

# Learning to Categorize Word Problems: Effects of Practice Schedules

Brian D. Gane & Richard Catrambone

Georgia Institute of Technology, School of Psychology, 654 Cherry St., Atlanta, GA 30332-0170

Email: bgane@gatech.edu, rc7@prism.gatech.edu

**Abstract:** Participants learned to categorize and solve probability word problems by studying worked examples and solving practice problems. The examples and problems were ordered according to either a blocked (sequential examples of the same problem category) or mixed (intermixed examples of different problem categories) schedule. Preliminary results suggest that when categorizing problems the blocked schedule might facilitate immediate performance, whereas the mixed schedule might facilitate delayed performance.

## Introduction

Solving word problems in science, technology, engineering, and mathematics (STEM) domains often requires determining the category of the word problem as a first step (e.g., selecting the relevant formulas or algorithm to solve the problem). Accurate categorization is difficult for novices because they have an incomplete base of conceptual categories and tend to concentrate on surface features (Chi, Glaser, & Rees, 1982).

Attempts to improve learners' ability to categorize word problems can be informed by psychological research on concept learning. Novices are essentially learning how to classify the exemplars that they view (i.e., the word problems) as instances of different categories (i.e., the problem types). This issue can be framed as "How do novices build a conceptual structure based on viewing example instances and what conditions during study might affect those processes?"

One important study condition is the order of examples. The *practice schedule* in which learners study or solve word problems can be varied such that multiple example instances from the same problem category are either presented sequentially (i.e., a *blocked* schedule), or distributed among other problem category examples (i.e., a *mixed* schedule). Research suggests that practice schedule can affect learners' conceptual structures (Hiew, 1977) and categorization accuracy (Gane & Catrambone, 2009; Kornell & Bjork, 2008).

These practice schedule conditions are relevant to instructional designers. Consider the task of creating a textbook presentation or homework assignment to teach multiple types of probability word problems. Should the textbook present worked examples according to a blocked or a mixed schedule? Likewise, should students complete homework problems that are arranged in a blocked or mixed schedule? These decisions become more difficult because practice schedules can differentially affect immediate and delayed performance (Schmidt & Bjork, 1992). This study examines how blocked and mixed practice schedules affect novices' immediate and delayed performance of categorizing and solving probability word problems.

## Method

### Participants

Undergraduates from the Georgia Institute of Technology who had not taken a college-level probability or statistics course participated. Data collection is ongoing; to date, data from 36 participants have been analyzed.

### Learning Materials

Participants studied four categories of probability word problems: permutation with replacement, permutation without replacement, combination with replacement, and combination without replacement. The learning materials consisted of four packets of word problems; each packet contained four worked examples followed by a *conventional problem* to solve. Thus, participants studied a total of 16 worked examples (four examples per category), and solved four conventional problems (one problem per category).

Worked examples had the word problem stem, a category label, the two features (*order*: permutation or combination, and *replacement*: with or without) that determine the category label, the category-appropriate formula, values inserted into the formula, calculations, and the final answer. Conventional problems had only the word problem stem.

### Design

The practice schedule manipulation had two levels: blocked and mixed. The same set of 16 worked examples was used in the blocked and mixed conditions, only the order differed. The blocked order had four worked examples from the same category presented consecutively; the mixed order never repeated a worked example from the same category consecutively. These two orders were operationalized by including four examples of

one category in a packet (blocked) or by including one example of each category in a packet (mixed). To reduce the influence of the *specific category ordering* used within the blocked and mixed schedules, we created two versions of each schedule with different category orders.

## Procedure

After studying the example packets, participants completed use, access, and transfer tests, in that order. In the *access test* participants chose the appropriate formula (but did not solve) the word problem. In the *use test* the formula was given; participants inserted the values and calculated the final answer. The transfer test was formatted like the conventional problems; we manipulated transfer distance by varying the number of probabilities used to calculate the final answer. The access, use, and transfer tests used a mixed schedule. Participants returned for Session 2 after 48 hours and completed another set of access, use, and transfer tests.

## Results

Preliminary results ( $n = 36$ ) are reported, but data collection is ongoing. Random assignment appeared to equalize domain knowledge and mathematical skills; there were no significant differences between conditions on either the probability pretest or math pretest.

All participants in the blocked conditions categorized all of the conventional problems correctly ( $M = 4.0$ ,  $SD = 0.0$ ), whereas participants in the mixed conditions did not ( $M = 3.6$ ,  $SD = 0.6$ ),  $F(1, 34) = 8.16$ ,  $p < .01$ . Practice schedule also affected categorization on the access tests. In Session 1, the blocked condition had higher categorization accuracy ( $M = 4.9$ ,  $SD = 1.1$ ) than the mixed condition ( $M = 4.5$ ,  $SD = 1.3$ ). In Session 2, however, the blocked condition ( $M = 3.9$ ,  $SD = 2.1$ ) had lower categorization accuracy than the mixed condition ( $M = 4.4$ ,  $SD = 1.5$ ). This reversal in means resulted in a significant interaction at the  $\alpha = .10$  level,  $F(1, 31) = 3.12$ ,  $p = .09$ . With a larger sample, we expect to find a significant interaction at a conventional  $\alpha$  level. The use and transfer test results did not show stable trends. We await a larger sample to determine if practice schedule affects categorization accuracy on the transfer test; previous research (Gane & Catrambone, 2009) suggests a small effect.

## Discussion

These preliminary results suggest that practice schedules might affect concept learning with word problems in STEM domains. Further, they suggest that this effect might change as a function of the retention interval. These results extend the general phenomenon of the *contextual interference effect* (Battig, 1979), to concept learning processes in a cognitive skill domain.

These results suggest that ideas from concept learning literature can inform instruction design decisions. For instance, to facilitate categorization while studying, a blocked schedule might be used. However, to improve delayed performance, a mixed schedule might be more effective. An interesting question is whether this type of ordering is practical to implement in a classroom environment, and how students' expectations of how lessons typically order materials (e.g., simple to complex, blocked first, then random, etc.) might affect their learning from different schedules.

We believe future work should focus on the processing differences between individuals in each condition. It would be helpful to collect qualitative data regarding whether participants in each condition are explicitly comparing worked examples to one another. Another interesting question is how individual differences in metacognitive processing might moderate practice schedule effects. For instance, individuals that effectively monitor their understanding and control their allocation of study effort might engage in the same type of processing regardless of the type of practice schedule.

## References

- Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (p. 23-44). Hillsdale, NJ: Erlbaum.
- Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence (Vol. 1)*. Hillsdale, NJ: Erlbaum.
- Gane, B. D., & Catrambone, R. (2009, April). *Practice schedules in cognitive skill acquisition: Effects of example order on categorization and problem solving*. Paper presented at the American Educational Research Association annual meeting, San Diego, CA (AERA '09).
- Hiew, C. C. (1977). Sequence effects in rule learning and conceptual generalization. *American Journal of Psychology*, 90, 207-218.
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science*, 19, 585-592.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3, 207-217.