Effects of Computer-Based Diagrams on Solving Word Problems: A Longitudinal Study*

Hidetsugu TAJIKA

Department of School Education (Psychology), Aichi University of Education, Kariya, Aichi 448–8542 Japan

Abstract

To examine effects of computer-based diagrams on solving two types of word problems over a one-year period, fifth-grade students were allocated to one of three groups. Students in the training group generated diagrams themselves using a computer and students in the other two groups did not. After having engaged in each activity, students in each group solved ratio word problems three times during a year. They also received a transfer test in the second and final test periods. The results showed that students in the training group outperformed students in the other two groups on solving ratio word problems over a one-year period but that students in the training group solved the transfer test problems as well as students in the other two groups. The computer-based diagrams as an effective cognitive strategy for helping students build mental models of ratio word problems are discussed.

Aims

Solving Processes of Arithmetic Word Problems

It is well known that arithmetic word problems are difficult for elementary school children to solve. Is it possible to teach them how to solve such word problems correctly? One of the approaches to teaching is to help students understand the problem solving processes for arithmetic word problems and help students build mental models suitable for solving word problems. As a result, students can generate the number sentences or equations easily and carry out the appropriate mathematical operations (Mayer, 1985, 1987, 1992, 1999; Nathan et al., 1992; Tajika, 1996).

Arithmetic word problem-solving processes can be analyzed into two subprocesses, the problem comprehension process and the problem solution process (Hinsley, Hayes, & Simon, 1977; Kintsch & Greeno, 1985; Paige & Simon, 1966; Riley, Greeno, & Heller, 1983). The problem comprehension process is referred to as consisting of two further cognitive subprocesses, translation and integration processes, whereas the problem solution process is also referred to as consisting of two subprocesses, planning–monitoring and execution processes (Mayer, 1985, 1987, 1992, 1999).

In the translation process, a student has to translate each sentence from the word problem into an internal representation. As the student reads an arithmetic word problem, the student tries to comprehend information in each sentence of the arithmetic word problem. In the integration process, the student has to select and combine information into a coherent representation (which I call a mental model) of the entire problem. The student tries to build a mental model from the propositions of the arithmetic word problem by integrating the information.

In the planning–monitoring process, the student has to break the arithmetic word problem down into a series of steps and generate a plan for solving it. The student tries to develop and check a solution plan based on the mental model which she or he has generated. In the execution process, the student has to carry out the mathematical operations specified in the plan.

In order to solve word problems, students have
to generate mental models suitable for solving word problems. In the integration process, students construct mental models for solving word problems by connecting information from word problems internally and activating prior mathematical knowledge existing in students' memories.

**Using a Computer to Help Students Build a Mental Model**

It is well known that computer-based tools can help children build their mental models for solving word problems (e.g., Jaspers & Van Lieshout, 1994; Nathan et al., 1992; Tajika et al., 1995, 1997).

Tajika et al. (1995) investigated the role of a computer as an understanding facilitator for solving ratio word problems. They conducted an experiment in which fifth-grade students in an elementary school answered questions designed to tap each of four subprocess of word problems stated above. Two kinds of ratio word problems were used in the experiment, easy problems and hard problems. Easy problems required only one computational step to solve them, whereas hard ones required two steps to solve them. Students were assigned to one of three solution groups, the diagram and relational picture group (called the diagram-relational picture group), the diagram group, and the control group. Students in the diagram-relational picture group and in the diagram group solved questions designed to tap the integration process of the ratio word problem. Two kinds of questions were developed to tap the integration process of each ratio word problem. Students in the diagram-relational picture group and in the diagram group were asked to fill in appropriate numbers in the diagram in the space provided. When they filled in an appropriate number in the diagram on a computer, the next space was provided in the diagram. Students in the control group only selected a correct equation and a correct answer to solve the word problem, instead of answering the diagram questions. Students in the diagram-relational picture group were then tested on the relational picture questions. Students in the diagram group selected a correct equation and a correct answer to solve the word problem, instead of answering the relational picture questions. After answering the relational picture questions, students in the diagram-relational picture group selected a correct equation and a correct answer to solve the word problem like students in the other two groups did. The results of Tajika et al. (1995) showed that students in the diagram-relational picture group outperformed subjects in the other two groups on solving hard ratio word problems.

