**When Best Intentions Go Awry: The Failures of Concrete Representations to Help Solve Probability 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, San Diego, April 2009

**Abstract:** We replicated the general results of two previous experiments showing that external representations hinder college students’ ability to solve probability word problems. In the present experiment, we converted our training and a portion of our practice materials to a multimedia format. The concrete-representations group was instructed in how to use Venn diagrams to solve probability problems involving non-mutually exclusive events; the control group was instructed only in the formulation of equations. Results indicated that the control group outperformed the concrete-representations group on transfer problems; additionally, there was a main effect for performance on transfer problems, favoring near-transfer problems over far-transfer problems. Contrary to our previous findings, in the present study cognitive load did not appear to be a factor in the lower performance of the concrete-representations condition.

**Theoretical Framework**

Recent research has called into question the efficacy of using concrete representations as an aid to solving mathematics problems. For example, Kaminski, Sloutsky, & Heckler (2008) found that college students who studied a mathematical concept using an abstract, symbolic representation outperformed their peers who learned the same concept using concrete representations (and even outperformed a third group who learned through both abstract and concrete representations). Koedinger, Alibali, & Nathan (2008) showed that concrete representations produce greater learning for simple mathematics problems but abstract representations are more efficacious when solving more complex mathematics problems.

Our own previous work (Beitzel, Staley, & DuBois, 2007, 2008) corroborates the general outcomes of these studies. We presented undergraduate students with training in either an abstract representation (i.e., an equation) or a concrete representation in the context of solving word problems involving two different types of probability: (a) joint probability of independent events and (b) total probability of non-mutually exclusive events. In two separate experiments, students who were trained to use the abstract representation completed posttest problems with a higher degree of accuracy than their counterparts who were trained to use the concrete representations.

In our second study (Beitzel et al., 2008) we hypothesized that the greater cognitive load (e.g., Van Merriënboer & Sweller, 2005) imposed by the use of the concrete representations might be responsible for the poorer performance of students using those representations. Our results confirmed this prediction; students using the concrete representations reported significantly higher levels of mental demand than students using only the abstract representations.
Purposes of This Experiment

One purpose of the current study was to replicate the Beitzel et al. (2007, 2008) studies to determine the robustness of the finding that concrete representations hinder student performance when solving probability problems. Another purpose of the study was to examine whether these results would hold when the instruction was offered in a multimedia format. A third purpose was to further investigate the possibility of an increase in cognitive load being responsible for any reduced accuracy when using concrete representations.

Method

Participants

The participants were 68 undergraduate students at an Eastern college enrolled in introductory educational psychology courses. They participated in the study for extra course credit.

Materials

We converted the instructional materials from the Beitzel et al. (2007, 2008) studies into a multimedia format. Appropriate visual information (e.g., an equation) was displayed on-screen while a narrator explained how to solve probability problems involving non-mutually exclusive events (see Figure 1 for a sample problem). The materials deployed the worked example procedure with fading that was developed by Renkl, Atkinson, and Maier (2000). The multimedia presentation included a tutorial on basic probability concepts, an exposition of how to solve probability problems involving non-mutually exclusive events, and eight word problems for practice. (Joint probabilities were not included in this study.) The instructional materials were identical across both treatments, except that the concrete-representations treatment included Venn diagrams (with explanations on their use) as external representations of the probability problems. All instructional materials described how to solve these problems in a series of steps that involved (a) constructing an appropriate representation (for the concrete-representation treatment only); (b) determining the number of favorable outcomes and total possible outcomes for each event; and (c) solving for the desired probability.

A 9-item pretest was developed to measure participants’ prior knowledge of basic probability concepts. Participants were asked to provide the solutions for simple probability problems involving a single event, such as, “When rolling a 6-sided die what is the probability that ‘2’ or ‘4’ will appear?”

The dependent measure was the number of correct solutions to 8 word problems representing near transfer (4 problems) and far transfer (4 problems). The 4 near-transfer problems included probability problems similar to the practice problems presented in the instructional materials and differed from the latter in terms of surface features or story line. The 4 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 all of these problems. Participants were asked to show all of their
work, but were not required to use any specific approach to solving the problems. All transfer problems had the same (or very similar) story lines as those in our earlier studies.

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 2 (procedural vs. concrete representations) × 2 (near-transfer vs. far-transfer problems) with repeated measures on the last factor. Pretest scores were used as a covariate to control for prior knowledge. Participants were randomly assigned to one of the two instructional treatment groups.

Procedure

Participants were tested in groups but completed their work independently. The presentation was structured such that the experiment was divided into four phases. In the first phase students were asked to complete a paper-and-pencil version of a brief demographic questionnaire and a nine-item pretest of word problems involving basic probability concepts. The second phase consisted of studying the multimedia instructional material that represented the treatment to which each participant was assigned. In the third phase a paper-based version of the NASA-TLX instrument was administered. Finally, participants completed the 8 near- and far-transfer problems.

