Long-term Use of Learning Environment for Problem-Posing in Arithmetical Word Problems

Tsukasa Hirashima[a], Takuro Yokoyama[b], Masahiko Okamoto[c], Akira Takeuchi[b],
[a]*Hiroshima University, Japan*
[b]*Kyushu Institute of Technology, Japan*
[c]*Osaka Prefecture University, Japan*
tsukasa@isl.hiroshima-u.ac.jp

Abstract: An experimental study on a computer-based environment for learning by problem-posing has been developed for the students. A long-term evaluation with the system was carried out for two months by testing it in a classroom composed by 99 students belonging to three second grade classes at an elementary school. Two systems were arranged in each class for the students and they were allowed to use the systems freely only during out-of-class time. We analyzed problem-posing logs in the system, the results of questionnaires, information extraneous problem test, and schema priming test for this experiment. From the output of this study, we found that: (1) our system improved the problem solving ability of low performing students, (2) the second grade students posed problems continuously using the system, and (3) both students and teachers answered questionnaires which showed that the problem-posing activity using this system was interesting and useful for learning.

Keywords: Problem-Posing, Problem Schema, Arithmetical Word Problem

Introduction

Learning by problem posing is well recognized as an important way to learn arithmetic or mathematics [1-4]. In order to realize learning by problem-posing, the way to assess and give feedback to posed problems is one of the most important issues. We have been investigating computer-based learning environments that can assess and give feedback to each posed problem automatically (we call this way of assessment “agent-assessment”, in contrast with "teacher-assessment" or "peer-assessment"[5, 6]). The framework of the learning environment is composed of (1) problem-posing interface, (2) problem diagnosis function, and (3) help function for correcting or completing the posed problems. We have already developed several learning environments for interactive problem-posing in arithmetical word problems [7-9]. We, then, have used them in several elementary schools and confirmed that a problem-posing exercise with the learning environment is effective to improve both problem-solving and problem-categorization abilities [10, 11].

The above researches, however, had two problems. The first problem is the effect of learning. The researches had a short-term learning paradigm with pre-post comparison. Particularly, they took only about one hour for students getting engaged in the problem-posing activity with the system. Therefore, it was unclear that whether students were able to gain stable abilities. The second problem is an intention of learning. In the previous studies, students were made to use the system within class time. If the system provides students with an attractive learning environment, they themselves use the system spontaneously.

So, as the next step of this research project, we planned to use the learning environment in an elementary school second grade for two months [12]. For this experiment, several computers were installed to set the learning environment in several classrooms and students were asked to pose problems freely by using the computers out-of-class time. Through this practical usage, we evaluated the effect of the learning environment and confirmed that whether it can be used by students by their own initiative.

In Section 1, a learning environment for problem-posing named: MONSAKUN (means "Problem-Posing Boy" in Japanese) is introduced. In the learning environment, a learner can pose arithmetical word problems that can be solved using an addition or subtraction, by combining several sentence cards. The framework of the practical use is explained in Section 2 and the results are reported in Section 3.

1. Learning Environment for Problem-Posing: MONSAKUN

The interface of problem-posing in MONSAKUN is shown in Figure 1. The area in left side, imaged blackboard, is "problem-composition area". At the top, a calculation expression has been given. A learner has to pose a problem which will be solved by the calculation expression, that is, by an addition or subtraction. Several sentence cards are presented at right side of the interface. In order to pose a problem, the learner selects several sentence cards and arranges them in a proper order. Although interpretation of each sentence is easy, the learner has to consider the relation among them to pose an adequate problem including the suitable relation for the calculation expression. This process is usually called “sentence integration” [13] where so-called “problem schema” plays a crucial role [14, 15]. Therefore, this activity is expected to sophisticate the problem schema.