Tajika et al. (1997) further replicated the results reported in Tajika et al. (1995) by adding a group in which students answered only relational picture questions. The results showed that students in the diagram-relational picture group and in the relational picture group outperformed students in the diagram group and in the control group on the hard type of ratio word problems. There was no difference in performance between the diagram-relational picture group and the relational picture group.

The results of the above two experiments (Tajika et al., 1995, 1997) suggest that the interaction between students and the computer can make it easier for the students to build a mental model of ratio word problems. The importance of the interactive activities with a computer has been documented (e.g., Cognition and Technology Group at Vanderbilt, 1996). By students' interactive activities with a computer I mean that students try to construct the mathematical concept by manipulating concrete materials (e.g., relational pictures) which exemplify the concept of part and whole of the ratio.

Recently Tajika and Sakamoto (1999) conducted an experiment in which students generated a diagram of the ratio word problem. When students in the self-generated diagram group received a ratio word problem, they first generated a diagram themselves and then solved the ratio word problem. The ratio word problems were classified into two semantic types of ratio problems, the second semantic type of ratio problems and the third semantic type of ratio problems. When P means a ratio of A to B, the second semantic type of ratio problems consists of asking to find A (A=B × P). The third semantic type of ratio problems consists of asking to find B (B=A ÷ P). Students were given one horizontal center line and the ratio word problem. They were asked to generate a diagram which was consistent with the ratio word problem. Tajika and
Sakamoto (1999) showed that there was a significant interaction between group and semantic type. They showed that students who generated diagrams of the third semantic type of the ratio word problems themselves outperformed students who did not. They also showed that there was no difference in performance on the second semantic type of ratio word problems between the two groups. The second semantic type of ratio word problems was solved more easily than the third semantic type of ratio word problems. As a result, there was no generation effect of diagrams. However, many students did not solve the second semantic type of ratio word problems including a two-step solution process. A harder second semantic type of ratio word problems is used in the present study.

The Present Investigation

Building on the results of these prior studies on ratio word problem solving, I had three primary purposes. The first purpose of the present study was to examine the effect of the training in which students solved ratio word problems by generating diagrams themselves using a computer. Previous research (Tajika et al., 1995, 1997) suggests that the interactive activity on which students manipulated relational pictures with a computer are effective in solving ratio word problems. When self-generated diagrams are designed to be an interactive learning environment with a computer, students can generate mental models for solving ratio word problems.

To this end, three groups were constructed as a group variable, the training group, the extra lessons group, and the control group. Students in the training group received two forty-five-minutes training sessions for generating the diagrams of ratio word problems themselves using a computer. Students in the extra lessons group received extra lessons on solving ratio word problems. Students in the control group solved two kinds of extra ratio word problems similar to the ratio word problems employed at each test period. Students in each group engaged in activities for the same amount of time. Judging from the results of previous research (e.g., Jaspers & Van Lieshout, 1994; Nathan et al., 1992; Tajika et al., 1995, 1997; Tajika & Sakamoto, 1999), it was hypothesized that students who received the training would solve more ratio word problems than students in the other two groups. When computer-based diagrams are effective, it was also hypothesized that students would solve more second semantic type of word problems than the third semantic type of word problems. Computer-based diagrams are assumed to be more effective on both semantic types of ratio word problems.

The second purpose was to investigate whether differences in solving performances among the three groups might be maintained over a one-year period. To this end, students in each group solved ratio word problems three times during the year. In the first period, students solved ratio word problems immediately after they engaged in one of three activities, manipulating a computer (the training group), receiving the extra lessons (the extra lessons group), or solving extra word problems (the control group), respectively. They had the second test period five months later after the first test period. Finally, they had the final test period about one year later after the first test period. On the basis of previous research, it was predicted that students in the training group would maintain higher performance on the ratio word problems than students in the other two groups.

The final purpose of the present study was to examine whether facilitative effects on the self-generated diagrams would be transferred to solving the other type of word problem. In this study, word problems used as the transfer test were identical to those used in a previous study (Mayer, Tajika, & Stanley, 1991). The word problems used as the transfer test were different from the two semantic types of ratio word problems. Each word problem used as the transfer test tapped translation, integration, and planning skills in the subprocesses of the word problem solving. The transfer test was carried out in the second test period and in the final test period. The results of previous research (e.g., Klahr & Carver, 1988; Lee, 1998) suggest that effective transfer requires a sufficient degree of understanding of original learning. When students in the training group solved ratio word problems using their mental models more correctly than students in the other two groups, positive transfer would be expected.