Results

As stated previously the experimental design was a 2 × 2 design with two levels of the between-subjects factor (procedural vs. concrete representations) and two levels of the within-subjects factor (near-transfer vs. far-transfer problems). 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, $F(1, 65) = 7.26, p < .01$, indicating that participants scored higher on the near-transfer problems ($M = 3.06$, $SE = 0.15$) than on the far-transfer problems ($M = 1.96$, $SE = 0.14$). There was also a main effect for condition, $F(1, 65) = 11.37, p < .01$, indicating that participants in the procedural condition ($M = 2.92$, $SE = 0.17$) outperformed participants in the concrete-representations condition ($M = 2.11$, $SE = 0.17$).

Pairwise comparisons (using the Tukey-Kramer procedure) demonstrated that for the near-transfer problems, the procedural group ($M = 3.60$, $SE = 0.21$) outperformed the concrete-representation group ($M = 2.51$, $SE = 0.22$), $p < .01$. For the far-transfer problems, there was no difference between the procedural group ($M = 2.23$, $SE = 0.20$) and the concrete-representation group ($M = 1.70$, $SE = 0.21$), $p > .05$.

A one-way analysis of covariance (with pretest scores as the covariate) was used to analyze the cognitive load data. There was no difference in participant-rated mental demand between the procedural group ($M = 52.01$, $SE = 4.17$) and the concrete-representation group ($M = 61.53$, $SE = 4.11$), $F(1, 62) = 2.65, p > .05$. However, a
negative correlation was found between mental demand and posttest scores, $r = -.25$, $p = .04$, indicating that the more mental demand participants experienced, the fewer probability problems they were able to solve.

We also qualitatively examined how students chose to approach the task of solving these word problems. Unlike our previous studies, we found relatively high frequencies of usage of the representational strategies taught in the instructional materials in the concrete-representation group (M = 85.0%, SE = 5.40%).

**Conclusions/Implications**

The results of the present study replicate the previous findings of Beitzel et al. (2007, 2008). Our results from this experiment and our previous studies corroborate other findings (e.g., Kaminski et al., 2008; Koedinger et al., 2008), showing that requiring college students to construct external, concrete representations while learning to solve mathematical word problems often results in lower performance than requiring them to apply the abstract mathematical procedure without using a concrete representation. Even though compliance with strategy usage was quite high (M = 85%) in the concrete-representations group, they were still unable to solve the probability problems as proficiently as the procedural group.

Because the near-transfer problems closely resemble the practice problems, the near-transfer problems could be measuring a procedural execution ability without requiring an extensive conceptual understanding of the probability concepts underlying these problems. On the other hand, successful solution of the far-transfer problems relies more upon conceptual understanding of the principles explained in the instructional portion of the experiment; however, these principles must be applied in novel ways in order to arrive at a successful solution on the far-transfer problems. If this analysis is correct, the procedural group acquired greater proficiency than the concrete-representations group at applying the correct solution procedure for problems like those included in the training. In fact, given the significant difference in favor of the procedural treatment, the use of concrete representations (in the concrete-representations group) hindered development of the knowledge necessary to solve these problems.

Further, neither of these treatments was able to substantially increase conceptual understanding of the basic principles to the extent necessary to perform well on the far-transfer problems. Given this pattern of results, neither of these instructional treatments seems to have heavily engaged students in deep processing. Future research should address this challenge. For example, one promising instructional procedure to promote deeper processing is the use of self-explanations (Atkinson & Renkl, 2007). Further research might explore whether external representations can provide scaffolding for more effective self-explanations when solving probability word problems.
References


**Figure 1.** Sample multimedia screen showing worked example from the concrete-representation treatment

**Problem:** Out of 20 students in Ms. Smith's study hall, 8 are taking Chemistry and 14 are taking Spanish. If 4 students are in both classes what is the probability that a randomly chosen student is taking either Chemistry or Spanish?

1: Draw a Venn diagram to represent the possible outcomes.

2: Determine the probability of each event

   Event 1 is choosing a student taking Chemistry
   a. Sample space = 20
   b. Favorable outcomes = 8
   P(student taking Chemistry) = 8/20

   Event 2 is choosing a student taking Spanish
   a. Sample space = 20
   b. Favorable outcomes = 14
   P(student taking Spanish) = 14/20

3: Determine the probability of the overlap of the 2 non-mutually exclusive events
   P(student taking Chemistry and Spanish) = 4/20

4: Determine the total probability by adding the probabilities of the individual events and then subtract the overlap of the events

   P(student taking either Chemistry or Spanish) = 
   P(student taking chemistry) + P(student taking Spanish) - P(student taking Chemistry and Spanish)

   P(student taking either Chemistry or Spanish) = 8/20 + 14/20 - 4/20
   P(student taking either Chemistry or Spanish) = 18/20 = 9/10