So, H. J. et al. (Eds.) (2020). Proceedings of the 28th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Middle-School Students' Behavior Pattern and Strategy Selection in Problem Solving: A Study Based on Data from PISA 2012

Yang LIU[a], Bo YANG[a], Lan WU[b], Baoping LI[ab*] & Shengquan YU[ab]
[a]*Advanced Innovation Center for Future Education, Beijing Normal University, China*
[b]*Faculty of Education, Beijing Normal University, China*
*libp@bnu.edu.cn

Abstract: Vary-one-thing-at-a-time (VOTAT) strategy is regarded as the optimal strategy in the knowledge acquisition stage of complex problem solving (CPS) in many studies. Based on the log-file data of the Climate Control task of computer-based assessment of the CPS in the 2012 cycle of the Programme for International Student Assessment (PISA), 388,931 pieces of records from 20,597 students who used VOTAT were collected for an in-depth analysis. The result of the latent class analysis identified three kinds of strategies, the depth-first strategy, breadth-first strategy and mixed-strategy. The Pearson's chi-squared Test and Welch variance analysis showed that the task performance of the students in the three groups varied. These results suggest that VOTAT strategy could be distinguished in a more elaborated way, which would contribute to portray the CPS process in detail, and so it is proved to play a key role in the CPS performance again.

Keywords: complex problem solving (CPS), learning analytics, VOTAT, depth-first strategy, breadth-first strategy, mixed-strategy

1. Introduction

As countries around the world are attaching great importance to the development of students' domain-general skills. The competence of CPS, which is regarded as the key of the future education (OECD, 2013), has attracted more and more attention of educators.

During the process of complex problem solving, the application of the VOTAT strategy usually makes it easier to find the answers. Although a lot of researchers found that the use of the VOTAT strategy was significantly correlated with higher performance in problem-solving tasks, most of them focused on a simple control action of variables, the continuity between behaviors was not taken into account, which would limit the understanding of CPS.

2. Literature Review

2.1 Definition and Assessment of Complex Problem Solving

Complex problem solving (CPS), also known as dynamic problem solving, dynamic decision making, interactive problem solving, and creative problem solving in PISA 2012 (Greiff, Molnár, Martin, Zimmermann, & Csapó, 2018), can be seen as finding solutions in dynamic tasks. There is insufficient information to solve the problem at the given state and the problem solver needs to integrate the information obtained in the process of exploration to solve the problem (Buchner, 1995). In terms of structure, CPS is mainly divided into two dimensions: knowledge acquisition and knowledge application (Funke, 2001).

2.2 VOTAT Strategy in CPS

Tschirgi (1980) proposed the VOTAT strategy, which is also known as the control of variables strategy in the manipulation-of-variables task (Wüstenberg, Greiff, & Funke, 2012; Croker & Buchanan, 2011; Kuhn & Dean, 2005). At the stage of knowledge acquisition, VOTAT represents the fact that the problem solver explores the independent influence of input variables on output variables by changing a single input variable. The empirical research showed that the use of the VOTAT strategy was significantly correlated with higher performance of complex problem solving (Wüstenberg, Stadler, Hautamäki, & Greiff, 2014; Greiff, Wüstenberg, & Avvisati, 2015).

However, Molnár et al (2018) found that the use of the VOTAT strategy did not always lead to higher performance and only conscious users of the VOTAT strategy proved to be the best solvers with non-conscious users of the VOTAT strategy the second and non-VOTAT strategy users the worst. Therefore, based on the common VOTAT strategy, this paper aims to make a more detailed exploration to further understand the cognitive pattern in the process of CPS, and increases the interpretability and educational significance of the strategy.

3. Research Questions

In this work, we aim to answer the following questions:

- Are there multiple strategies linked to VOTAT in the exploration of problem solving?
- What are the differences in the performance of students using different strategies in both the specific task and overall problem-solving proficiency?

4. Research Design

4.1 Test Task and Data

This paper opted for using the data (retrieved from https://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm) obtained from the CPS test unit, Climate Control in PISA 2012 as it provided the authoritative log-file of VOTAT. There are about 30,820 students from 42 countries or economies participated. Only the process data of manipulating input variables with a total of 477,258 entities was collected. In the data, some students did not use VOTAT in the knowledge acquisition phase and their data was deleted and finally, 388,931 pieces of records from 20,597 students were available for subsequent analysis.

