The Cognitive Load Cost of Constructing Representations When Learning to Solve Mathematical Word Problems

Brian D. Beitzel, Richard K. Staley, and Nelson F. DuBois
SUNY Oneonta

Poster presented at the annual meeting of the American Educational Research Association, New York, March 2008

Abstract: We investigated the use of external representations in helping college students solve probability word problems. Two types of multiple-event probability problems were presented. The specific-representations group was instructed in how to use tree and Venn diagrams; the general-representation group was trained to use a matrix to solve both types of probability problems. The control group was instructed only in the formulation of equations. Results indicated that the control group outperformed the specific-representations group on near transfer problems; there was no difference between the general-representation group and the other two groups on transfer problems. Results further indicated higher cognitive load in the specific-representations condition, supporting the contention that a split-attention effect was responsible for the lower performance of that condition.

> *The difference between understanding and not understanding is in the nature of the representation. . . . Good understanding involves achievement of a coherent representation* (Greeno, 1977, p. 44).

Theoretical Framework

Students can be taught how to create effective external representations for specific types of mathematics problems and these representations can help them to understand the nature of the problems and how to solve them. Several studies have demonstrated that teaching students to construct their own external representations leads to greater problem-solving performance. Lewis (1989) found that college students taught how to diagram mathematical word problems showed greater pre- to post-test gains than either another treatment group or a no-training control group. Stern, Aprea, and Ebner (2003) found that college students directed to create a linear graph based on a text passage answered questions about the passage better than students given the graph. Cox (1999) has suggested that constructing external representations has effects similar to those of self-explanation. Such external representations of mathematical word problems focus attention on critical attributes of the problem and their inter-relationships, thus helping students to develop a deeper understanding of the problem schemas underlying such problems.

Beitzel, Staley, and DuBois (2007) examined the effects of student-generated external representations within a worked example treatment designed to teach students how to solve two types of complex probability problems. The treatment transitioned from worked examples to problem solving using the backward fading procedure described by

Renkl and Atkinson (2003). Beitzel et al. postulated that external representations can vary in how directly their perceptual features illustrate the schematic relationships in a problem. The more salient these schematic relationships are in the external representation, the more thoroughly students should learn the problem schema and the better they should perform in solving novel problems based on that problem schema.

Beitzel et al. developed two treatments involving the construction of external representations. In the specific-representation group, participants were taught to construct an external representation (e.g., a Venn diagram) that more directly illustrated the schematic relationships in the problem than did the generic representation that students were taught to construct in the general-representation group. A third group (called the procedural group) served as the control group and was given instruction using the same mathematical procedure for solving the problems, except that the step of constructing an external representation was omitted. It was predicted that the worked-example instructional procedure with a specific representation would result in greater performance on a transfer task than either a worked example procedure with a general representation or no representation.

Results of the study showed that the group constructing a specific representation actually solved fewer posttest problems than the procedural group that created no representation as they learned to solve the problems. The performance of the general-representation group did not differ significantly from either the specific-representation or the procedural groups. This is a potentially important finding. Can a commonly endorsed heuristic for problem solving—"Create a diagram of the problem to help understand it"—actually interfere with learning to solve problems?

Purposes of This Experiment

One purpose of the current study is to replicate the Beitzel et al. (2007) study to determine the robustness of this finding. The other purpose of the study is to examine a possible explanation for these results. In our opinion the most plausible explanation for the results is the split attention effect. This occurs when a person has to integrate information from two or more sources in order to learn something with the result being higher cognitive load that interferes with learning (Sweller, van Merrienboer, & Paas, 1998). The present study incorporated a measure of cognitive load to determine if the specific-representation treatment actually interferes with learning by increasing the cognitive load experienced by participants.

Method

Participants

The participants were 99 undergraduate students at an Eastern college enrolled in introductory educational psychology courses. They participated in the study for extra course credit.

Materials

Instructional material was developed explaining how to solve two types of multiple-event probability problems: (a) joint probability of independent events (with replacement) and (b) total probability of non-mutually exclusive events. The materials deployed the worked example procedure with fading that was developed by Renkl, Atkinson, and Maier (2000). The instructional material consisted of a tutorial on basic probability concepts, an exposition of how to solve problems of each of the two problem types, and sixteen practice problems – eight word problems for each of the two types of probability problems. The instructional materials were identical across all treatments, except that the general-representation treatment included matrix representations of the problems (with instructions on how to construct a matrix); and the specific-representations treatment included tree diagram representations for the joint probability of independent events problems and Venn diagrams for the total probability of non-mutually exclusive events problems. All instructional materials verbally described how to solve each type of problem in a series of steps that involved (a) constructing an appropriate representation (for the two representation treatments only); (b) determining the number of favorable outcomes and total possible outcomes for each event, (c) solving for the desired probability.

