle and London: University of Washington Press, 1982).

[6] S. H. Xu, F. Lau, and Y. Pan, *A Computational Approach to Digital Chinese Painting and Calligraphy.* (Hangzhou: Zhejiang University Press and Berlin, Heidelberg: Springer 2009).

[7] W. Cohn, *Chinese Painting* (New York: Hacker Art Books 1978).

[8] Kozma, R. B. (1991). Learning with media. *Review of Educational Research*, *61*, 179–211

[9] Roxana Moreno, Alfred Valdez, Cognitive Load and Learning Effects of Having Students Organize Pictures and Words in Multimedia Environments: The Role of Student Interactivity and Feedback. *ETR&D*, Vol. 53, No. 3, 2005.

[10] CNN NEWS, Queue for hours to get tickets to “Along the River During the Qingming Festival”: http://www.cnngo.com/hong-kong/life/200000-tickets-sold-out-first-day-glimpse-along-river-during-qingming-festival-666493

[11] Ukiyo-Hakkei, ODD JOB. http://www.youtube.com/watch?v=WDdlEj6UUNQ

[12] Happiness Two. Lee Lee Nam. 2012. http://logoartproject.naver.com/artist5

[13] Happiness Four. Lee Lee Nam. 2012. http://logoartproject.naver.com/artist5

Using Self-generated Drawings to Support Writing and Storytelling in Language Learning

Calvin C. Y. LIAO*, Yi-Chian LEE, Maureen WU, Tak-Wai CHAN
Graduate Institute of Network Learning Technology, National Central University, Taiwan
*Calvin@cl.ncu.edu.tw

Abstract: Previous study indicated that self-generated drawings strategy possible will play a key role on students' learning to speak and write from emergent literacy to conventional phase. Hence, this study developed a self-generated drawings environment to support young children writing and storytelling. We implemented an activity integrated of drawing, writing, and storytelling. In particular, each student used the tablet PC to share a story with drawing and writing product. Besides, some implications about three future directions were also discussed.

Keywords: self-generated drawings, writing, storytelling, language learning

1. Introduction

In the last three decade, the relationship between drawing and writing has been explored in relation to children's literacy development (Caldwell, & Moore, 1991; Norris, Mokhtari, & Reichard, 1998). Some studies have suggested that student-generated drawings can supported for a variety of learning activities (van Meter, & Garner, 2005). In particularly, these suggestions included construction of knowledge representation (van Meter, & Garner, 2005), preparation for narrative writing (Caldwell, & Moore, 1991), and pre-writing strategy (Norris, Mokhtari, & Reichard, 1998). On the other hand, for a long time, oral narrative is considered important for interpersonal interaction and social communication (Vygotsky, 1986). Some studies considered that the oral narrative about the drawing and writing could scaffold children's topical understanding, such as using drawings to planning and discussion (van Meter, & Garner, 2005). That means that oral narrative about the drawing and writing allow children to share more detailed and accurate information with each other; they become acutely aware of the contrast between the pre-drawing-and-writing and the post-drawing-and-writing. In particularly, children determine whether modify their original illustrations, recreate new drawings, or add new text that more accurately reflects their drawing and writing. Above learning activities mean that student's self-generated drawings strategy have potential value for oral language and written language of primary children. In other words, self-generated drawings strategy possible will play a key role which students learn to speak and write from emergent to conventional phase. By contrast, few empirical studies of drawing construction on speaking and writing aspect have been systematic implementation in formal school, and much remains unknown about this learning process. Moreover, children in Taiwan lack the opportunity to practice oral narration and express their opinions (Chang, & Ku, 2008). Norris *et al.* (1998) also indicated that there has been limited formal study about the role of drawing in the writing process of children in primary grades 1-3. If we could combine drawing, writing, and storytelling which will be more meaningful for children's language learning. Hence, this study attempts to support students drawing, writing, and telling their products in order to improve the writing skills and storytelling abilities of children.

2. Through Self-generated Drawings Strategy to Support Young Children Writing and Storytelling

Previous study developed an on-line drawing environment to support young children writing and to share their products, entitled Crazy Brush (Lee, Liao, & Chan, 2010). We implemented a drawing, writing, and storytelling activity. These activities in the digital drawing process can simplify automatic sketch segmentation, which can be used to support students in creating products from drawings. In particular, each student used the tablet PC to share a story with drawing based on text of writing. We hope the Crazy Brush system could let students draw their scribble and have fun, and meanwhile students would write a story about their scribble. In other words, students would find the interesting element from the process of drawing their scribble and transfer into writing, and presenting their products. The main sources of products include two ways: one is reading and the other is life experience. The Crazy Brush system would provide a writing channel and students could write a story through describing others' scribble. Through the system, students could mutually share ideas and maybe they could get more creative inspirations and ideas which will not be limited to reading or experience. Crazy Brush system utilizes three strategies to enhance students' learning and motivation. We hope system design is flexibility and it could support various activities, whether in the classroom or after school. The Crazy Brush system function contains three parts: *drawing*, *writing*, and *portfolio (for storytelling).*

2.1 Drawing, Writing, and Storytelling

Drawing: The system provides the most basic functions, for example: brush, eraser, color, and etc., see Figure 1. Using the eraser tool can easy clear unwanted drafts and the undo tool can quickly return to the previous step. Paper cannot provide these functions. The button design principle is intuitive which could avoid children's cognitive overload. Students could scribble through system function to name their scribble. The system will also investigate creator the sources of ideas: (1) imagination, (2) their own experience, (3) reading experience, and (4) other. After the completion of product, the students could choose whether to share their scribble picture with others and the scribbles could become one of storytelling resources.

Writing: Students have to write a story in light of choosing their own scribble. Students can use keyboard input words or handwrite on the screen. Students have to input a story topic and the system will search related words. Students can write a story using related words. In this part, the system provides video and sound recording functions and students could telling a story according to their write story. After story writing or telling, the students could choose whether to share with others and to publish to portfolio.