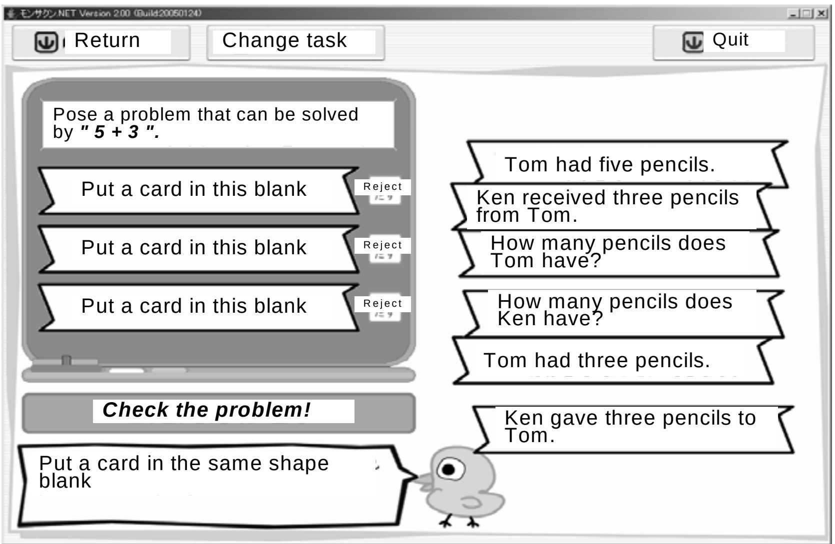

Figure 1. Problem-Posing Interface of MONSAKUN (Currently, it can deal only in Japanese language. All words were translated into English for this paper).

A sentence card is put into a blank in the problem-combination area. There are three blanks in Figure 2, a learner should select three cards from the card set at right side and arrange them in a proper order. A learner can move a card by drag & drop method in the interface. When a learner pushes "diagnosis button" under the problem-composition area, the system diagnoses the combination of sentences. The results of the diagnosis and message to help the learner's problem-posing is presented by another window.

As exercises of problem-posing, four problem-structure types and three blank types were prepared. The four problem-structure types were composed of (I) change problem-increase, (II) change problem-decrease, (III) combine problem, and (IV) compare problem. This categorization is a typical one used in the elementary schools. Three exercise blank types were composed of (i) one blank, (ii) two blanks, and (iii) three blanks. In total, 300 exercises were prepared. The number of combinations of the sentence cards is more than five even in the easiest exercise with one blank. As for the most difficult exercise with three blanks, there are two hundred and ten combinations of the sentence cards. Therefore, it is almost impossible to complete correct problems at random. This diagnosis function in MONSAKUN is described in more details [11].

2. Experimental Use of MOSAKUN in Classroom

2.1 Purposes and Measurement

In this experiment, several computers with MONSAKUN were arranged in classrooms. For nine weeks, 99 students belonging to three different classes of second grade elementary school was allowed to use the system freely during out-of-class time. Purposes of this experiment: (1) to examine the learning effect of a long-term use of problem-posing with MONSAKUN, and (2) to confirm whether students could use MONSAKUN on their own free will.

For the first purpose, we used (a) extraneous problem test and (b) schema priming test. (a) An extraneous problem includes extraneous information that is not necessary to solve the word problem [16, 17]. In the following example, the third point has been considered as an extraneous one. {1. There are two apples. 2. There are three oranges. 3. There are seven bananas. 4. How many apples and oranges are there in total?}. Because this problem has some difficulty mainly in the process of sentence integration where the problem schema plays a crucial role, this test is suitable to examine the effect of this learning environment.