Method

Participants

Participants were 80 fifth grade students in three elementary schools when this study started. However, the final sample included 66 students (30 girls, 36 boys), who participated in every test
Design

A 3 x 2 x 3 factorial design was carried out in the study. The first factor was group, a between-subject factor with three levels: students who had training about generating computer-based diagrams (the training group), students who had extra lessons about ratio word problems (the extra lessons group), and students who had solved a similar type of extra ratio word problems as the critical test problems (the control group). The second factor was problem type, a within-subject factor that had two levels: the second semantic type of word problems and the third semantic type of word problems. The last factor was time of test, a within-subject factor with three levels: the immediate test period after the training (February), the second test period (July), and the final test period (next February).

Materials

Two types of word problems were used in the study, one a critical test which consisted of ratio word problems and the other a transfer test which was devised by Mayer et al. (1991). The pretest was also used in addition to both tests stated above.

The critical test of ratio word problems included an eight-page booklet that consisted of 8 ratio word problems, including 4 problems of the second semantic type and 4 problems of the third semantic type. Each ratio word problem was selected from Tajika and Sakamoto (1999). For example, one ratio word problem of the second semantic type used in this study was as follows: “Masao is a fifth grader. There are 90 fifth graders in Masao’s school. The number of students in Masao’s class is 0.3 times as many as the total number of fifth graders. How many students are there in Masao’s class?” One ratio word problem of the third semantic type used in this study was as follows: “Yasuko has a red tape and a white tape. The white tape is 60 cm long and is 0.6 times as long as the red tape. How long is the red tape?”

The transfer test included an 18-item four-page mathematical problem-solving test booklet that tapped translation, integration, and planning skills as described by Mayer et al. (1991). One of the six items involving translation from the mathematical problem-solving test was: “Which number sentences is correct? Masao has 5 more marbles than Yasuko. a. Masao’s marbles + 5 = Yasuko’s marbles, b. Masao’s marbles + 5 = Yasuko’s marbles, c. Masao’s marbles + Yasuko’s marbles = 5, d. Masao’s marbles = 5.” One of the six items involving integration from the mathematical problem-solving test was: “Which numbers are needed to solve this problem? Sayuri had 500 yen for lunch. She bought a sandwich for 230 yen, an apple for 100 yen, and milk for 80 yen. How much money did she spend? 500, 230, 100, 90. One of the six items involving planning skills from the mathematical problem-solving test was: “Which operations should you carry out to solve this problem? If it costs 50 yen per hour to rent roller skates, what is the cost of using the skates from 1:00 p.m. to 3:00 p.m.? a. subtract, then multiply, b. subtract, then divide, c. add, then divide, d. multiply only.” The transfer test tapped translation and planning skills was in multiple choice formats, with four response alternatives. The transfer test tapped integration was the task in which students were instructed to circle all the numbers they thought of being needed. The pretests which included two kinds of tests used in the control group consisted of 8 ratio word problems similar to those used by the critical test, respectively. However, each pretest was easier than the critical test.

Procedure

Students were tested as part of their intact class in groups, at three times over a one-year period (in February, July, and next February). Prior to the first critical test in February, each student answered the pretest and engaged in one of three activities according to her or his group. Students in the training group had two forty-five-minutes training sessions about generating computer-based diagrams themselves. They tried to generate a diagram suitable for each ratio word problem using a computer in which only a center line was presented and they could manipulate it actively. Students in the extra lessons group had two forty-five-minutes extra lesson sessions. They had two classes about ratio word-problem solving. Students in the control group solved extra word problems yielding twice as many as students in the training group and in the extra lessons group. Students spent the same time on each activity. Students in the training group and in the control group did not receive feedback concerning the correct answers after each task was accomplished. One week after each activity session had finished, each student had the critical test in the first test period. Twenty-two students served in each group.
period. Students were allowed forty minutes to solve the critical test. In the second and third test periods, students took the same critical test as the first test period. After that, students also took the transfer test in the second and third test period. Students were allowed fifteen minutes to solve the transfer test.

Results

Pretest scores

Students solved 8 ratio word problems as a pretest prior to this study. The maximum score of the pretest was 16. Mean scores were 12.26 (SD=3.33) in the training group, 12.83 (SD=3.20) in the extra lessons group, and 13.67 (SD=3.24) in the control group. There was no difference of scores of the pretest among three groups, F (2,63) =1.04, ns, MSE=10.61.