Storytelling: The portfolio function would collect students' scribble and story, and students could review all their products. They could press the button and watch others' creative writing or speech. Paper cannot record students scribble process, but the system provide this function. In portfolio, students can see others scribble process, and they can learn from each other. In addition, the system also provides recommend feature, students can vote their favorite creation. We hope to make writing more fun and let students do the initiative writing and speaking.

2.2 *The Design of Activity Flow: Drawing, Writing, and Storytelling*

This section describes the Crazy Brush system how to use in the classroom learning activities. In the activity, teachers provide certain topic of activities. Through these topics, students need to paint and write a story, see Figure 1. The learning activity includes 3 steps. Step 1: students need to think and organize the ideas in mind, and then draw pictures. Step 2: students need to describe and explain these pictures, and then write down some scripts according to previous ideas. Step 3: students need to share and present these scripts and pictures with others.

Figure 1. Drawing, Writing, and Storytelling in Crazy Brush.

3. Remarks

This study developed a self-generated drawings environment to support young children's writing and storytelling. We implemented a drawing, writing, and storytelling activity. We come up with three future directions in which deepening the *creation*, *sharing*, and *assessment* of story making could be implemented. First, we should focus on story creation in order to increase the content and structure of story. Secord, we should focus on story sharing in order to provide the opportunity and interaction of students. Finally, we should focus on story assessment in order to enhance the sentence usage and article structure.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support (NSC 101-2811-S-008-009, NSC 101-2631-S-008-003, & NSC 99-2511-S-008-002-MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan. In addition, the authors would like to be grateful for suggestions and comments provided by Wan-Chen CHANG.

References

[1] Caldwell, H., & Moore, B. H. (1991). The Art of Writing: Drawing as Preparation for Narrative Writing in the Primary Grades. *Studies in Art Education*, *32*(4), 207–219. doi:10.2307/1320873

[2] Chang, W. C., & Ku, Y. M. (2008). Children's oral narrative evaluation and assistance. *Journal of Education Research*, *173*, 74-84.

[3] Lee, Y. C., Liao, C. C. Y., & Chan, T. W. (2010). Crazy Brush: Designing the Scribbles Environment to Improve Children' Interest in Writing. In S. L. Wong et al. (Eds.), Proceedings of the 18th International Conference on Computers in Education (*pp.73-75*). Putrajaya, Malaysia: Asia-Pacific Society for Computers in Education.

[4] Norris, E., Mokhtari, K., & Reichard, C. (1998). Children's use of drawing as a pre-writing strategy. *Journal of Research in Reading*, *21*(1), 69–74. doi:10.1111/1467-9817.00043

[5] van Meter, P., & Garner, J. (2005). The Promise and Practice of Learner-Generated Drawing: Literature Review and Synthesis. *Educational Psychology Review*, *17*(4), 285–325. doi:10.1007/s10648-005-8136-3

[6] Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Mobile Learning Numeracy in Ontario

Serene Jin Yan HE, Maiga CHANG, & KINSHUK
School of Computing and Information Systems, Athabasca University, Canada
*maiga@ms2.hinet.net

Abstract: The growth of mobile technologies and devices has penetrated the world substantially over the last 5 years, and a large percentage of these users are young users. This introduces profound changes to traditional teaching practices which may initially be seen as low-key, but it will eventually create visible side effects in the education domain that cannot be ignored. As we push for effective instruction to be differentiated, we can use mobile technologies to enhance learning. This research will present the mobile device functions suitable for mobile learning and map it to learning topics of Ontario's K-6 Numeracy Curriculum.

Keywords: Mobile learning, instructional strategy, e-learning, Numeracy curriculum, grade K-6, Ontario Ministry of Education

Introduction

Recent research is more focused upon mobile learning and the huge potential it has in the language learning field (Mahruf, Shohel and Power, 2010), it also enhances student learning of mathematics (Rains and Clark, 2011). Although there are many researches on mobile learning, the educational value in the use of mobile technology is still limited (Morgan, 2010) there is no standard or model that formally addresses these relationships. The objective of this research will seek to identify which mobile technologies are most suitable to be integrated in instructional strategies to deliver knowledge and content.

1. Ontario's K-6 Numeracy Curriculum and Instructional Strategy

The numeracy domain has always been an area of strong focus because it an essential part of our life. The learning goals of numeracy as stated by the Ontario Ministry of Education (2003) is to achieve beyond computation skills, instead students are to be able to solve problems, process information from various sources and access and use quantitative information to make knowledgeable decision. Table 1 maps out the learning topics within the domain of numeracy according to the curriculum set up by the Ontario Ministry of Education (2003, 2008a, 2008b, 2008c, 2008d).

Table 1. Numeracy Learning Topics

Learning Domain	Learning Topic
Numeracy	Counting, Operational Sense, Quantity, Relationships, Representation, Geometry and Spatial Sense, Data Management and Probability, Measurement, Patterning/Algebra

Ontario MoE (2003) identified that effective instruction is crucial to the successful learning of mathematical knowledge and skills. There are five categories of instructional strategy: Direct instruction, Indirect instruction, Experiential learning, Independent study and Interactive instruction. These five categories are inter-related as the instructional mode can be used in more than one strategy. Table 2 shows a snippet of the instructional strategies and its instructional modes.

Table 2. Instructional Strategies and its Instructional Mode (Partial)

Instructional Strategy	Instructional Mode
Direct Instruction	Structured Overview, Lecture, Demonstrations, Guided and Shared
Indirect Instructions	Problem Solving, Case Studies, Reading for Meaning, Cloze Procedure
Experiential Learning	Field Trips, Narratives, Conducting Experiments, Simulations

2. Discovering Relations

Educators recognize the potential of mobile technologies as a learning tool in education. We need to establish how these mobile device functions relate to instructional modes and how the relationship is useful to educators. There are different instructional strategies with its own set of instructional modes that can overlap with each other. Mobile device functions can be used to assist in the delivery of these instructional modes. New and improved devices hardware capabilities, speed and innovative program applications using Web 2.0 technologies (e.g., wikis, Twitter, YouTube, MySpace) have made mobile devices more promising as an educational learning tool (Park, 2011).