The schema priming test has been proposed in this research, as another test to measure the availability of the problem schema in problem solving. Several studies have reported that if a question sentence is previously submitted to a subject before presenting all the sentences, he/she can quickly judge whether the sentence is enough to answer the question or not, because his/her problem schema has been activated by the question sentence beforehand[18]. In order to judge, the subject is required to complete an internal representation of the problem. In the schema priming test, a question is first presented as prime information, and then the all problem sentences are presented as target information. The response time to judge whether the problem can be solved or not, is measured as an indicator of availability of the problem schema. If the availability of the problem schema is high, it is expected to be activated quickly, and then the judgment of the problem sentences would be completed in short time. In Figure 2, all the sentences are presented in the schema priming test. Before presenting the whole sentences as target

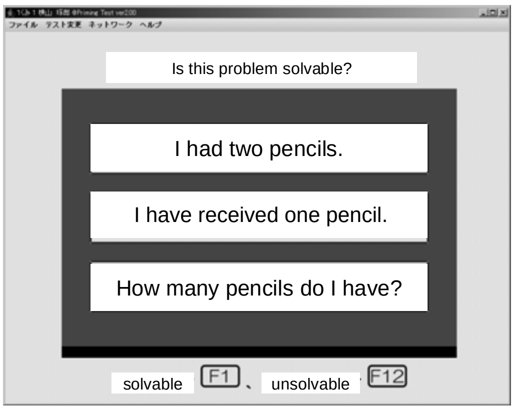

Figure 2. Schema Priming Test

information, only the third sentence "How many pencils do I have?" was presented as the prime information. A student is required to push "F1" key when the student judges the problem is solvable, and to push "F12" key when the student judges the problem is unsolvable.

For the second purpose, we examine the total number of problems posed by individual students and the results of questionnaire.

2.2 Situation of the Experimental Use

Before this experimental use, two class times (90 min. in total) were taken as introductory use of MONSAKUN in a computer room of an elementary school where each student could able to use one computer. Then, two computers installed with MONSAKUN were placed in each class (six computers were used in total). Here, about fifteen students were assigned for a computer. The class teachers were involved only in making rules to the students for sharing the systems but not for directing the students to use them. This experimental use was carried out for nine weeks including 46 school working days. We conducted a pre-test just after the introductory use, an intermediate test during the mid-way of the period, and a post-test at the end of the period. In both the pre- and post-tests, students took the extraneous problem and schema priming tests. However, in the intermediate test, they took only the schema priming test because of time limitation as per the school schedule.

3. Evaluation of MONSAKUN

3.1 Results of System Use

A total of 8,386 problems was posed by the six systems. In a day, 30.4 problems in average was posed with a system. Figure 3 shows the number of students and the number of used days. In average, 8.5 days was used by a student. In summary, three students used a system for a day and each of them posed ten problems. Table 1 shows the number of posed problems for every two weeks. In the middle period, the number of posed problems was reduced but more than ten problems were posed at least. Therefore, it could be confirmed that the system was used continuously in the period. Table 2 shows the results of

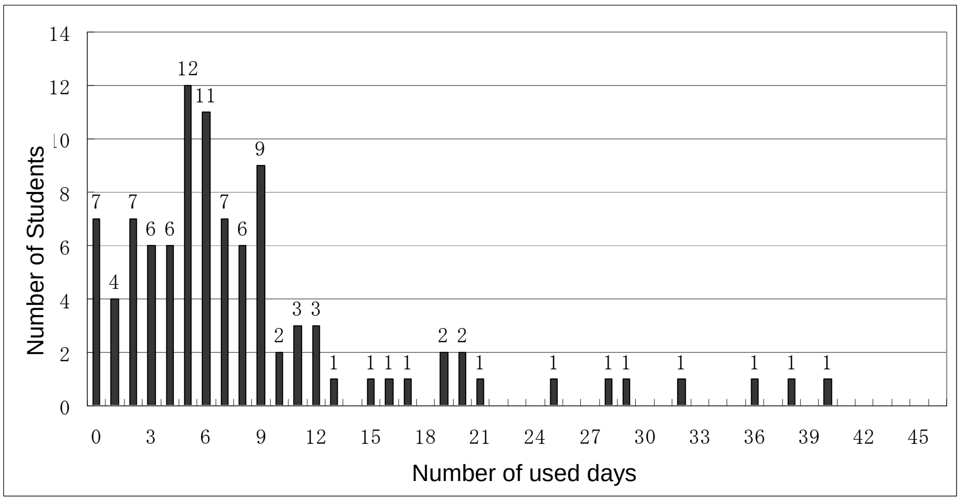

Figure 3. Number of used days of students.

questionnaire distributed at the end of the experiment. It was found that, most of the students felt that the usage of MONSAKUN made arithmetic enjoyable. It was also very useful for learning arithmetic, and students are interested to use it more often in the near future. Teachers who were involved in these activities also agreed with the students.

