Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Educational Practice for Interpretation of Experimental Data Based on a Theory

Hitomi SAITO[a*], Kazuhisa MIWA[b], Nana KANZAKI[c], Hitoshi TERAI[b,d], Kazuaki KOJIMA[e], Ryuuichi NAKAIKE[f], Jyunya MORITA[g]

[a]*Faculty of Education, Aichi University of Education, Japan*
[b]*Graduate School of Information Science, Nagoya University, Japan*
[c]*Department of Early Childhood Education, College of Nagoya Women's University, Japan*
[d]*JST/CREST, Japan*
[e]*Learning Technology Laboratory, Teikyo University, Japan*
[f]*Graduate School of Education, Kyoto University, Japan*
[g]*School of Knowledge Science, Japan Advanced Institute of Science and Technology, Japan*
* hsaito@auecc.aichi-edu.ac.jp

Abstract: Interpreting experimental data based on a psychological theory requires understanding the mechanisms or factors underlying cognitive processes and acquiring an attitude for interpreting evidence from a theoretical perspective. In this study, we designed and practiced teaching and learning activities using cognitive models to foster both requirements in an introductory course of cognitive science. Fifty-three undergraduate students attended the course. During practice, students constructed a computational model on the process of semantic memory and conducted simulations using their model. We evaluated changes in learner interpretation of experimental data from pretest to posttest. The results of the practice showed that students' interpretations of experimental results for semantic memory changed from pretest to posttest. However, their interpretations of the results of other experiments did not show much difference between pretest and posttest.

Keywords: Cognitive Model, Evaluate evidence, Mental model, Educational practice

1. Introduction

Evaluating and interpreting evidence is one of the most important activities in the scientific discovery process. Klahr & Dunbar (1988) proposed scientific discovery as a dual search (SDDS) consisting of a set of three basic components (search the hypothesis space, test hypotheses, evaluate evidence) to guide searches within and between two problem spaces. Searching the hypothesis space is the process of generating a hypothesis. Testing the hypothesis involves searching the experimental space. Evaluating evidence is the process that mediates between the two spaces. This process assesses the fit between theory and evidence and guides further searches.

Interpreting experimental data based on a psychological theory requires understanding the mechanisms or factors underlying cognitive processes and acquiring an attitude for interpreting evidence from a theoretical perspective. In this study, we designed teaching and learning activities using cognitive models to foster both of them in an introductory course of cognitive science.

People build mental models to understand the mechanisms of complex objects and phenomena. A mental model is a mental representation of how a system (object) operates (Norman, 1983). Studies on scientific education or human–computer interaction (HCI) have investigated how learners or users construct a mental model and how to assist them in building such models (White, 1993; Biswas, Leelawong & Schwartz, 2005; Kulesza, Stumpf, Burnett & Kwan, 2012).

Previous studies show that a mental model is effective for understanding the target phenomena or control systems. In addition, interactions between learners (or users) and the target (such as natural phenomena or systems) are important for building mental models.

When learners base a mental model on human cognitive processes, the target system is the human cognitive system. Because human cognitive processes are invisible, they are difficult for learners to control and observe. Therefore, we used a cognitive model running on computers instead of humans. We expected that interacting with a cognitive model by constructing and running models would help learners build a mental model based on a target cognitive process.

Furthermore, using a cognitive model would facilitate learners to acquire an attitude of interpreting experimental results based on psychological theory. Anderson (1993) and Miwa (2009) argue that cognitive models mediate between data and theory. An explanation of theory level is too abstract to compare it directly with experiment data such as performance or behavior. In contrast, cognitive models are specific enough and can be compared with experimental data. Cognitive models provide prediction results that assume cognitive processes in same tasks or situations as in the experiment. That is, theories can be evaluated for their validity when a cognitive model is constructed.