Educators can deliver the learning topics like measurement using GPS to see distance traveled from point A to B, or using GPS to calculate length of time before reaching destination. Students can use the camera to capture a chosen location with human traffic during different times of the day as part of a data management topic. Figure 1 below demonstrates the mapping concept for the relation between Learning Topics, Learning Styles, Instruction Mode and Mobile learning mode.

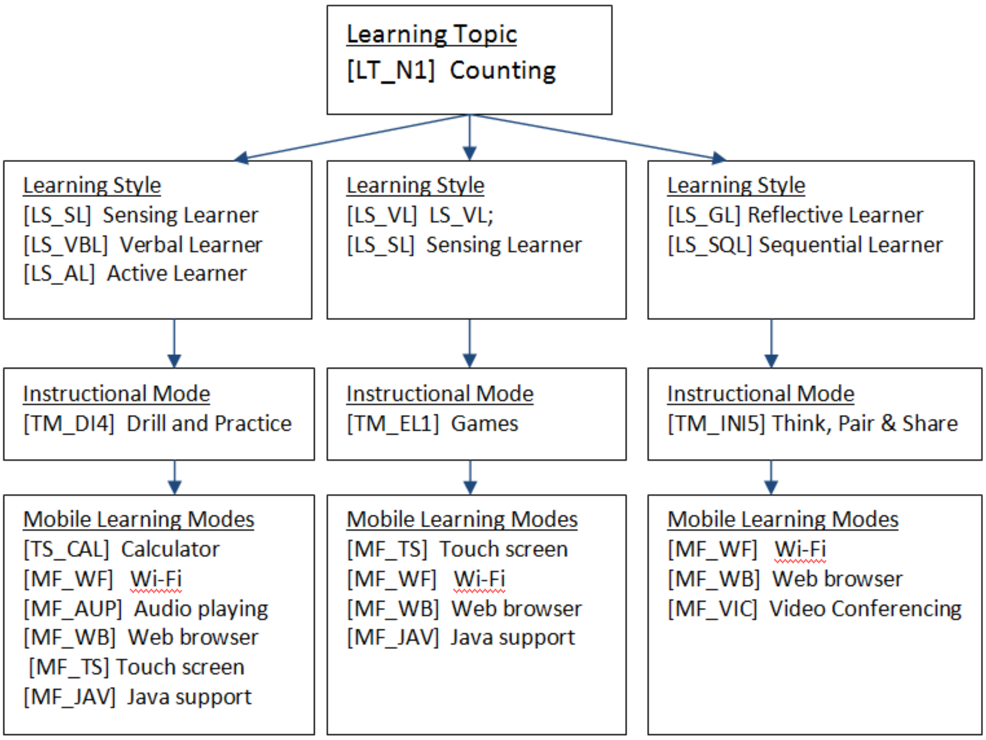


Figure 1. Mapping of Learning topic (Counting), Learning Styles, Instructional Mode and Mobile Learning Mode

3. Discussions

In Figure 2 below, a tally of mobile learning mode across all instructional modes was total and charted. There is clear indication based upon the list of instructional modes; there are some mobile learning modes that are more frequently considered useful. The chart has Bluetooth at a zero count, it indicates that it not considered useful for any delivery of instructional mode or learning domains.

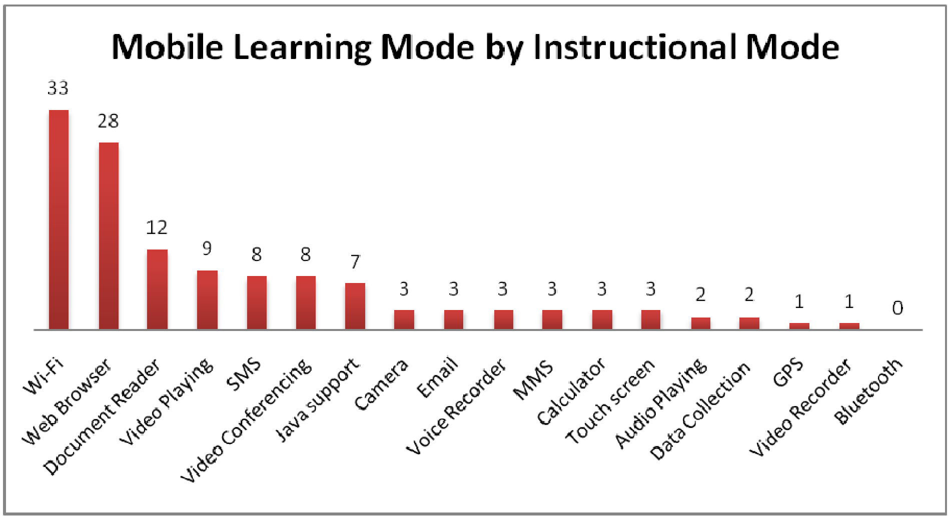


Figure 2. Mobile Learning Mode by Instructional Mode

Acknowledgements

The authors wish to thank the support of NSERC, iCORE, Xerox and the research related gift funding provided to the Learning Communities Project by Mr. Allan Markin.

References

Mahruf, M. , Shohel, C., & Power, T. (2010). Introducing mobile technology for enhancing teaching and learning in Bangladesh: teacher perspectives. *Open Learning: The Journal of Open, Distance and e-Learning, 25*(3), 201-215.

Morgan, H. (2010). Using handheld wireless technologies in school: Advantageous or disadvantageous? *Childhood Education, 87*(2), 139-142.

Ontario Ministry of Education. (2003). *Early Math Strategy: The report of the Expert Panel on Early Math in Ontario*. Ministry of Education, Ontario, Canada.

Ontario Ministry of Education. (2008a). *A Guide to Effective Instruction in Mathematics, Kindergarten to Grade 6 - Patterning and Algebra*. Ministry of Education, Ontario, Canada.

Ontario Ministry of Education. (2008b). *A Guide to Effective Instruction in Mathematics, Kindergarten to Grade 6 - Measurement*. Ministry of Education, Ontario, Canada.

