SEEKING RELEVANCE: SHOULD A DIAGRAM BE NOTICED WHEN SOLVING A MATHEMATICS PROBLEM?

Dr. Brian D. Beitzel*

**Brian D. Beitzel, Professor, Department of Educational Psychology and Counseling, SUNY Oneonta, Oneonta, NY, USA, E-Mail: brian.beitzel@oneonta.edu*

ABSTRACT

Abstract—Diagrams are commonly used (and encouraged) for assisting learners with solutions to mathematics problems. Unfortunately, the empirical literature on postsecondary student use of diagrams is not uniformly promising. This study employed eye-tracking equipment to explore the extent of undergraduate students' use of diagrams when solving probability word problems. Results showed that students spent relatively little time looking at accompanying diagrams while solving word problems, and some portions of the diagram were noticed much more than others. Additionally, students who were trained to use diagrams achieved *lower* problem-solving performance than their peers who were trained to use equations alone. These findings suggest that educators should be cautious about encouraging students to draw diagrams as a solution strategy.

Keywords—diagrams, eye-tracking, mathematics representations

INTRODUCTION

Mathematics educators frequently employ diagrams as instructional aids to help students solve many kinds of mathematics problems. Indeed, diagrams are so ubiquitous that their utility is seldom even questioned.

The use of diagrams is explicitly endorsed by several professional groups in mathematics. For example, the Conference Board of the Mathematical Sciences (2012) emphasized that mathematics teachers should model "strategic use of appropriate tools" (p. 1) — and one of the "appropriate tools" listed is diagrams.

The Committee on the Undergraduate Program in Mathematics (2015, p. 74) listed six "core competencies for quantitative reasoning" (which were adopted from other sources); one of these "core competencies" is representation — and diagrams are listed as an example. No cautions are provided that diagrams may possibly inhibit student success when reasoning quantitatively.

Other groups that have encouraged the use of diagrams in mathematics include the National Mathematics Advisory Panel (2008) and the National Council of Teachers of Mathematics (2000). Thus, there is no shortage of professional support for utilizing diagrams in mathematics instruction.

The above-referenced endorsements indicate that diagrams are presumed to be beneficial to student learning. Yet surprisingly little scholarly work (e.g., Clinton, Alibali, & Nathan, 2013; Cooper, Sidney, & Alibali, 2018) has been conducted to confirm these assumed advantages — and not all of the findings have been positive.

In a series of studies, my colleagues and I have documented a positive impact of diagram training on undergraduate students' solutions to conditional-probability word problems (Beitzel, Gonyea, & Staley, 2013, 2014; Beitzel & Staley, 2011, 2015). However, it has been empirically established that conditional-probability problems are more difficult for undergraduates than total- or joint-probability problems (Sedlmeier & Gigerenzer, 2001). For difficult mental tasks, it can be helpful to have the assistance of a device such as a diagram to "offload" some of the work onto an external representation. But what happens to the effectiveness of the diagram when it is adopted for word problems that are easier to solve?

© ICBTS Copyright by Author(s)

Our empirical work has demonstrated that diagrams do not uniformly improve problem-solving performance. In another series of experiments (Beitzel, 2018; Beitzel, Boss, & Gonyea, 2015; Beitzel & Staley, 2016; Beitzel, Staley, & DuBois, 2011a, 2011b; Beitzel, Staley, Holmes, & Snow, 2017), training undergraduates to use diagrams to solve total- and joint-probability problems did not result in improved performance. In none of these experiments did diagrams improve problem-solving performance. And worse, in most of them, performance *declined* when students had been trained to use diagrams! This is surely a devastating pattern of results for advocates of diagrams in mathematics.

Given these findings, the research question for the present study is, how are students visually interacting with Venn diagrams when solving total-probability problems? Since diagrams have been demonstrated to decrease problem-solving accuracy for this type of problem, what portion(s) of the diagram are students looking at as they solve these problems? Or are they not using the diagram much at all? The eye-tracking equipment employed for this study assisted us in addressing these questions.

METHODOLOGY

2.1. Participants

The participants were 58 undergraduate students (41 women, 17 men) at an eastern U.S. college. The largest proportion of participants was freshmen (67%), followed by sophomores (21%), juniors (10%), and seniors (2%). The mean self-reported GPA was 3.14, and most participants (72%) reported having taken no more than one mathematics class in college.