Critical ratio word problem scores

The mean scores and standard deviations (SDs) of critical ratio word problems in each group are shown in Table 1. The ratio word problem scores are divided into two sub-scores, the scores of the second semantic type of word problems (second type) and the scores of the third semantic type of word problems (third type). The maximum score of each semantic type was 8. A 3 (group) x 2 (semantic type) x 3 (time of test) analysis of variance (ANOVA) was conducted on the mean scores of Table 1. There were significant main effects of group, F (2,63)=4.47, p <.05, MSE=13.44; semantic type, F (1,63)=5.74, p <.05, MSE=2.41; and time of test, F (2,126)=13.47, p <.01, MSE=1.32. The supplementary Tukey's HSD test (with p less than 5 %) revealed that students in the training group solved more word problems than those in the other two groups. There was no difference between the extra lessons group and the control group. Students solved the second semantic type of word problems more than the third semantic type of word problems. The supplementary Tukey's HSD test showed that students solved word problems more in the second test period than in the first test period. However, there was no difference of scores between the second test period and the final test period. There was also a significant interaction between semantic type and time of test, F (2,126)=3.13, p <.05, MSE=1.88. The supplementary Tukey's HSD test revealed that students solved more third semantic type of word problems from the first test period to the second test period.

Table 1

<table>
<thead>
<tr>
<th>Semantic Type</th>
<th>First Period</th>
<th>Second Period</th>
<th>Final Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Semantic</td>
<td>5.23</td>
<td>5.36</td>
<td>6.00</td>
</tr>
<tr>
<td>Third Semantic</td>
<td>3.64</td>
<td>4.62</td>
<td>4.58</td>
</tr>
<tr>
<td>Extra Lessons Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Semantic</td>
<td>4.35</td>
<td>4.41</td>
<td>4.59</td>
</tr>
<tr>
<td>Third Semantic</td>
<td>2.55</td>
<td>3.32</td>
<td>3.73</td>
</tr>
<tr>
<td>Control Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Semantic</td>
<td>4.18</td>
<td>4.36</td>
<td>4.41</td>
</tr>
<tr>
<td>Third Semantic</td>
<td>2.64</td>
<td>3.50</td>
<td>3.41</td>
</tr>
</tbody>
</table>

(Note). The maximum score of each cell was 8.

Next, the first test to second test gains and the second test to the final test gains for each group were analyzed to examine change processes of students' performance over a one-year period. Each gain is obtained from the data of Table 1 by subtracting the first test scores from the second test scores in corresponding semantic types of ratio word problems and by subtracting the second test scores from the final test scores in corresponding semantic types of ratio word problems. As shown in Table 1, the first test to second test gains seem to be greater than the second test to final test gains. A 3 (group) x 2 (semantic type) x 2 (interval) ANOVA was conducted on the results. There was only a significant main effect of interval, F (1,63)=4.09, p <.05, MSE=1.41. The performances improved in the first test to second test interval more than in the second test to final test interval.

Transfer test scores

The mean scores and SDs on the transfer test in each condition are shown in Table 2. The transfer test scores are divided into three sub-test scores, translation scores, integration scores, and planning scores. The maximum score of each test type of problems was 6. A 3 (group) x 3 (test type) x 2 (time of test) ANOVA was conducted on the mean scores of Table 2. There were significant main effects of test type, F (2,126)=3.16, p <.05, MSE=3.07; and time of test, F (1,63)=15.95, p <.01, MSE=0.57. The supplementary Tukey's HSD test revealed that students solved more translation type of problems than the other two types of problems. Students also solved more transfer test problems in the final period than in the second period. There were no
other significant main effects and interactions.