This role of a cognitive model could be effective for bridging the gaps between learners' interpretations of experimental data and their understanding of theories. Therefore, in this study, we designed teaching and learning activities that assisted learners in developing an attitude for interpreting experimental results based on the given theory by constructing cognitive models and comparing the behaviors of the cognitive models with the experimental results.

In this study, a computational model running on a production system was used as a cognitive system to interact with learners. We used a web-based production system architecture called "DoCoPro" (Nakaike, Miwa, Morita & Terai, 2011), which is easy to build and is used to simulate computational models for beginners.

2. Practice

2.1 Class and Learning Contents

We designed teaching and learning activities for an intensive introductory course on cognitive sciences at the Aichi University of Education titled "Cognitive Science 1." Students learned about the methodologies and key topics of cognitive science and cognitive psychology. Fifteen lessons were conducted over a 5-day period. Cognitive models were used for six lessons conducted during the first two days. Fifty-three undergraduate students majoring in informational science (31), science (13), clinical psychology (6), and other fields (3) attended the course. We analyzed data from fifty students who attended all lessons on cognitive modeling.

We used a classical study of semantic memory developed by Collins & Quillian (1969) as learning content for these classes. Collins & Quillian conducted experiments using a true-false reaction-time task to test their hypothesis on the organization of semantic memory, which stated that semantic memory was organized in a hierarchical structure, as is shown in Figure 1. The true-false reaction-time task in their experiment used sentences such as "A canary can fly," to which participants answered true or false. If the hypothesis was true, the reaction time of a true-false task would change depending on the number of links between the nodes.

Figure 2(a) shows the results of Collins & Quillian's experiments. This figure shows that the reaction time increased as the number of links increased. Many researchers have examined Collins & Quillian's hypothesis, and some studies propose other hypotheses about the organization of semantic memory. However, Collins & Quillian's work often appears in many cognitive psychology textbooks. In addition, their model is simple enough to allow beginners to build a computational model. Therefore, we selected their study for our course's learning content.

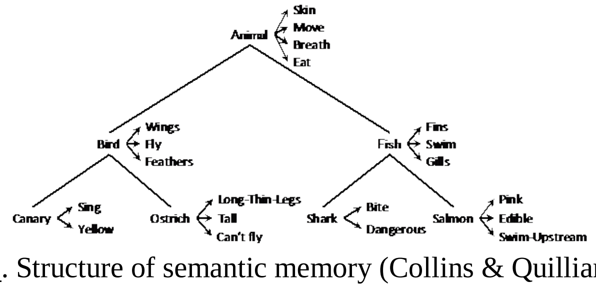
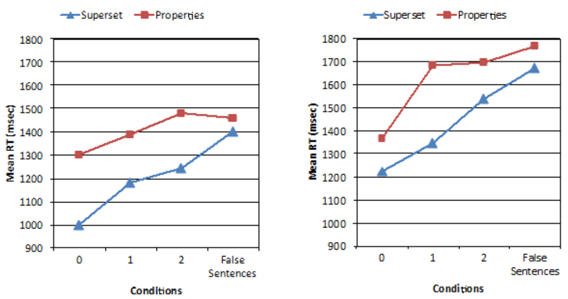


Figure 1. Structure of semantic memory (Collins & Quillian, 1969).

2.2 Learning Environment

Students used the web-based production system “DoCoPro” during the class. Figure 3 shows the interface window. On the left side of the window are displayed educational materials, and on the right side are displayed the components of the production system. Students can build and execute computational models as educational materials. The details of DoCoPro have been described in a previous study by Nakaike, Miwa, Morita & Terai (2011). Educational materials were constructed to help students learn at their own comfortable speed. These materials also included some quizzes to allow students to check their understanding of the materials.



(a) Collins & Quillian (b) Our practice

Figure 2. Results of two experiments.