Table 1. Number of Posed Problems.

Period	1-2 weeks	3-4 weeks	5-6 weeks	7-9 weeks
Number of Posed Problems	2666	1149	722	3849

Table 2. Results of Questionnaires.

Question \ Answer	Yes	No	No idea
Do you think MONSAKUN make arithmetic enjoyable?	84	1	6
Do you think MONSAKUN is a game?	52	21	18
Do you think MONSAKUN is useful for arithmetic learning?	82	3	6
Do you like to use MONSAKUN more often?	80	2	9
Do you think you could make problems easier than before?	75	3	13

Since, more than fifteen students shared one system, students could not always use the system whenever they like. Besides, the available time of the systems was only out-of-class time. Considering these restrictions, above results suggest that the second grade students were continuously able to pose problems with MONSAKUN by their own will, and accepted its usage for learning with enjoyment.

3.2 Learning Effect

We have analyzed the results of 85 students whose data were posed as problems, and every score of the test was completely gathered. In the analysis, based on the average score (= 8.32) of extraneous problem test, the students were divided into two groups: a high-score and low-score groups. Then, the students were also divided into high-posed group and low-posed group based on the median (= 77) of the number of posed problems by each student. As a result, we found that the number of a high-score & high-posed group is thirty-two, a high-score & low-posed group is twenty, a low-score & high-posed group is twelve, and a low-score & low-posed group is twenty-one.

3.2.1 Results of Extraneous Problem Test

The extraneous problem test was composed of twelve problems. When a student correctly wrote both an expression and an answer, one point was given. Hence, total score was twelve points. The results of the extraneous problem test in the pre/post-tests are shown in Table 3.

In order to examine the validity of the group categorization, the scores of the pre-test were analyzed by two-factor analysis of variance with the score (high-score vs. low-score) and the number of posed problems (high-posed vs. low-posed). It was found that only the main effect of score was significant ($F(1, 81) = 195.42$, $p<.01$), and the score of the high score group was higher than the low score group. This result indicates that the categorization with the score is adequate. Because the main effect of the number of posed problems was not significant, it was confirmed that the number of posed problems has no effect on the scores of the pre-test.

Table 3. Mean Scores of Extraneous Problem Test.

Condition	Pre-test	Post-test
High-score/High-posed (*n* =32)	11.06 (SD=1.16)	10.96 (1.28)
High-score/Low-posed (*n* =20)	10.65 (1.18)	10.50 (2.25)
Low-score/High-posed (*n* =12)	4.41 (2.78)	7.25 (2.52)
Low-score/Low-posed (*n* =21)	4.14 (3.04)	4.71 (3.81)

The scores of the post-test were also analyzed in the same way. It showed that the main effect of the scores was significant ($F(1, 81) = 67.81$, $p<.01$), and the score of the high score group was higher than the score of the low score group. The main effect of the number of posed problems also had significant effect ($F(1, 81) = 6,78$, $p<.05$), and the score of the high-posed group was higher than the low-posed group. Besides, because the interaction between the score and the number of posed problems showed a tendency of significance ($F(1, 81) = 3.21$, $p<.10$), Tukey's HSD post hoc test was used to compare all the four groups. There was no significant difference between the high-score/high-posed group and high-score/low-posed group, but the low-score/high-posed group had a significant low score when compared to the high-score/high-posed and high-score/low-posed groups (both $p<.01$). The low-score/low-posed group had a significant low score not only from both the high-score/high-posed and high-score/low-posed groups (both $p<.01$), but also from the low-score/high-posed group ($p<.05$). Even though the score of low-score/high-posed group had no significant difference from the score of the low-score/low-posed group at the pre-test, it became significantly higher than one of the low-score/low-posed group at the post-test. These results suggest that the long-term use of the system had a strong effect to improve the performance of problem-solving of the extraneous problems.