Ontario Ministry of Education. (2008c). *A Guide to Effective Instruction in Mathematics, Kindergarten to Grade 6 - Geometry and Spatial Sense*. Ministry of Education, Ontario, Canada.

Ontario Ministry of Education. (2008d). *A Guide to Effective Instruction in Mathematics, Kindergarten to Grade 6 - Data Management and Probability*. Ministry of Education, Ontario, Canada.

Park, Y. (2011). A Pedagogical Framework for Mobile Learning : Categorizing Educational Applications of Mobile Technologies into Four Types. *International Review of Research in Open and Distance Learning, 12*(2), 78-102. Elsevier. Retrieved June 4, 2012, from http://www.irrodl.org/index.php/irrodl/article/view/791/1788

Rains, J., & Clark, L. (2011). A Brief Overview on Using Technology to Engage Students in Mathematics. *Current Issues in Education, 14*(2). Retrieved June 4, 2012, from http://cie.asu.edu/ojs/index.php/cieatasu/article/view/786/215

The Exploration of Online Engagement Data in LMS as Predictors to E-Learning Outcomes

Ching-Rong LEE
National Open University, Taiwan
*crlee@mail.nou.edu.tw

Abstract: In this study, it is proposed an approach to utilize the students' online engagement data, in terms of the "counts", collected by the LMS. Data about 364 students who learned online throughout a semester was analyzed. Due to the skewed and peaked distribution, the negative binomial regression was applied to the data analysis. The test scores and time spent in e-learning produce the significant effects on the log of the counts of the LMS login, the counts of course studying, as well as the counts of the e-pages read. It was shown that using the count outcome variables can form the relationships with the predictors in a linear model.

Keywords: E-learning, Online engagement, Learning Management System

Introduction

American National Survey of Student Engagement (ANSS) addresses the concern of the amount of time and effort that students devoted to their studies and other educationally purposeful activities. The key concept of "Engaged Learning" can also be applied to e-learning (Thurmond, Wambach, and Connors, 2002). The disengaged e-learners is a challenge for the instructors who try to make extensive use of various pedagogies (Hiltz, 2004). If the e-courses had numerous students, the measures of the student online engaged behaviors became time consuming. Prior studies in e-learning adopt the survey-oriented approaches (Robinson & Hullinger, 2010), such as exploring students' attitudes about e-learning, overall students' satisfaction, and the participative effects on learning outcomes. It is argued that the most reliable evaluation of educational program's effectiveness is derived from performance-based measures as a data-driven approach (Kirkpatrick, 1994). This argument is well supported by the fact that learners' sense of engagement with courses is more dependent on their connection with learning materials than with the instructor or peers (Conrad, 2002). In this study, the way of students' online learning behaviors counted and tracked by the LMS (Learning Management System) is applied to be a data-driven method for e-learning educators.

Purpose and Research Questions

LMS is capable of meeting three pedagogical features: (a) a repository of course materials, (2) communication facilities, and (3) a platform for communication over the Internet. One advantage is often ignored by the instructors is the functions of data collection, in terms of counting, tracking, and recording of students' online behaviors. Most of the LMS counts the data, such as the occurrence of the login, the occurrence of the studying course materials, the occurrence of the pages read, the time spent in e-learning, and so on. When the dependent variable is a count variable, the Poisson or negative binomial distributions are commonly used to represent its distribution. It is in the hope that significant predictors could be identified by using this new analytic methods. Three research questions were formed: (1) which engagement data can be useful for prediction; (2) what the differences existed between the predictors' impacts, (3) which a regression model has a good fit of the data.

Research Methods: A Case Study and Regression Model

A course entitled "Introduction to Information Management" was offered by Taiwan National Open University Taiwan (TNOU). In spite of attending face-to-face group tutoring of total 8 hours on four weekends, e-learners studied via an e-learning platform through 18 weeks. Both of the midterm and the final exams in the format of the paper-pencil test were administered in proctor-based classrooms nationwide.

The Descriptive Data about the E-learners' Online Behaviors

There were 443 distance students nationwide enrolled this course. 54% were males. In the end of the semester, 64 students dropped, accounting for 14% of the drop rate. 15 students' records were excluded as outliers. Table 1 shows a great variability of the students' data. On the average, a student accessed this course for study 17 times, 83 pages, and 13 hours throughout 18 weeks. The counts of LMS login accumulated the prior and the login of other e-courses. It is noted that the positive skew and peaked kurtosis failed to represent a normal distribution (see Table 1 & Figure 1).

Table 1: The descriptive data of the students' backgrounds and online behaviors

N=364	Age	Level of Years	Midterm Scores	Final Scores	Count of LMS Login	Count of Freq. in Studying	Count of Pages Read	Time(hr) Spent E-learning
Minimum	18.0	1.0	8.0	2.0	0	0	0	0
Maximum	72.0	19.0	100.0	100.0	1210.0	147.0	634.0	186.9
Mean	39.9	3.9	58.8	72.1	134.7	17.0	83.8	13.3
Std Dev.	9.9	3.9	16.2	19.6	176.2	22.0	113.2	22.5
Skewness	0.3	2.3	-0.2	-0.6	2.4	2.6	1.9	2.7
Kurtosis	-0.1	5.0	-0.1	-0.4	7.0	8.4	4.2	10.9

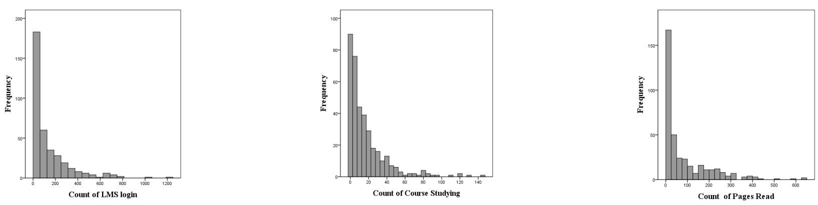


Figure 1: Frequency histogram of engagement data (LMS login, course study, pages read)

Negative Binomial Model: An Over-Dispersion Poisson Regression

Negative binomial regression is known as a log-linear model, with the dependent variable, a count variable, has a Poisson distribution. The logarithm of its expected value can be modeled by a linear combination of the predictors. Over-dispersion is occurred when the problems of excess zeros in which a subgroup of respondents who would never display the behavior are included in the sample. In this study, the regression analyses were performed for the three models, in which each included a different dependent variable as follows: Y1 is the counts of LMS login, Y2 is the counts of course studying, and Y3 is the counts of pages read, respectively. Age, gender, level of college years, time spent in e-learning, the midterm test scores, and the final exam scores served as the predictors.