2.2. Materials

2.2.1. Pretest

A nine-item pretest was used to measure participants' prior knowledge of probability concepts varying in difficulty from simple probability problems (e.g., "When a coin is flipped what is the probability that a 'head' will appear?") to more complex probability problems. The total score from the pretest was used as a covariate to control for prior knowledge when analyzing the posttest data. The reliability of the pretest data is an ω_t (omega total; McDonald, 1999) value of .72.

2.2.2. Instructional materials

Instructional materials were developed that explained some general probability concepts and then explained total probability, with examples. Two sample problems dealing with total probability were introduced and their solutions were identified and explained. Participants in the diagram condition received instruction on how to use a Venn diagram as an aid in constructing an equation to solve total-probability problems; participants in the no-diagram condition received instruction on how to generate an equation, without any reference to a diagram.

A series of practice problems followed the sample problems; the practice problems were presented in a faded worked-example format (Renkl & Atkinson, 2003), with the first problem fully worked out, the second problem mostly worked out, etc., until the final problem was left for participants to solve independently with only the steps labeled.

2.2.3. Equipment

The entire experiment was presented on a computer so that eye-tracking equipment could monitor the gaze patterns as participants encountered the instructional materials. The computer monitor was a 22in display at a resolution of 1920 × 1080 pixels. The eye-tracking hardware was the 60Hz binocular FOVIO device (2014); the software used to collect and analyze the eye-tracking data was EyeWorks™ version 3.21 (2017).

2.2.4 Posttest

A 6-item posttest was constructed to measure participants' ability to solve total-probability word problems without assistance. Scorers were blind to condition and assigned scores based on conceptual accuracy of the solutions, ignoring errors in rounding or fraction simplification. The reliability of the posttest data is □□ = .94.

2.3. Design

This experiment used a two-group (diagram vs. no-diagram), between-subjects design.

2.4. Procedure

Participants completed the experiment individually, in a laboratory setting with an experimenter present. All tasks were delivered on a computer. After a brief calibration with the eye-tracking equipment, a demographic survey was administered; then the pretest was administered, followed by the instructional materials that included two sample problems and five practice problems. The final portion was the posttest. At several points within the experiment, participants were asked to rate how much mental effort the tasks demanded (but those data are not reported here.)

RESULTS

3.1. Problem Solving

There were no between-group differences on the pretest, $F(1,56) = 1.95$, $p = .17$, $MSE = 3.453$, indicating that participants in these two groups did not differ in their prior knowledge of probability.

An analysis of covariance (ANCOVA) was used to analyze the problem-solving data, with pretest scores (representing prior knowledge of probability) as the covariate. There was a main effect of condition, $F(1,55) = 17.83$, $p < .01$, $MSE = 3.872$, $d = 1.13$, indicating that participants in the no-diagram condition ($M = 4.51$, $SE = 0.39$) outperformed their counterparts in the diagram condition ($M = 2.28$, $SE = 0.35$).

3.2. Eye-Tracking

To answer the major research questions for the present study, this analysis will focus on the first two practice problems that were presented as part of the instructional materials — and only on data from the diagram condition. Although it would be desirable to examine participants' diagrammatic gazes during the posttest, the diagrams on that instrument were generated by the participants and thus could not be examined in the aggregate using standard eye-tracking techniques.

Table 1
Mean Percentage of Gaze Time in Each Major Region of the Diagram for Selected Practice Problem Steps

Problem	Step	Region I	Region II	Region III
1	1	2.2%	1.9%	1.5%
1	2	0.9%	2.5%*	1.3%
2	2	1.0%*	2.5%	2.8%

*$p < .05$ compared to the two other regions.