3.2.2 Results of Schema Priming Test

The reaction time of the schema priming test was measured at the time when a student pressed the answer key. The results of pre-, intermediate- and post-tests are shown in Table 4.

Table 4. Mean Reaction Time of Schema Priming Test.

Condition	Pre-test	Intermediate-test	Post-test
High-score/High-posed (*n* =32)	16.32(SD=7.61)	14.96 (6.14)	19.19(27.34)
High-score/Low-posed (*n* =20)	21.52 (14.97)	17.14 (11.76)	13.89(6.73)
Low-score/High-posed (*n* =12)	15.74(9.85)	13.82(3.84)	12.30(3.45)
Low-score/Low-posed (*n* =21)	34.05(21.90)	25.51(23.29)	24.08(16.52)

The reaction time was analyzed by three-factor analysis of variance with the score (high-score vs. low-score), the number of posed problems (high-posed vs. low-posed) and the test period (pre-test vs. intermediate-test vs. post-test). As a result, the main effect of the score was not significant, but the main effect of the number of posed problems was significant ($F(1, 81) = 7.62$ $p<.01$). Besides, because the interaction between the score and the number of the posed problems was significant ($F(1, 81) = 6.24$, $p<.05$), Tukey's HSD post hoc test was used to compare the all four groups. The results indicated that the reaction time of the low-score/low-posed group was significantly late ($p<.05$). Therefore, it is suggested that the low-score/high-posed group showed the similar availability of the problem schema with the high-score/high-posed and high-score/low-posed group, but the availability of the low-score/low-posed group was lower than the ones of the other groups. Because the test period showed a tendency of significan*ce* ($F(2, 162) = 2.62$, $p<.10$), Tukey's HSD post hoc test was carried out. As a result, there was a tendency of significance ($p<.10$) between the reaction times of the pre-test and the post-test. Because a reduction tendency of the reaction time was appeared without any regard to the score or the number of the posed problems, it is suggested that the effect came from the experience to take the schema priming test rather than to use the system.

3.3 Considerations

The main target of learning by problem-posing implemented in MONSAKUN is the sentence integration process in problem solving for word problems. In the integration process, several investigations have been carried out to indicate that the internal representation of problems are generated with a problem schema, and mistakes of the problem solving mainly comes from failures in completing integrated representation of the whole problem [19]. Following this idea, students in the two groups with low-score had difficulty in the integration process. The results of the extraneous problem test indicated that students in the low-score/high-posed group improved their ability through the long-term use of the system. The analysis suggested that even though students have difficulty in the integration process, they could able to improve their performance, if they posed a lot of problems continuously. On the other hand, we could not confirm the effect for students in the high-score group.

The results of schema priming test indicated that reaction time of the students in the low-score/low-posed group were significantly slow. This suggests that their availability of the problem schema was low. However, results of the low-score/high-posed group not providing the same suggestion. Considering with that the low-score/high-posed group improved their ability but the low-score/low-posed group did not. It is possible to interpret that the students in the low-score/high-posed group had already had ability to use the problem schema to some degree at the pre-test. Therefore, they could pose problems actively and improve their ability through the use of the system. In contrast, it is also interpreted that because the students in the low-score/low-posed group did not have enough ability to use the problem schema, they could not pose problems actively and did not improve their performance. Of course, this is only an interpretation and it is necessary to examine the results of schema priming test in more details as future works.