Results and Discussions

In Table 2, the parameter estimates shows the negative binomial regression coefficients for each of the predictor variables along with their standard errors. Across the three models, the predictor variables, the midterm scores, the final scores, and time spent in e-learning, are statistically significant. It explained that, for instance, each one-unit increase on midterm scores, the expected log count of the Y1, Y2, Y3 increase by 0.01 times, 0.014 times, and 0.011 times respectively. The midterm scores have a stronger affect (B =0.014) on the counts of course studying (i.e., Y2) than Y1 and Y3. Each one-unit increase on time spent in e-learning, the expected logarithm of the pages read (i.e., Y3) increase by 0.044 pages. The categorical variable Level=1st year has a coefficient of -1.195, which has statistically significant effect only on Y1. As compared to senior students, the count of LMS login for the freshman was less. The time spent in e-learning for the male students is decreased by 0.274 hours, less than the female. The LMS login can be well predicted by the factors of age, midterm, final exam, time spent, and the levels of the college years. The college years may affect only on the frequency of using the LMS. Test scores and time spent in e-learning can be well used for predicting the students' online engagement.

Table 2: Parameter estimates of negative binomial regression by three dependent variables

Parameters N=364	df	B of Y1	Std Error	Wald Chi-Square	B of Y2	Std Error	Wald Chi-Square	B of Y3	Std Error	Wald Chi-Square
(Intercept)	1	2.845	0.321	78.42***	0.296	0.352	0.71	1.330	0.334	15.81***
Gender=male	1	0.007	0.108	0.00	-0.001	0.112	0.00	-0.274	0.109	6.38**
Gender=female		0			0			0		
Level=1st year	1	-1.195	0.150	63.20***	-0.068	0.160	0.18	0.052	0.154	0.11
Level=2nd~3rd yr	1	-0.161	0.122	1.74	-0.089	0.131	0.46	-0.047	0.124	0.14
Level=4th year		0			0			0		
Age	1	0.018	0.006	9.56**	0.010	0.006	2.70	0.023	0.006	14.31***
Midterm scores	1	0.010	0.004	7.10**	0.014	0.004	13.16***	0.011	0.004	7.47**
Final scores	1	0.007	0.003	4.92*	0.009	0.004	6.80**	0.007	0.003	4.92*
Time spent	1	0.022	0.003	53.25***	0.027	0.003	75.52***	0.044	0.004	152.90***

Note: *** stands for ρ -value less than the significance level at 0.001; ** : 0.01; *: 0.05.

Conclusion

This study has analyzed the data about e-learners' engaged behaviors in use of a negative binomial regression. The data collected by LMS in the case study shows that the effects of the predictors, such as age, test scores, and time spent on the counts variables. The findings provide a good fit of the new method and the count data types.

References

[1] Conrad, D. (2002). Engagement, Excitement, Anxiety, and Fear: Learners' Experiences of Starting an Online Course. American Journal of Distance Education, 16(4), 205-226.

[2] Hiltz S. R. (1994). Learning Together Online. Lawrence Erlbaum Association. New Jersey.

[3] Kirkpatrick, D. (1994). Evaluating Training programs. Berrett Koehler, San Francisco, CA.

[4] Robinson, C. & Hullinger, H. (2010). New Benchmarks in Higher Education: Student Engagement in online learning. Journal of Education for Business, 12,3, 100-120.

[5] Thurmond, V., Wambach, K., & Connors H. (2002). Evaluation of student satisfaction: Determining the impact of a Web-based environment by controlling for student characteristics. The American Journal of Distance Education, 60, 169–189.

e-Learning for Business English

Masnidar TANJUNG*
Muhammadiyah University of Prof. DR. Hamka, Indonesia
*nitanvirgo54@yahoo.com

Abstract. This paper discusses teaching Business English using *E*-Learning media technology; the technology of which the interconnection and networking (Internet) as the global network of computers is used, that enables people to connect with each other. This is a kind of web enhanced course using the internet to support the quality of learning process done in the classroom. The teacher in this case functions as a facilitator in searching the sites related to the teaching materials, and create learning experience that is interesting and joyful set of learning activities; and which is not of the less importance is having the capability to access to the internet by the teacher him/herself. Why should this kind of media be used is that by the reason where we can see the value of this media as a language teaching and learning resource that can motivate and provide a source of authentic and varied language teaching and learning activities and materials on Business English related to the skills of Listening, Speaking, and Writing which are all supported by the Vocabulary as the language component, as well. This medium is excellent, where the learners do exercises on line, in terms of some typical English expressions used in the business world that can be in forms of Putting Words in order, Paraphrasing Meaning, Mix and Match, Completing the Sentence with Choice, Rearranging the lines, True False, Listening by Clicking on the Name the Words you Hear, Listening to the Text by Clicking on, and the form of Y/N for Is It Exactly What You Hear.

Keywords: *E*-learning, Business English, Facilitator, Joyful Learning

Introduction

It has been stated formally that learning foreign languages, including English, has become important subjects among the other subjects taught at schools in Indonesia nowadays. It is the fact that some teachers who teach English, including Business English often encounter difficulties in making their students successful in learning. Failure in teaching might be caused by the weakness of aspects in learning as well as in teaching; such as lack of books, poor equipment and teaching aids, poor student motivation and inadequate teacher's qualifications as well.

As a matter of fact, and based on the author's experience in teaching English, she still has not been able to get satisfactory results yet. She found that the students were not highly motivated and rather lazy during the period of learning, and this influenced their achievements, that may be resulted by the way that she presented the lesson. She realized then that the way of teaching she used has some weaknesses from the teacher's point of view, as well as from the students. Due to this problem, she tried to look for the way out of solving the problem during the process of teaching and learning activities. She tried to change her technique in teaching that might create good result in teaching; by applying E-Learning technology in teaching Business English, and in turns, the teaching can be more enjoyable and beneficial.