3.2.1. Practice Problem #1, Step 1

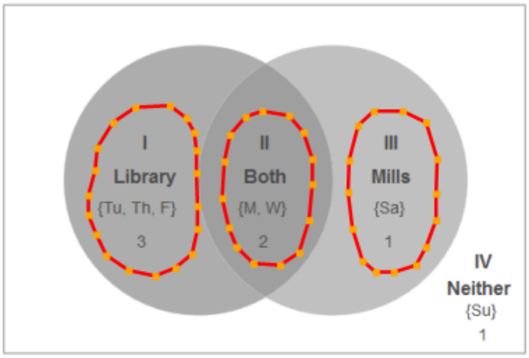


Figure 1
Areas of interest for practice problem #1, step 1

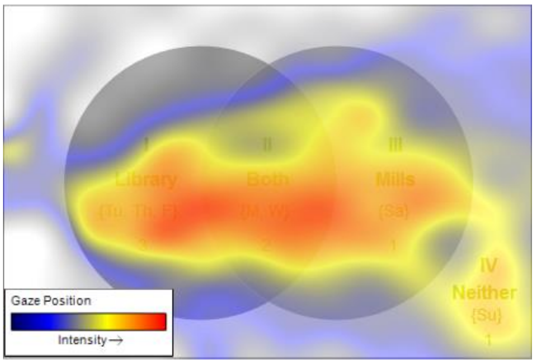


Figure 2
Length of participants' gaze at the diagram in practice problem #1, step 1

Practice Problem #1 presented a word problem along with a completed Venn diagram and an explanation of how the values within the diagram had been derived (Step 1 in solving the problem). Figure 1 shows the diagram as well as the researcher-constructed boundaries for calculating the amount of time participants spent examining each region of the diagram. Figure 2 shows the intensity of participants' gaze as they visually interacted with the diagram.

There was no difference in gaze time among the three regions defined in Figure 1, $F(2,90) = 1.11$, $p = .33$,

$MSE = 3.029$ (see Table 1 for the average gaze time per region, as a percentage of gaze time recorded on each screen). Combining the three regions together, participants spent only 6% of their gaze time in these areas while they remained on this screen. The remainder of their gaze time was spent on the other portions of the screen (e.g., the problem text, instructions, etc.; see the Appendix). In other words, participants did not spend very much time looking at the diagram.

3.2.2. Practice Problem #1, Step 2

Figure 3
Areas of interest for practice problem #1, step 2

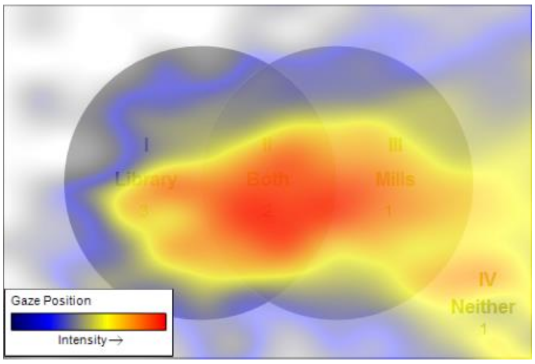


Figure 4
Length of participants' gaze at the diagram in practice problem #1, step 2

On the subsequent screen, the instructions for Step 2 prompted participants to generate and solve an equation for this problem using the values derived from Step 1 (see Figures 3 and 4).

There was a significant difference in gaze time among the three regions defined in Figure 3, $F(2,90) = 11.44$, $p < .01$, $MSE = 1.849$. Participants spent more of their time in Region II (the overlap region) than in Region I ($p < .01$) or Region III ($p < .01$).

Across all three regions combined, participants spent only 5% of their gaze time on this screen in these areas. So again, not much time was spent looking at the diagram, relative to other portions of the screen.

3.2.3. Practice Problem #2, Step 1

Practice Problem #2 had a similar structure, except that for Step 1 (not shown), a blank diagram with labels was provided, and participants were asked to fill in the values. Because there were no values displayed for participants to visually examine in the diagram, Step 1 will not be further examined in this paper.

3.2.4. Practice Problem #2, Step 2

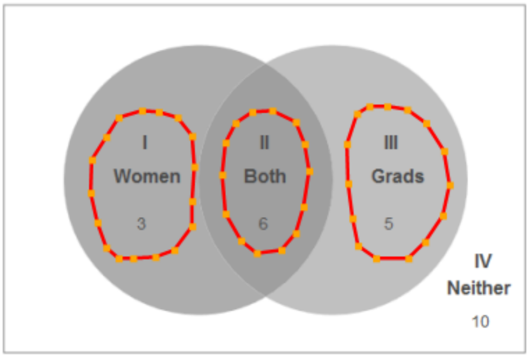


Figure 5
Areas of interest for practice problem #2, step 2

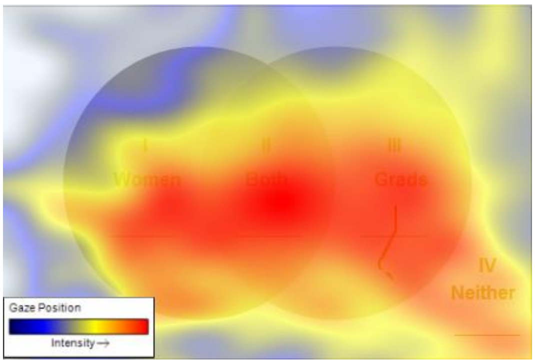


Figure 6
Length of participants' gaze at the diagram in practice problem #2, step 2

The correct values were provided in Step 2 (see Figures 5 and 6) in order to avoid any issues with miscalculations from Step 1 on the previous screen.