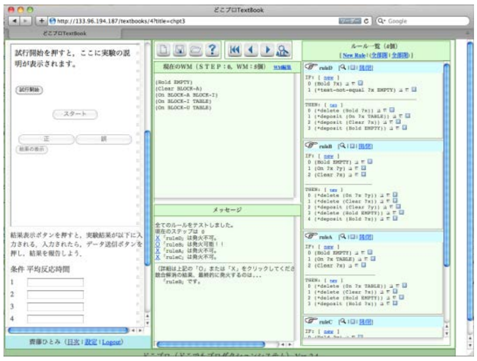


Figure 3. Interface window in DoCoPro.

2.3 Schedule of lessons

There were six lessons in our practice. Each lesson lasted for 90 min. In the first two lessons, we introduced approaches using cognitive modeling and taught students about the basics of the production system using DoCoPro. Students were given an ID and a password to log into the system. They learned about the key concepts of the production system, such as working memory, production rule, production rule firing, and bounding variable values using a block-stacking task.

Lessons 3–6 contained the main phases of our study. In Lesson 3, we introduced semantic memory, and students performed an experimental task similar to that conducted by Collins & Quillian (1969). The experiment was conducted over the Internet. After the experiment, we explained Collins & Quillian’s (1969) study. Next, students were asked to interpret the two experimental results in a posttest. Problem 1 consisted of interpreting the results of the above-explained experiment that they performed. Students were shown Figure 2(b) and required to describe the results and explain why they occurred. Problem 2 consisted of interpreting the results of an experiment that students did not perform. They were shown a brief overview of the experiment and its results and required to describe the same thing that they had investigated in Problem 1.

During Lessons 3, 4, and 5, students built a cognitive model based on the theory proposed by Collins & Quillian (1969). Students built a model at their own pace based on educational materials provided to them along with examples of descriptions of the knowledge structure in working memory and explanation of the rules. Students created the model to fill a gap in the given production rules based on the rules expressed verbally. After creating the model, the students then checked its operation through some true–false tasks.

In Lesson 6, students performed simulations and took a posttest. Before conducting simulations, they selected problems from a problem list to use in the true-false tasks in their simulations. The problem list consisted of 48 problems in each condition (supersets and properties). Students were asked to choose eight sentences and simulate four levels—Link1, Link2, Link3, and False sentences—in each condition. Then, they conducted simulations.

2.4 Pretest and posttest

There were two subjects in the pretest and posttest. Task 1 was conducted to directly evaluate the effects of practice. Task 2 was conducted to evaluate whether the knowledge that the students gained was applicable to the interpretation of other experimental results. The pretest and posttest used the same task. Each test lasted for 10 min.

In Task 1, we presented the experimental results of Collins & Quillian, and the students described the results and interpretation of the experiment.

In Task 2, we presented an outline and results of the experiment conducted by Bower (1979), which investigated the effect of schema on the remembrance of a text, and the students described the results and interpretation of the experiment. The outline of Bower's experiment was as follows. A subject read 18 kinds of texts and was asked to recall the actions described in the texts. The subject was assigned to three conditions containing three, two, or no similar stories in 18 kinds of texts. The action that the subject remembered was classified into 1) the action described in the text and usually performed (stated script actions); 2) the action not described in the text but usually performed (unstated script actions); and (3) the action not described in the text and seldom performed (other actions). The experimental results suggested that the remembrance rate of an action not described in the text but usually performed increases with an increase in the number of similar stories read.

3. Results

We analyzed students' descriptions in the pretest and posttest to evaluate whether building a cognitive model and simulation helps learners to interpret experimental results based on the given theory. First, we compared students' descriptions of Task 1 in the pretest with those in the posttest to directly confirm the learning effects of our practice. Next, we investigated students' descriptions of Task 2 to confirm that learning activities influence the interpretation of other psychological experiments.

3.1 Analysis of learners' descriptions in Task 1

We classified learners' descriptions in Task 1 in terms of descriptive content and level of interpretation. Descriptive content contained three subcategories: (a) descriptions focused on differences in the number of links; (b) descriptions focused on differences between the supersets and properties; and (c) descriptions focused on differences between false sentences and others.