### Table 2
Mean scores (Ms) and standard deviations (SDs) for each condition on the transfer test

<table>
<thead>
<tr>
<th>Condition</th>
<th>Period</th>
<th>Second</th>
<th></th>
<th>Final</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Training Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation</td>
<td>4.91</td>
<td>1.27</td>
<td></td>
<td>5.41</td>
<td>1.03</td>
</tr>
<tr>
<td>Integration</td>
<td>4.86</td>
<td>0.81</td>
<td></td>
<td>5.00</td>
<td>0.80</td>
</tr>
<tr>
<td>Planning</td>
<td>4.14</td>
<td>1.51</td>
<td></td>
<td>4.82</td>
<td>1.40</td>
</tr>
<tr>
<td>Extra Lessons Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation</td>
<td>5.09</td>
<td>1.28</td>
<td></td>
<td>5.18</td>
<td>1.11</td>
</tr>
<tr>
<td>Integration</td>
<td>3.95</td>
<td>1.36</td>
<td></td>
<td>4.68</td>
<td>0.92</td>
</tr>
<tr>
<td>Planning</td>
<td>4.50</td>
<td>1.64</td>
<td></td>
<td>4.84</td>
<td>1.52</td>
</tr>
<tr>
<td>Control Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Translation</td>
<td>4.86</td>
<td>1.39</td>
<td></td>
<td>4.95</td>
<td>1.15</td>
</tr>
<tr>
<td>Integration</td>
<td>4.59</td>
<td>0.94</td>
<td></td>
<td>4.73</td>
<td>0.96</td>
</tr>
<tr>
<td>Planning</td>
<td>4.55</td>
<td>1.41</td>
<td></td>
<td>4.77</td>
<td>1.44</td>
</tr>
</tbody>
</table>

(Note): The maximum score of each cell was 8.

### Discussion

There are three main findings in the present study. First, students in the training group outperformed students both in the extra lessons group and the control group on solving ratio word problems. The result supports the hypothesis that students who have solved the ratio word problems by generating diagrams themselves with a computer outperform students who have not generated diagrams. The extra study lessons did not have an effect on students’ performances on ratio word problems.

Students in the training group generated diagrams of each semantic type of word problems using a computer before they solved the critical ratio word problems. Students were presented only a center line on a computer screen. They engaged in drawing a diagram that tapped each semantic structure of the word problems by operating the computer actively. They did not receive feedback about whether each diagram they had generated was correct or incorrect. Students in the control group also did not receive positive or negative feedback of their solutions of the extra ratio word problems. It was expected that, by engaging in a learning environment that encourages them to build a mental model of word problems using a computer, students in the training group would demonstrate more superior solutions of ratio word problems than students in the other two groups. The result of this study supports this prediction. Students in the training group solved ratio word problems as well as students who solved ratio word problems by generating a relational picture using a computer (Tajika et al., 1995, 1997).

The result also showed that students solved more second semantic type of ratio word problems than the third semantic type of ratio word problems. When some harder second semantic types of ratio word problems were used, self-generated diagrams were effective on the second semantic type of ratio word problems as well as the third semantic type of ratio word problems.

Second, the superiority of the scores of students in the training group was maintained over a one-year period. The supplementary Tukey's HSD test reveals that students in the training group outperformed students in the other two groups on the scores of the ratio word problems over a one-year period. It is clear that to generate computer-based diagrams has long-term effects. The obvious reason for the superiority of the scores of the students in the training group is that they have solved more word problems after the training at the first test period and that they have maintained this superiority over a one-year period. The results of the first test to second test gains and the second test to final test gains are consistent with the explanation. The results show that there are no differences among three groups on the first test to second test gains and the second test to final test gains. Students in the training group outperformed students in the other two groups only at the first test period.

Third, there were no differences among three groups on the transfer test. There are at least two possible explanations for the results of the transfer test. One explanation attributes the locus of the similarity of the scores among three groups to a format of the transfer test used in the present study. The transfer test used in this study was in a multiple-choice format. Each test item consists of four response alternatives. Students only selected the correct one from four items and made an easy decision about which item was correct.

The second explanation is that training activities using a computer result in effective performance on ratio word problem solutions but not effective transfer. Previous research on the transfer of problem solving emphasizes the similarity between conditions of original learning tasks and conditions of transfer tasks (Gick & Holyoak, 1983). Students who generated diagrams them-
selves using a computer could perform the similar tasks very well but they could not apply their mental models of ratio word-problem solving to solving new unrelated problems.

Finally, subjects in the training group did not have any feedback about whether the diagrams generated during the training were correct or not. Students in the training group received 8 center lines and generated 3.12 diagrams correctly during the first training session and 4.36 diagrams correctly during the second training session. If students in the training group had more training time for generating diagrams and receiving knowledge of their results, they would have more easily built their mental models for solving ratio word problems.

In summary, it was found that students who themselves generated diagrams tapped each semantic type of ratio word problems outperformed students who did not generate diagrams on ratio word problems over a one-year period. The computer-based diagrams are regarded as an effective cognitive strategy for helping students build mental models of ratio word problems. In contrast, it was also shown that there was no difference among the three groups on the transfer test. The present study points out the need for examining other kinds of transfer tests that have similar types of word problems as the ratio word problems.

References