1. An Overview of E-Learning Technology

Why was this kind of technology chosen to be used in teaching Business English, was inspired by:

1) E-learning concept accessed from Wikipedia [1], the free encyclopedia stated that there is a trend to move towards blended learning services, where computer-based activities are integrated with practical or classroom-based situations. E-learning is naturally suited to distance learning and flexible learning, but can also be used in conjunction with face-to-face teaching, in which case the term Blended learning is commonly used. E-Learning pioneer Bernard Luskin argues that the "E" must be understood to have broad meaning if e-Learning is to be effective. Luskin says that the "e" should be interpreted to mean exciting, energetic, enthusiastic, emotional, extended, excellent, and educational in addition to "electronic" From the learner's point of view this provides the ability to pause and rewind and gives the learner the advantage of moving at their own pace, something a classroom cannot always offer.
2) Learning concept proposed by Pearson Brown who has been teaching people from all sectors of business to communicate more effectively in English for over twenty-five years. He has written on line exercises for his students to develop their language skills, which are all copyright (c) 2006 /2007 Caroline and Pearson Brown. He can send out free English lessons by email from time to time. Resources for learning, include Business Vocabulary and Presentations, Business Letters and the Sample of Business Letters, Negotiations in English, Import-Export why he chose to create on line learning was based on his experience of learning Chinese by himself instead of in a classroom, and found it fun.
3) The Program of *Nabiku Idolaku* is a kind of book presented in two languages: English and Indonesian; with the aim of which to enrich and enhance the language mastery of the children. In addition, the use of Smart E-Pen also aims at giving the learners pronounce the language in English and Indonesian as well. This kind of program is called Pelangi Mizan Smart E-Pen [2], and contains text by which can be read by the Smart E-Pen the pen of which can also play musical illustration, and dialogue as well. This program is also regarded as having harmony between visual and audio that can increase the children's imagination.

2. The Concept of E-Learning Proposed

What is meant by the E-Learning technology in teaching Business English is that, this is a kind of **Business English** especially related to international communications. It is a part of English for Specific Purposes and can be considered a specialism within English language learning and teaching. Business English means different things to different people, and for some, it focuses on vocabulary and topics used in the worlds of business, trade, finance, and international relations. For others it refers to the communication skills used in the workplace, and focuses on the language and skills needed for typical business communication such as presentations, negotiations, meetings, small talk, socializing, correspondence, report writing, and so on. What is meant by E-learning in business concerns about where the students have the opportunity in doing exercises on line and in up-to-date program of exercises; most of which are taken from Pearson online exercise related to the materials mentioned above.

3. Learning Materials

Business vocabulary, idioms and expressions:

- Idioms, Language of Recruitment, Hire and Fire
- Jobs & related verbs and Adjectives, and Intensifiers
- Words related to management and Common phrasal verbs
- Parallel Structure

Listen and Learn How to Talk in the Office.

- Starting a Presentation and Rhetorical skills
- Greeting and Arranging an Appointment
- Talking on the Phone and Negotiations

Writing Skill

- Applying for a job, and Preparing for a Job Interview
- Useful Business Language expressions used at work
- A Design Procedure for Routine Business Documents
- Effective Writing and Tone in Business Writing
- Business Letters, Technical Reports & Report Abstracts

4. Research Finding and Recommendation

In implementing teaching Business English by applying E-Learning Technology, the writer actually wants to look for the benefit of teaching using this technique of teaching; and what she has done is arranging the pre-test (before the treatment was applied) and the post-test (after the treatment was applied); and then examined the result by applying **t-test** formula.

The writer then got an exciting conclusion that the students taught Business English by using E-Learning technology can get higher scores in learning that can be proved by the significant difference of the **t-test** with the result of **3.75** (the table of t-ratio says that with df 60 is 1.67 which is higher than the t-table), that is at **0.05** level of significance. In that it is recommended that for the teachers are suggested to be creative in looking for better learning strategies that can fulfill the needs of the students, and make the learning enjoyable as well, which is not of the less important.

References

[1] Wikipedia, the free encyclopedia 2012: Business English
[2] *www.bukuanakmuslim.com/2011/**nabiku-idolaku**/*Pelangi *Mizan Publishing*

The Correlation between the Students' Attitude towards ICT with the Students' Achievement in Learning English

Endy SJAIFUL ALIM *
Muhammadiyah University Prof. Dr Hanka, Indonesia
*esaitb@yahoo.com

Abstract. This paper is a research report concerning Computer-Assisted Language Learning (CALL), used as medium of teaching English; in terms of looking for significant correlation between the Students' attitude towards ICT with the students' learning achievements in English. This is a kind of experimental research; where the data processing was performed by using descriptive statistics that aimed to know; the mean, the median, the range, the standard deviation, the variance, the minimum and the maximum score, and the presentation of the data forms to show the frequency of distribution and in a histogram. Based on the calculation of the correlation score, the research variables have positive correlation and it can be said that there is quite strong and significant correlation with the significance level of 5% (Sig. < 0.05). It means that all the three hypotheses stated or presented in this study could be accepted.

Keywords: *ICT*, Attitude, Achievement, English Learning,

1. Introduction

As more and more Information and Communication Technology (ICT) resources become available both to support English education and for Web-based learning, it becomes increasingly interesting to map the ICT - based resources available for students learning English and their attitudes towards the utilization of the resources. Increasing availability of computers lead to new ways of delivering the English language learning using interactive multimedia learning modules and possibly by integrating ICTs into other methods of English language teaching and learning. The use of ICT has been increasing in the school or university over the last years, but the possibilities of new technology, particularly in Indonesia, are still not being used in favor of pedagogical changes.