The levels of interpretation were determined by referring to the types of knowledge used in science as proposed by Clement (1989). Clement classified knowledge used in science into four levels: observations, empirical low hypothesis summarizing an observed regularity, exploratory model hypothesis, and formal principles. On the basis of these types, we defined three levels of interpretation:

- Descriptions of the facts shown in the experimental results (Fact): This level corresponds to the observations of Clement's (1989) levels of knowledge.
- Descriptions of interpretations not based on the theory and model (No Model): This level corresponds to Clement's empirical low hypothesis.
- Descriptions of interpretations based on the theory and model (Model): This level corresponds to Clements' exploratory model hypothesis and formal principles.

Table 3(a) shows the number of descriptions in each level in the three descriptive contents included in Task 1. Label "O" means the number of students who did not provide descriptions on the descriptive content. In the pretest and posttest, a chi-square test about the total number of descriptions was conducted to analyze the differences in each category (M, NM, F, and O). As a result, the number of descriptions of the four categories were significantly different in the pretest and posttest ($\chi^2(3)$ = 11.350, p < .01). The results of a residual analysis of the interpretations based on the theory and model (M) show that there were few descriptions in the pretest but significantly more descriptions in the posttest. In the interpretations not based on the theory and model (NM), there were more descriptions in the pretest but few in the posttest.

Next, the same analysis was conducted for every type of descriptive content. As a result, in the false sentence, the number of descriptions had a significant difference in the pretest and posttest ($\chi^2(3)$ = 9.541, p < .05). As a result of residual analysis, in the interpretations based on the theory and model (M), there were marginal differences, with there being few descriptions in the pretest and more descriptions in the posttest. In the interpretations not based on the theory and model (NM), there were more descriptions in the pretest and few in the posttest. In supersets and properties, the number of descriptions were marginally different in the pretest and posttest ($\chi^2(3)$ = 6.256, p < .10). As a result of residual analysis, in the interpretations based on the theory and model (M), there were marginally significant differences, with few descriptions in the pretest and more in the posttest. In the number of links, there were no significant differences in the number of descriptions in the pretest and posttest.

These results show that students' descriptions changed from the interpretation in a way that is not based on the theory and that the model of interpretation in the posttest was based on the theory and model.

Table 3: Number of descriptions in each level in the three descriptive contents.

(a) Task 1

	Pretest				Posttest			
	M	NM	F	O	M	NM	F	O
The number of links	9	16	15	10	12	13	12	13
The supersets and the properties	7	18	14	11	17	12	14	7
The false sentences and others	8	13	15	14	16	3	19	12
The total number of descriptions	24	47	44	35	45	28	45	32

(b) Task 2

	Pretest				Posttest			
	M	NM	F	O	M	NM	F	O
Number of stated script actions	12	29	8	1	19	20	7	4
Number of unstated script actions	5	20	8	17	8	19	7	16
Other actions	0	12	21	17	4	8	15	23
The total number of descriptions	17	61	37	35	31	47	29	43

3.2 Analysis of learners' descriptions in Task 2

We classified learners' descriptions in Task 2 according to descriptive content and level of interpretation. Descriptive content contained three subcategories: (a) descriptions focusing on differences in stated script actions; (b) descriptions focusing on differences in unstated script actions; and (c) descriptions focusing on differences in the other action. The level of interpretation was determined to be one of three levels based on the same criteria as was the analysis of Task 1.

Table 3(b) shows the number of descriptions in each level in the three descriptive contents in Task 2. Label "O" means the number of students who did not provide descriptions on the descriptive content. In the pretest and posttest, a chi-square test about the total number of descriptions was conducted to analyze the differences in each category (M, NM, F, and O). As a result, in the total number of descriptions, the number of descriptions had a marginal difference in the pretest and posttest ($\chi^2(3) = 7.688$, $p < .10$). As a result of residual analysis, in the interpretations based on the theory and model (M), there were marginal differences, with few descriptions in the pretest and more in the posttest.