The dependent measure was the number of correct solutions to 16 word problems representing near transfer (7 problems) and far transfer (9 problems). The 7 near-transfer problems included both types of problems presented in the instructional materials and differed from the latter in terms of surface features or story line. The 9 far-transfer problems were different in both surface and deep structure from the practice problems; however, the skills acquired during the instructional phase were sufficient to solve both the near- and far-transfer problems. For example, two of these far-transfer problems were combination problems requiring application of both joint probability and total probability equations. Participants were asked to show all of their work, but were not required to use any specific approach to solving the problems.

Cognitive load was measured by the Mental Demand scale of the NASA-TLX instrument (Hart & Staveland, 1988). Sweller et al. (1998) have reported that subjective rating scales (like the NASA-TLX) are sensitive to small variations in cognitive load and are more reliable and valid than other competing measures.

Design

The experimental design was a 3 (specific representations vs. general representation vs. procedural) × 2 (near-transfer problems vs. far-transfer problems) with repeated measures on the last factor. Pretest scores were used as a covariate to control for prior knowledge. All participants were randomly assigned to one of the three instructional treatment groups.

Procedure

Participants were tested in groups but completed their packets independently. The packets were structured such that the experiment was divided into four phases. In the first phase students were asked to complete a demographic questionnaire and a nine-item pretest of word problems involving basic probability concepts. The second phase consisted of studying the instructional material that represented the treatment to which each participant was assigned. In the third phase the NASA-TLX instrument was administered. Finally, participants completed the 16 near- and far-transfer problems.

Results

As stated previously the experimental design was a 3 × 2 design with 3 levels of the between-subjects factor (procedural vs. specific representations vs. general representation) and two levels of the within-subjects factor (near-transfer vs. far-transfer). An analysis of covariance was conducted, with prior knowledge of probability concepts (represented by pretest scores) as the covariate. There was a main effect for transfer, $F(1, 92) = 9.63$, $p < .01$, indicating that participants scored higher on the near-transfer problems (M = 63.9%, SE = 3%) than on the far-transfer problems (M = 44.9%, SE = 2.7%). There was also a main effect for condition, $F(2, 92) = 6.84$, $p < .01$.

Pairwise comparisons (using the Tukey-Kramer procedure) demonstrated that for the near-transfer problems, the procedural group (M = 74.2%, SE = 5.2%) outperformed the specific-representation group (M = 46.3%, SE = 5.1%), $p < .05$. There was no difference between the specific-representation group and the general-representation group (M = 71.3%, SE = 5.5%) or the procedural and general-representation groups, all *p*s > .05. There were also no differences among any of the groups on the far-transfer problems, all *p*s > .05.

A Fisher LSD approach (Levin, Serlin, & Seaman, 1994) was used to analyze the cognitive load data. A one-way analysis of covariance was conducted, with pretest scores as the covariate. There was a main effect for condition, $F(2, 90) = 4.47$, $p = .01$ (two participants did not complete the cognitive load measure). Pairwise comparisons demonstrated that participants in the specific-representation group (M = 61.28, SE = 3.81) rated the mental demand to be greater than participants in the general-representation group (M = 44.57, SE = 4.10). The ratings of mental demand in the procedural group (M = 52.12, SE = 3.98) did not differ statistically from the ratings in the other two groups.

We also qualitatively examined how students went about the task of solving these word problems. We found relatively low frequencies of usage of the representational strategies taught in the instructional materials in both the specific-representation group (M = 32.6%) and in the general-representation group (M = 41.7%).