Growing schools, however, have tried to implement new ways of teaching and learning using, or facilitated by, information and communication technology (ICT), and in many cases integration of ICT has been used to support changes in teaching and learning methods. Access to computers, at home and on campus or schools and even internet access are both important. The issue of students' attitude toward ICT, students' familiarity with computer, and students' learning achievement in English delivered via Computer known as Computer – Assisted Language Learning (CALL) are important. This is because even provided with a very sophisticated learning media, students are not automatically gain promising learning achievement; their attitude towards the media – computers and their familiarity need to be taken into account as these factors may possibly be potential variables that affect their learning achievement [1].

In addition, based on previous studies conducted by language researchers such as Fitzelle and Trochim [2], students should have positive attitudes in order to be successful in their language learning. In term of using CALL, familiarity toward computer, which is used as medium of delivering the English language learning, is also needed to be put into consideration. Familiarity is something that needs to be possessed by students because for some students in particular situations and conditions using computers has not been a part of their daily life. In other words, they are not quite familiar with computers as they do not use computer very often. Hence, to investigate whether the two variables correlate to the students' English learning achievement an empirical research needs to be conducted.

2. The Objectives of the Research

The research is aimed at investigating:

1. Whether there is a positive relationship between Students' attitude towards ICT and students' learning achievements in English.
2. Whether there is a positive relationship between students' familiarity with computer and students' learning achievements in English.
3. Whether there is a positive relationship between students' attitude towards ICT and students' familiarity with computer towards students' learning achievements in English.

3. The Description of the Data

From the research data on students' attitude toward ICT (X), the students' familiarity with computer (Y), and students' learning achievement in English delivered via CALL (DYNED) as shown in the appendix 1, the data processing are performed using descriptive statistics. Descriptive statistics is used to know; mean, median, range, standard deviation, variance, minimum and maximum score, and the presentation of data forms is given to show the frequency of distribution and in histogram.

This result provided that the descriptive statistics of variables of the students' attitude toward ICT (X_1), students' familiarity with computer (X_2), and students' learning achievement in English delivered via CALL (DYNED) have positive correlation. The students' attitude toward ICT variable was explained by number of respondents (N) = 35 who filled attitude questionnaires with the score; mean = 66.2571 and standard deviation = 5.77185, respondents (N) = 35 who filled familiarity questionnaires with score; mean = 71.2286 and standard deviation 5.07639, respondent (N) = 35 who answered the tests with score; mean = 82.2851 and standard deviation = 2.67602.

4. Conclusion

Based on the research hypothesis and the data analysis, the research about the relationship between students' attitude toward ICT and students' familiarity with computer with students' learning achievement in English conducted at *SMA Negeri 3 Setiabudi* can be summarized as follows:

1. The relations $X_1 – Y$ is positive and significant, with $r_{y1} = 0.838$ with alpha level of 5% which indicates that if students' attitude toward ICT increase, the result learning achievement in English will increase. And vice versa if the students' attitude toward ICT is low, the result will be low. Because it is significantly less than 0.05 , the relationship between students' attitude toward ICT with students' learning achievement in English can be generalized to all population. Partial correlation

between X_1 toward Y, where X_2 as a control is obtained r_{y12} = 0.838, so it is significant and can be applied to all population. Meanwhile, the contribution R Square is 0.702 , it means that about 70.2% of students' learning achievement in English result can be explained by the variables of students' attitude toward ICT and students' familiarity with computer while the rest 70.2% is determined by other variables.

2. The relationship between students' familiarity with computer (X_2) with Y is positive and significant, with r_{y12} = 0.838, at alpha level of 5% which showed a correlation. This shows that if familiarity with computer increases the learning achievement in English will increase. Similarly, if the familiarity with computer decreases, the ability of students' learning achievement in English will decrease. Because it is significantly less than 0.05, it means that the relationship between students' familiarity with computer with students' learning achievement in English can be generalized to all population. Meanwhile, the contribution R. Square is 0.702, it means that about 70.2% of learning achievement in English result can be explained by familiarity with computer.
3. The relationship between X_1 and X_2 with Y is positive and significant, with r_{y12} = 0.838 at alpha level of 5%, it showed high correlation. This shows that if the students' familiarity with computer increases the students' learning achievement in English will increase. Similarly, if the students' familiarity with computer decreases, the students' learning achievement in English will decrease. Because it is significantly less than 0.05, it means that the relationship between students' familiarity with computer with students' learning achievement in English can be generalized to all population. Partial correlation of X_2 with Y, where X_1 as the control is obtained r_{y12} = 0.838, it means significant and can be applied to all population. Meanwhile, the contribution R. Square is 0.702, it means that about 44% of students' learning achievement result can be explained by the familiarity with computer.

References

[1] Best, J. W. (1981). *Research in Education*. New Jersey: Prentice – Hall Inc.

[2] Fitzelle, G. T., & Trochim, W. M. K. (1996). *Survey Evaluation of website instructional technology: Does it increase student learning?* Retrieved from http://trochim.human.cornelledu/webeval/webques/webques.htm

[3] Hartoyo (2008). *Individual Difference in Computer-Assisted Language Learning (CALL).* Semarang: Pelita Insani Semarang.

[4] Saleh, M. (2001). *Pengantar Praktek Penelitian Pengajaran Bahasa*. Semarang: *IKIP* Semarang Press.

Intimacy to Others: Scale Based Internet

Resty Wales PRATANDANI[a]*, Subhan El HAFIZ[a]
[a]*Faculty of Psychology, Universitas Muhammadiyah Prof. Dr. HAMKA, Indonesia*
*resty.wales@yahoo.co.id

Abstract: Human interactions often lead to feelings of deep interest called intimacy [1]. The ability of this intimacy which was researched by creating scale of intimacy for the community so that they can evaluate themselves using self-learning media. Scale tested using a survey administered to 35 participants. Data were analyzed using Alpha Cronbach Reliability test. The result, the remaining 25 of 36 items had good reliability with an alpha score of 0.873. Scales that have been tested disseminated through the internet social media technology that can be connected to many people. Alpha Cronbach score indicate that the scale is sufficiently reliable to measure intimacy.

Keywords: Intimacy Scale, Internet social media.