In the analysis for every type of descriptive content, in the unstated script actions, the number of descriptions had a marginal difference in the pretest and posttest ($\chi^2(3) = 6.700$, $p < .10$). This was the same result as that of the analysis of the total number of descriptions, as shown through residual analysis. There was no significant difference in other descriptive content (the stated script actions: $\chi^2(3) = 5.100$, *n.s.*; the other action: $\chi^2(3) = 0.815$, *n.s.*).

These results show the tendency that students' descriptions change from the interpretations not based on the theory and model to the interpretations that are based on the theory and model in the posttest.

4. Discussion and Conclusion

In this study, we designed and practiced teaching and learning activities using computational models to assist in the interpretation of experimental data based on a given theory. In our experiment, students constructed a computational model for the process of semantic memory and conducted simulations using their model. The results of the practice show that students' interpretations of experimental results for semantic memory in Task 1 changed from pretest to posttest. These results indicate that the learners could elaborate upon their mental model for the target cognitive processes and interpret experimental data based on the given theory through learning activities.

Compared to Task 1, there were marginal differences between the pretest and posttest in Task 2. Task 2 was an experiment regarding schema memory and is difficult to interpret using the theory of semantic memory. As a result, it was difficult for learners to build a mental model for cognitive processes in Task 2 and interpret experimental data based on the theory and model. In addition, some descriptions in the interpretations based on the theory and model (M) interpreted experimental results based on the theory of semantic memory that the students learned. These descriptions resulted in the learners' misapplication of the theory of semantic memory.

In the educational practice, learners constructed a cognitive model based on the theory proposed by Collins & Quillian. They were also given explanations about how to represent the knowledge structure and production rules from educational materials. Learners translated the given theory into a computational model running in the production system. As a result, these learning activities affected the interpretations of experimental data based on the given theory, as shown by the results of Task 1. However, the effect on the interpretation of experimental data was not so large as to require learners to construct individual theories or models. Therefore, the challenge will be to assist learners in generating individual theories and models.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research B (23300100) from The Ministry of Education, Culture, Sports, Science and Technology.

References

Anderson, J. R. (1993). *Rules of the Mind*. Lawrence Erlbaum Associates, 365 Broadway, Hillsdale, NJ 07642.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & Vanderbilt, T. T. A. G. at. (2005). LEARNING BY TEACHING: A NEW AGENT PARADIGM FOR EDUCATIONAL SOFTWARE. *Applied Artificial Intelligence*, *19*(3-4), 363–392.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220.

Clement, J. (1989). Learning via model construction and criticism. In G. Glover, R. Ronning, & C. Reynolds (Eds.), *Handbook of creativity: Assessment, theory and research* (pp. 341–381). New York: Plenum Publishers.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240–247.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive science*, *12*(1), 1–48.

Kulesza, T., Stumpf, S., Burnett, M., & Kwan, I. (2012). Tell me more?: the effects of mental model soundness on personalizing an intelligent agent. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1–10). New York, NY, USA: ACM.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner and A. L. Stevens (Eds.), *Mental models* (p. 7-14). Hillsdale, NJ: Lawrence Erlbaum.

Nakaike, R., Miwa, K., Morita, J., & Terai, H. (2011). Development of a Web-based Production System for Introductory Cognitive Science Classes. *Transactions of the Japanese Society for Artificial Intelligence*, *26*(5), 536–546.

Miwa, K. (2009). Cognitive Models as Hypothesis Deduction Machine, Cognitive Simulator, and Data Analytic Machine(<Special Issue>Model-Based Approach in Cognitive Science). *Journal of Japanese Society for Artificial Intelligence*, *24*(2), 229–236.

Miwa, K., Terai, H., Morita, J., Nakaike, R., & Saito, H. (2012). Class Practise of Cognitive Science by Creating Models. *Transactions of the Japanese Society for Artificial Intelligence*, *27*(2), 61–72.