In a separate qualitative analysis, we examined the algebraic equations participants used to solve the near- and far-transfer problems. After categorizing the equations as either appropriate or inappropriate (with scores of 1 or 0, respectively), a Fisher LSD approach was used to look for inter-group differences. The omnibus ANCOVA (with pretest scores as the covariate) was significant, $F(2, 92) = 8.56$, $p < .001$, so pairwise comparisons were conducted. Participants in the specific-representations group were found to have produced fewer correct algebraic equations (M = 37.2%, SE = 4.2%) than participants in either the procedural group (M = 61.2%, SE = 4.3%) or the general-representation group (M = 54.9%, SE = 4.6%). The number of correct equations in the procedural group did not differ from those of the general-representation group.

A follow-up analysis was also conducted to investigate whether there were any main effects of correct problem solutions collapsed across the two transfer problem types. Again using the Fisher LSD approach with an omnibus ANCOVA, $F(2, 92) = 6.31$, $p < .01$, pairwise comparisons indicated that participants in the specific-representations group arrived at fewer correct solutions (M = 41.3%, SE = 4.2%) than participants in either the procedural group (M = 61.5%, SE = 4.3%) or the general-representation group (M = 56.8%, SE = 4.5%). There were no differences in the number of correct solutions between the procedural and general-representation groups. Not surprisingly, this pattern mimics that found for the qualitative analysis of the number of equations constructed correctly in each condition.

Conclusions/Implications

The results of the present study replicate the previous findings of Beitzel et al. (2007). It is becoming clear that requiring students to construct external representations while learning to solve mathematical word problems can result in lower performance than applying the mathematical procedure without an external representation.

In addition, our cognitive load data suggest that the decreased performance in the specific-representation group may be due to the high mental demand of constructing the external representation. The specific-representation group reported higher levels of cognitive load than either the procedural group or the general-representation group. Since the only difference among the three experimental groups was in the representations they were taught to construct, the most logical conclusion from these data is that the mental burden of constructing the specific representations did not pay off in higher performance.

Even though there were twice as many practice problems in the present study as in the earlier study (Beitzel et al., 2007), the posttest performance of the participants in these three conditions was substantively unchanged. Therefore, additional practice did not appear to reduce the burden on working memory for the specific-representation condition.

The low usage of external representations in the specific-representation group may indicate that participants did not perceive that these representations were useful to them. In future research we would like to computerize this task to help ensure that participants utilize the strategy that they were trained to use; perhaps if external representations were used more consistently, the outcome would be more positive.

References

Beitzel, B. D., Staley, R. K., & DuBois, N. F. (2007, April). *Do representations help college students solve mathematical word problems?* Poster presented at the annual meeting of the American Educational Research Association, Chicago.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9,* 343-363.

Greeno, J. G. (1977). Process of understanding in problem solving. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory* (Vol. II, pp. 43-83). Hillsdale, NJ: Lawrence Erlbaum.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in psychology* (pp. 139-183). Oxford, England: North-Holland.

Levin, J. R., Serlin, R. C., & Seaman, M. A. (1994). A controlled, powerful multiple-comparison strategy for several situations. *Psychological Bulletin, 115*, 153-159.

Lewis, A. B. (1989). Training students to represent arithmetic word problems. *Journal of Educational Psychology, 81*, 521-531.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*, 15-22.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13*, 191-203.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Figure 1. Tree diagram for the specific-representation condition for the problem, "Suppose Dave has 2 blue shirts, 3 white shirts, and 1 yellow shirt hanging in his closet. He also has 4 red ties, 2 green ties and 2 yellow ties hanging there. What is the probability that if he picks a shirt and then a tie at random from the closet they will both be yellow?"

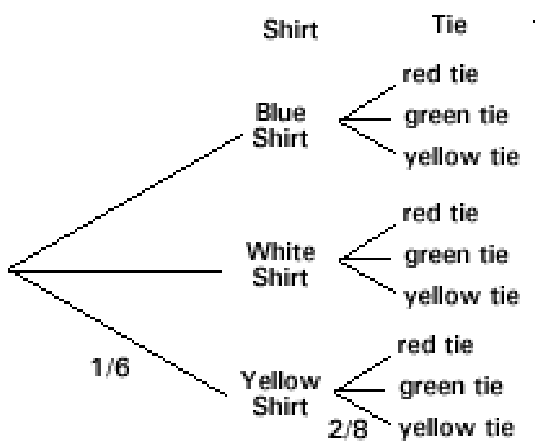

Figure 2. Matrix for the general-representation condition for the problem in Figure 1.

	Yellow Shirt - First	Yellow Tie - Second
Favorable outcomes	1	2
Sample Space	6	8