Background

Nowadays, the social technology of the internet becomes the needs of human life. Almost every activity can be easily done using the internet. Interaction is one of human activity that often done via internet.

Through the internet that can connect people, it enabling for every individuals to be acquainted so not a few people who make up a connection via social media account that they have made. The connection can begin from an interaction between individuals. Then within certain time cause the feeling of affinity.

After emerging sense of interest, each individual will undergo an intense relationship. In psychology emerging sense of interest is phase of creating interpersonal relation by communication and interaction while intense relationship is the phase of strengthening the relationship through direct acknowledgement, positive feeling, etc. [2]. Through such an intense relationship that will lead to deeper feelings between individuals. The deep sense is called intimacy [1].

For example, individuals found similar of interest and pleasure to others. Then he or she will communicate each other in a certain time, in which the process can lead to feelings of affinity. They were chatting, writing comment on other wall, send messages, and more. These activities purpose is to make their relationship closer and deeper.

The other phenomenon that is often found in interaction via internet is that relationship building, war in social media account, breaking up of couples in love, as a mass gathering to support a particular motion, and more [3, 4]. This shows that the using of the internet for user account social media (Facebook) has a large impact to the development of human life as social beings. In consequence, the researchers interest to measure intimacy scale of individuals via internet. The internet user is very important in this study.

1. Preliminary study

1.1. Method

The research used survey method. Thirty five participants were included in the study.

1.2. Result

The study used quantitative method with the number of respondent is 35 students. After analyze data using Alpha Cronbach reliability, the result shows 25 of 36 item had good reliability. This study shows that need any program measurement of intimacy scale via computer technology, internet and the other electronic media.

2. Computer Model

Based on the earlier study, we made computer model that can be access and distributed via internet. The computer model is made to spread the intimacy scale. The deployment scale of intimacy can only access by the people who have a Facebook account. Furthermore, the scale of intimacy will be re-testing to more variety of ages and cultures to be more reliable and useful scale.

References

[1] Newman, B. M., & Newman, P. R. (1979). An Introduction to the 7th Psychology of Adolescence. Illionis : The Dorsey Press.

[2] Rakhmat, J. (2005). *Psikologi Komunikasi*. Bandung : PT Remaja Rosdakarya.

[3] Susanto, A. D. (2012). Didepak dari Daftar Pertemanan Facebook Menyakitkan. Retrieved from: http://www.merdeka.com/teknologi/didepak-dari-daftar-pertemanan-facebook-menyakitkan.html.

[4] Health Day News. (2012). Can Facebook prolong post-breakup pain? Retrieved from: http://www.wowktv.com/story/19653549/can-facebook-prolong-post-breakup-pain-healthday-news.

Hardiness Test: Find Your Score about Hardiness

Siti KHADIJAH[a*], Fahrul ROZI[a]
[a]*Faculty of Psychology, Universitas Muhammadiyah Prof. Dr. HAMKA, Indonesia*
*khadijah.siti86@gmail.com

Abstract: Many psychologists believe knowledge will guide one's life. Therefore knowledge about hardiness might be correlate to their own hardiness life. Hardiness [1] is a constellation of personality characteristics that cause an individual more powerful, durable, stable, and optimistic in the face of stress and reduce the negative effects encountered. Thus, the objective of this study is creating hardiness test so one's can learn more about hardiness. The test was conducted with 35 participants. The result showed that most of participants had low understanding (48.5%) and average understanding (51.4%) about hardiness. This result also indicates that people need programs to increase understanding about hardiness using media such as internet.

Keywords: hardiness, adolescent, Facebook

Background

Adolescence is the most complex period because teenagers undergo significant changes physically and psychologically. Physical changes characterized by emergence of secondary sex characteristics, namely height growth and genetalia organ development in the male and female. It also causes psychological changes, the emotions tend to be unstable [2]. Along with the times and technology, teens change in different ways as well, and it is not rare that this causes stress for the teenagers themselves. Many found that teenage who are stressed will resolve the problem to negative things so that need to develop a hardiness personality.

Hardiness personality is associated with people who have made commitments in life, they have control, and see difficult or new situations as a challenge, not a threat. Teenage with hardiness personality will always be optimistic, be able to cope with stress, and be able to adapt to events that cause stress.

Nowadays, the technology is more advance and internet is increasing widespread as well. The applications on the internet also varied, one of them is social network like Facebook. Most teens in Indonesia have an account on Facebook. Therefore, as researchers, we have developed an idea that combines hardiness measurement with Facebook use, so that the teenagers easily access and respond to questionnaires on hardiness. The use of Facebook is quite significant in this study so providing information about hardiness can give great benefit to the teenager users of Facebook.

1. Preliminary Research

1.1. Method

This study used a survey method. The participants included 35 students in Universitas Muhammadiyah Prof. DR. Hamka.

1.2. Result

This study is used quantitative methods with the number of respondents were 35 students. The score of the hardiness test which have been done by the respondents was 48.5% with the low score, 51.5% with the medium score, and 0% with the high score. Calculation of scores is divided into three, namely low, medium and high. The score is derived from the total of questions difficulties divided by three grades of scores, namely low, medium and high. Based on the research, the majority have a hardiness personality in a moderate level. The weakness of this test is that it includes a number of hardiness distracters that less appropriate and needs repair. This study shows that there needs to be a test of hardiness personality through Facebook so the teenagers can figure out what the personality that must be improved.

2. Computer Model

This test is distributed through Facebook. This test will sent through message in Facebook to the user teenagers of Facebook. They will do the tests to give a blue on the answer that they think is correct. Then they were asked to send back the test that has been done and the result will be provided through the message as well. In the future, this hardines tests will be developed to be used by workers to deal with job demands.

References

[1] Kobasa, S. C., Maddi, S. R., & Kahn, S. (1982). Hardines and Health: A prospective study. *Journal of Personality and Social Psychology*. *42*, 162-177.

[2] Mönks F. J., Knoers, A. M. P., & Haditono, S. R. (2002). *Psikologi Perkembangan : Pengantar dalam Berbagai Bagiannya (cetakan 14).* Yogyakarta : Gadjah Mada University Press.