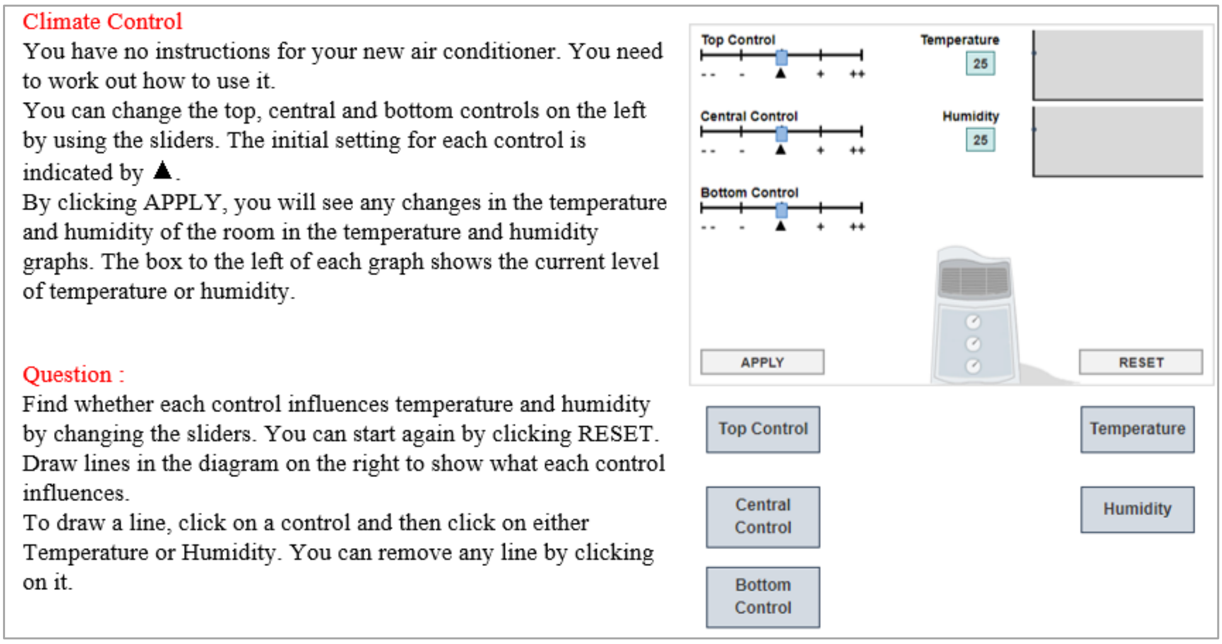


Figure 1. The Test Task of Climate Control in PISA 2012

4.2 *Variables and Scoring*

4.2.1 Coding the Behavior in the Climate Control Task

All the behaviors of participants were recorded in the log-file. In this dynamic interaction system, information would not be generated in the final drawing behavior. Therefore, only manipulating data in the process of exploration such as clicking the "reset" tab and adjust the control button will be coded and analyzed.

The subsequent behavior of VOTAT was counted and labeled by whether a certain behavior occurred and the behavior which occurred was labeled with 1 while the behavior which did not occur was labeled with 0. There are four main types of behaviors:

- The previous VOTAT behavior was repeated and no input variable values were changed. This kind of behavior sequence was labeled as "e (0_step_0_class_0_value)".
- Based on the previous input variable settings, only different values were tried changed on the same control variable. This kind of behavior sequence was labeled as "ev (1_step_0_class_1_value)".
- Based on the previous input variable settings, only one of the other variables was changed. This kind of behavior sequence was labeled as "ec (1_step_1_class_0_value).
- Based on the previous input variable settings, a new input variable is manipulated and the previous control variable was restored to the initial value. This kind of behavior sequence was labeled as "ecv+ (2_step_1_class_1_value)".

Other behaviors are not used in this study because their counts tend to zero-inflate. Table 1 shows some of the behavior sequences.

Table 1. *VOTAT and Its Follow-up Behaviors*

	[2,0,0]	[1,0,0]	[0,1,0]	[0,0,1]	[1,0,1]	[2,0,1]
[1,0,0]	ev	e	ecv+	ecv+	ec	---
[0,1,0]	ecv+	ecv+	e	ecv+	---	---
[0,0,1]	ecv+	ecv+	ecv+	e	ec	ec

Note: The column represents the VOTAT behaviors and the row represents the exploration behaviors after the VOTAT behaviors. The three values in brackets show the input variable values of the top control, central control and bottom control in the Climate Control task, respectively.