There was a significant difference in gaze time among the three regions defined in Figure 5, $F(2,90) = 9.08$, $p < .01$, $MSE = 3.047$. Participants spent less time in Region I than in Region II ($p < .01$) or Region III ($p < .01$); however, there was no difference between the time spent in Region II and the time spent in Region III, $p = .54$.

Once again, participants spent a minimal proportion of their time (6%) examining these three regions, relative to the time spent looking at the other portions of the screen.

DISCUSSION

The introduction of the Venn diagram as a solution strategy for total-probability word problems caused inferior performance in the diagram group, relative to the no-diagram group in which participants were trained to use equations alone. This deterioration of problem-solving performance in the diagram condition was anticipated from prior studies (Beitzel, 2018; Beitzel et al., 2015; Beitzel & Staley, 2016; Beitzel et al., 2011a, 2011b; Beitzel et al., 2017), yet reiterates the message that diagrams cannot be recommended universally to undergraduates seeking to improve their problem-solving skills.

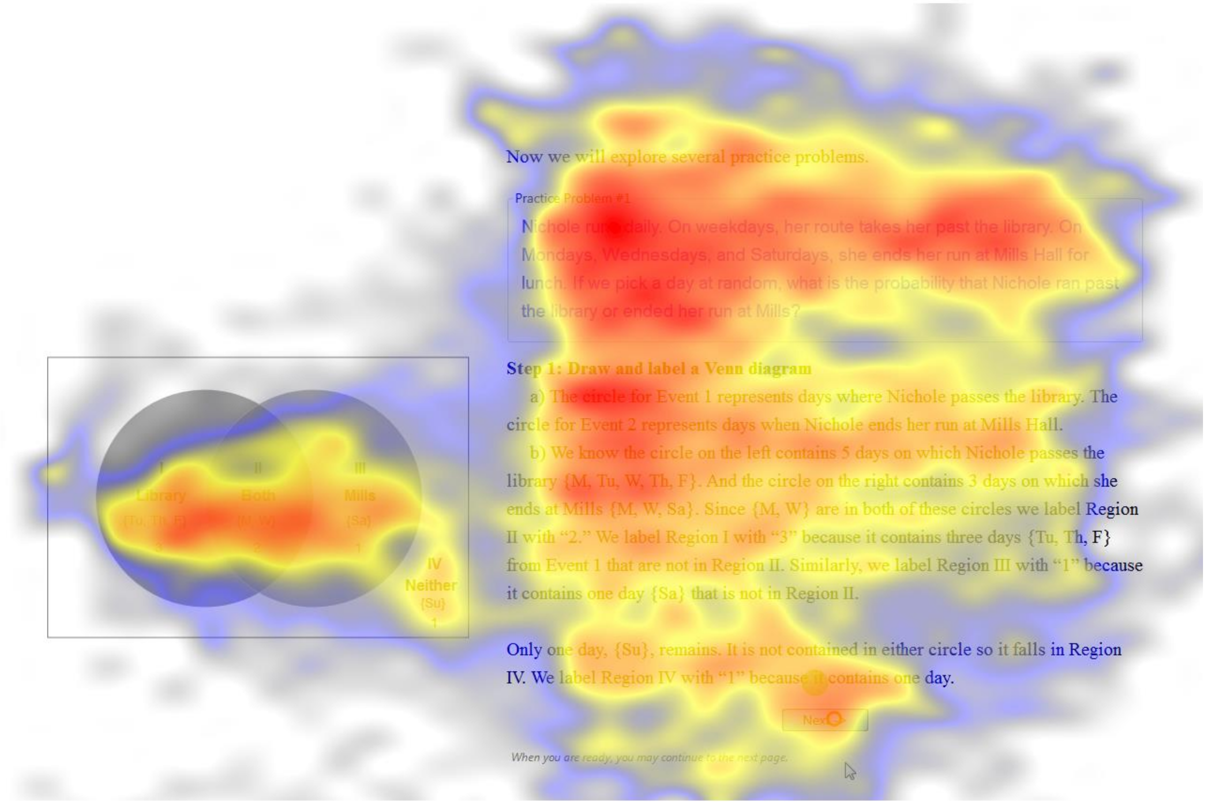
Regarding the central research questions for this study, the eye-tracking data demonstrate that participants are not spending a great deal of time visually interacting with the diagrams during the training phase of the experiment. In the three diagrams examined here, participants spent only 6%, 5%, and 6%, respectively, of their screen time gazing at the diagrams — and *still* their ultimate problem-solving performance was not a match for their peers who did not experience these diagrams.