4. Concluding Remarks

In this study, we have developed a computer-based environment for learning by problem-posing for the elementary school students. The experiment was carried out with the aim of long-term evaluation using the system for two months in a classroom composed by 99 students belonging to three different classes of second grade elementary school. The results showed that: (1) our system improved the problem solving ability of low performing students, (2) the second grade students posed problems continuously using the system, and (3) both students and teachers answered questionnaires that the problem-posing activity using this system was interesting and useful for learning. However, thoroughly the concrete reasons for low performing students' gradual improvement were not examined, which is very important to make clear the role and effect of learning by problem-posing. It is imperative to study the sophistication of the model of learning by problem posing and extension of the target domain for our future works.

Acknowledgements

This research was partially supported by the Japanese Ministry of Education, Science, Sports, and Culture under Grant-in-Aid for the Scientific Research (C) (2).

References

[1] Ellerton, N.F. (1986) Children's Made Up Mathematics Problems: A New Perspective on Talented Mathematicians. *Educational Studies in Mathematics*, Vol.17, pp.261-271.

[2] Brown, S.I., Walter, M.I. (1993) *Problem Posing: Reflections and Applica*tions, Lawrence Erlbaum Associates.

[3] Silver, E.A., CAI, J. (1996) An Analysis of Arithmetic Problem Posing by Middle School Students, *Journal for Research in Mathematics Education*, Vol.27, No.5, pp.521-539.

[4] English, L.D. (1998) Children's Problem Posing Within Formal and Informal Contexts, *Journal for Research in Mathematics Education*, Vol.29, No.1, pp.83-106.

[5] Barak, M., Ragael, S. (2004). On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. *International Journal of Human-Computer Studie*s, 61, 84-103.

[6]]Yu, F. Y. Liu, Y. H., Chan T. W. (2004). A web-based learning systems for question-posing and peer assessment, *Innovations in Education and Teaching International*, 42(4), 337-348.

[7] Nakano, A., Hirashima, T., Takeuchi, A. (1999) Problem-Making Practice to Master Solution-Methods in Intelligent Learning Environment, *Proc. of ICCE'99*, pp.891-898.

[8] Hirashima, T., Nakano, A., Takeuchi, A. (2000) A Diagnosis Function of Arithmetical Word Problems for Learning by Problem Posing, *Proc. of PRICAI2000*, pp. 745-755.

[9] Nakano, A., Hirashima, T., Takeuchi, A. (2000) A Learning Environment for Problem Posing in Simple Arithmetical Word Problem, *Proc. of ICCE2000*, pp.91-99.

[10] Nakano, A., Hirashima, T., Takeuchi, A. (2002) An Evaluation of Intelligent Learning Environment for Problem Posing, *Proc. of ITS2002*, pp.861-872.

[11] Hirashima, T., Yokoyama, T., Okamoto, M., and Takeuchi A. (2007) Learning by Problem-Posing as Sentence-Integration and Experimental Use, *Proc. of AIED2007*, pp.254-261.

[12] Hirashima, T., Yokoyama, T., Okamoto, M., and Takeuchi A. (2008) An Experimental Use of Learning Environment for Problem-Posing as Sentence-Integration, *Proc. of ITS2008*, pp.687-689.

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

[14] Mayer, R. E. (1981) Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science*, 10: 135-175

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

[16] Muth, K. D. (1984) Solving arithmetic word problems: Role of reading and computational skills. *Journal of Educational Psychology*, Vol.76, pp.205-210.

[17] Muth, K. D. (1992) Extraneous information and extra steps in arithmetic word problems. *Contemporary Educational Psychology*. Vol. 17, pp.278-285.

[18] Robinson, C. S., & Hayes, J. R. (1978) Making inferences about relevance in understanding problems. In R. Revlin & R. E. Mayer (Eds.), *Human Reasoning*, 195-206. Washington: Winston.

[19] Cummins, D. D., Kintsch, W., and Reusser, K., & Weimer, R. (1988) The role of understanding in solving word problems. *Cognitive Psychology*, 20: 405-438