4.2.2 Students' Performance in Problem-Solving Tasks

In the Climate Control task, scores would be awarded if the correct relationship between the input and output variables was given. In addition to this task which involves interactive behaviors of the students in the process of problem solving, some other tasks such as logical reasoning and information retrieval were also employed in PISA 2012 to assess students' problem-solving ability and an overall score was given. For the overall problem-solving proficiency, PISA used an imputation methodology to derive 5 plausible values, the plausible values were random draws from the marginal posterior of the latent distribution for each student (OECD, 2014). This paper used the first plausible value as the overall score of the students (cf., Greiff et al., 2015).

5. Results

5.1 *Exploratory Latent Class Analysis of Various Follow-up Behaviors of VOTAT*

The latent class analysis was conducted using Mplus. In this paper, three latent class models were obtained. Table 2 indicates the fit indices.

Table 2. *Fit Indices for Latent Class Analyses*

latent classes	G^2	χ^2	AIC	BIC	aBIC	entropy	class proportions
1	2342.892	2534.746	103451.954	103483.815	103471.103	---	---
2	776.554	799.610	101895.616	101967.302	101938.701	0.910	0.083/0.917
3	311.767	316.997	101440.829	101552.340	101507.849	0.946	0.604/0.296/0.1

As Table 2 indicates, the 3-class model fits the data best since all of its fit indices appear smallest among the three models with the entropy of 0.946 in Table 2 implying that the 3-class model is reliable. In addition, each class is named according to the latent class probability and conditional probability of four main behaviors after VOTAT. As Figure 2 illustrates, the latent class 1 shows high probability in ev, which indicates that the students using this class preferred to explore different values of the same control variable after VOTAT. The latent class 2 shows the same probability as class 1 in e, ev and ec but a higher probability in ecv+, which indicates that the students using this class explored different values of the same control after VOTAT, as well as a single control of the others. In the latent class 3, the high probability of ec means different control variables were explored.

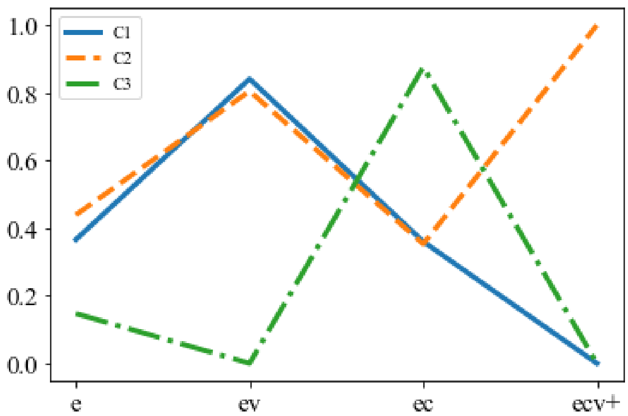
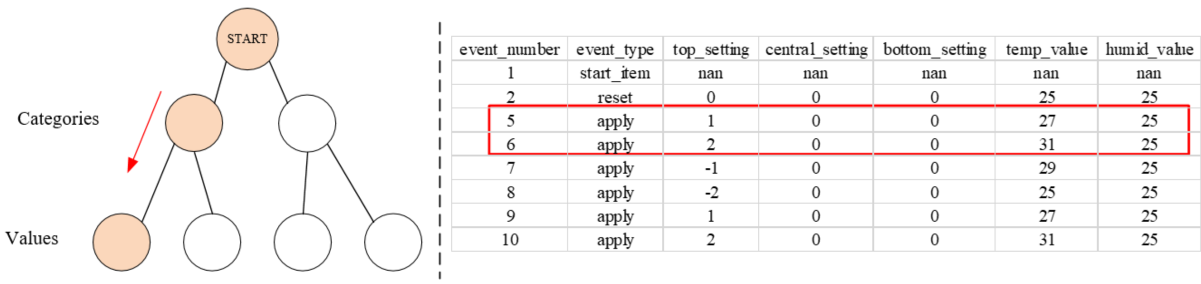


Figure 2. Conditional Probability Distribution of Three Latent Classes

The depth-first search and breadth-first search are traversal algorithms to search graph structure and tree structure in computer science (Kreher & Stinson, 1999; Jungnickel, 2008). Depth-first search starts from the initial node, and extends the search to the next level of child node sequentially, which is similar to the behavior pattern of the latent class 1 in the previous analysis. Breadth-first search starts from the initial node and first cover the neighboring nodes, which is similar to the behavior pattern of the latent class 3. Therefore, the latent class 1, 2 and 3 are named depth-first strategy, mixed-strategy and breadth-first strategy, respectively. For example, the schematic diagram of a problem solver who used depth-first strategy as shown in Figure 3, and the code ev is marked with a red border.



event_number	event_type	top_setting	central_setting	bottom_setting	temp_value	humid_value
1	start_item	nan	nan	nan	nan	nan
2	reset	0	0	0	25	25
5	apply	1	0	0	27	25
6	apply	2	0	0	31	25
7	apply	-1	0	0	29	25
8	apply	-2	0	0	25	25
9	apply	1	0	0	27	25
10	apply	2	0	0	31	25

Figure 3. The Schematic Diagram and Exploration of Depth-First Strategy

5.2 Performance Differences of Using the Three Strategies in the Climate Control Task

Pearson's chi-squared test was carried out to investigate the difference of three strategies in the performance of the Climate Control task, as Table 3 shows. The result indicated that there is a significant difference among the students who used different strategies (χ^2 = 2534.374, p = 0.000***). The correlation coefficient of 0.351 indicated a moderate association, which means the identification of three strategies is reliable (Rea & Parker, 2014).

Table 3. *Pearson's Chi-Squared Test for the Three Different Strategies in the Climate Control Task*

	Depth-First	Breadth-First	Mixed-Strategy	Sum
Correct	9170	432	4684	14286
Incorrect	3253	1627	1431	6311
Sum	12423	2059	6115	20597
$\chi^2 = 2534.374$		P = 0.000***		V = 0.351

It was found that students using the mixed-strategy performed the best with 77.112% of them correctly finishing the task. In comparison, students with breadth-first strategy did the worst with only 17.58% of them correctly completing the task. The performance of the students with the depth-first strategy also did a good job with 75.923% them succeeding. There is one thing worth noting that the percentage of the students correctly finishing the task was calculated based on the final sampling weight in PISA and therefore, it would not completely agree with the sample distribution in Table 3.

5.3 Differences of the Three Strategies in the Overall Problem-Solving Performance

Moreover, to explore students' overall performance in CPS, the data was first analyzed by using the normality test. The Kolmogorov-Smirnov test revealed that the data did not follow the Gauss Distribution (D = 0.016, p = 0.000***). Then, the homogeneity of variances test showed the assumption also failed (Bartlett's K-squared = 61.426, p = 0.000***). Therefore, the data could not be analyzed using the conventional ANOVA method and the Welch's test was used. The result showed a moderate effect with the effect size of 0.069 (Welch F = 848.805, p = 0.000***, $1 - \beta > 0.99$). Table 4 indicates the result of the subsequent Games-Howell's multiple comparison.

Table 4. *Games-Howell's Multiple Comparisons of the Three Strategies in the Overall Problem-Solving Performance*

Multiple Comparisons		Mean Difference	Standard Error	Sig	95% confidence interval lower-bound	upper-bound
Depth-First	Mixed-Strategy	-21.4249	1.287	0.000	-24.4416	-18.4083
Depth-First	Breadth-First	61.8619	1.9064	0.000	57.3915	66.3323
Mixed-Strategy	Breadth-First	83.2868	2.0213	0.000	78.5475	88.0262

The result complied with the preceding finding about students' performance in the Climate Control task and individuals with mixed-strategy students performed the best, the depth-first strategy students the second and the breadth-first students the worst.

6. Discussion and Conclusion

This paper has identified three different types of strategies linked to VOTAT by using the latent class analysis to explore the follow-up behaviors of VOTAT, the depth-first strategy, breadth-first strategy, mixed-strategy. Similar findings were mentioned, such as engineering design (e.g., Ball, Evans, Dennis, & Ormerod, 1997) and information seeking (e.g., Heinström, 2005), but in the research of CPS, related research was limited (at the best of our knowledge). This research proved that CPS performance was strongly related with the strategies after applying VOTAT, which would be worth exploring.

The students with the mixed-strategy outperformed the other two kinds in terms of CPS performance, the depth-first strategy ranked the second and the last was the breadth-first strategy. A hypothesis may exist that the difference among the three groups may be related to the cognitive load. From the perspective of cognitive psychology, problem solving is defined as information-seeking in problem space from the initial state to the target state (Newell & Simon, 1979). As the breadth-first strategy involves the exploration in both variable types and values, the cognitive load is increasing rapidly with multiple attempts. However, the capacity of human information processing is limited

(Miller, 1956) and this is why students with the breadth-first strategy would be more likely to fail. By contrast, the depth-first strategy only involves the attempts in different variable values, which delivered a low cognitive load and therefore possible better performance. Furthermore, the mixed-strategy group achieves a more efficient way of information integration with a lower level of cognitive load by focusing on only one variable and value in each step, which can be considered as a higher-level variant of the VOTAT strategy.

According to the previous findings, strategies play a crucial role in solving problems. Teachers could actively guide students in using the effective strategies and continuous training can be offered to help students improve metacognitive skills.

References

Ball, L. J., Evans, J. S., Dennis, I., & Ormerod, T. C. (1997). Problem-solving Strategies and Expertise in Engineering Design. *Thinking & Reasoning, 3*(4), 247–270. doi: 10.1080/135467897394284

Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. In P.A. Frensch & J. Funke (Eds.), Complex problem solving: *The European perspective* (pp. 27-63). Hillsdale, NJ: Erlbaum.

Croker, S., & Buchanan, H. (2011). Scientific reasoning in a real-world context: The effect of prior belief and outcome on children's hypothesis-testing strategies. *British Journal of Developmental Psychology, 29*(3), 409–424. doi: 10.1348/026151010x496906

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning, 7*(1), 69–89. doi: 10.1080/13546780042000046

Greiff, S., Molnár, G., Martin, R., Zimmermann, J., & Csapó, B. (2018). Students' exploration strategies in computer-simulated complex problem environments: A latent class approach. *Computers & Education, 126*, 248–263. doi: 10.1016/j.compedu.2018.07.013

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education, 91*, 92–105. doi: 10.1016/j.compedu.2015.10.018

Heinström, J. (2005). Fast surfing, broad scanning and deep diving. *Journal of Documentation, 61*(2), 228–247. doi: 10.1108/00220410510585205

Jungnickel, D. (2008). Graphs, Networks and Algorithms. *Algorithms and Computation in Mathematics*. doi: 10.1007/978-3-540-72780-4

Kreher, D. L., & Stinson, D. R. (1999). Combinatorial algorithms. *ACM SIGACT News, 30*(1), 33–35. doi: 10.1145/309739.309744

Kuhn, D., & Dean, D. (2005). Is Developing Scientific Thinking All About Learning to Control Variables? *Psychological Science, 16*(11), 866–870. doi: 10.1111/j.1467-9280.2005.01628.x

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. doi: 10.1037/h0043158

Molnár, G., & Csapó, B. (2018). The Efficacy and Development of Students' Problem-Solving Strategies During Compulsory Schooling: Logfile Analyses. *Frontiers in Psychology, 9*. doi: 10.3389/fpsyg.2018.00302

Newell, A., & Simon, H. A. (1979). *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.

OECD. (2013). *PISA 2012 Assessment and Analytical Framework*, 119–137. doi: 10.1787/9789264190511-6-en

OECD. (2014). *PISA 2012 Results: Creative Problem Solving (Volume V)*. doi: 10.1787/9789264208070-table57-en

Rea, L. M., & Parker, R. A. (2014). *Designing and conducting survey research: A comprehensive guide*. San Francisco: Jossey-Bass.

Tschirgi, J. E. (1980). Sensible Reasoning: A Hypothesis about Hypotheses. *Child Development, 51*(1), 1. doi: 10.2307/1129583

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence, 40*(1), 1–14. doi: 10.1016/j.intell.2011.11.003

Wüstenberg, S., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The Role of Strategy Knowledge for the Application of Strategies in Complex Problem Solving Tasks. *Technology, Knowledge and Learning, 19*(1-2), 127-146. doi:10.1007/s10758-014-9222-8