We also learn from these data that participants are focused on certain portions of the diagrams. For example, there are portions of Figure 6 that received very few gazes at all. The data in Table 1 are also instructive here, documenting that participants are selective about how much time they spend examining each of the three major regions of the diagram. In Practice Problem #1, Step 2, the overlap region received more attention than the other two regions; but in Practice Problem #2, Step 2, the far-left region received less attention than the other two regions.

Should a diagram be noticed when solving a mathematics problem? If the problem-solvers are undergraduates and the mathematics problem is a total-probability word problem, the clear answer is *no*. Even the minimal noticing documented in the present study appears to be sufficient to deteriorate performance. Thus, at least for this particular situation, the notion that diagrams contribute to problem-solving success should be reconsidered.

APPENDIX

SAMPLE PROBLEM SCREEN



ACKNOWLEDGEMENT

This research was supported in part by the Spencer Foundation.

REFERENCES

Beitzel, B. D. (2018, January). *The risks of using diagrams in mathematics problem solving: Common sense is not sufficient.* Paper presented at the annual meeting of the Hawaii International Conference on Education, Honolulu, HI.

Beitzel, B. D., Boss, R. C., & Gonyea, N. E. (2015, April). *Diagrams may not be your friend: Empirically discriminating their problem-solving efficacy.* Poster presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Beitzel, B. D., Gonyea, N. E., & Staley, R. K. (2013, April). *The more the merrier? An investigation of multiple external representations on problem-solving performance.* Poster presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Beitzel, B. D., Gonyea, N. E., & Staley, R. K. (2014). *Illusions of using a diagram: An investigation of two types of mathematical representations.* Poster presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Beitzel, B. D., & Staley, R. K. (2011, April). *Sometimes they are effective: Using external representations to help solve probability word problems.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Beitzel, B. D., & Staley, R. K. (2015). The efficacy of using diagrams when solving probability word problems in college. *The Journal of Experimental Education, 83*, 130-145.

Beitzel, B. D., & Staley, R. K. (2016). *Paging the mathematician general: Empirical warnings for diagram use.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Beitzel, B. D., Staley, R. K., & DuBois, N. F. (2011a). The (in)effectiveness of visual representations as an aid to solving probability word problems. *Effective Education, 3*, 11-22.

Beitzel, B. D., Staley, R. K., & DuBois, N. F. (2011b). When best intentions go awry: The failures of concrete representations to help solve probability word problems. *Educational Research Quarterly, 34*(3), 3-14.

Beitzel, B. D., Staley, R. K., Holmes, M. S., & Snow, A. L. (2017, April). *The eyes have it: Examining a physiological measure of cognitive load.* Poster presented at the annual meeting of the American Educational Research Association, San Antonio, TX.

Clinton, V., Alibali, M. W., & Nathan, M. J. (2013). The effects of diagrams and questioning-while-reading on learning from a statistics lesson. In M. V. Martinez & A. Castro Superfine (Eds.), *Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 341-348). Chicago, IL: University of Illinois at Chicago.

Committee on the Undergraduate Program in Mathematics. (2015). *Curriculum guide to majors in the mathematical sciences.* Washington, DC: The Mathematical Association of America.

Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II.* Providence, Rhode Island: American Mathematical Society.

Cooper, J. L., Sidney, P. G., & Alibali, M. W. (2018). Who benefits from diagrams and illustrations in math problems? Ability and attitudes matter. *Applied Cognitive Psychology, 32*, 24-38.

EyeWorks (Version 3.21) [Computer software]. San Diego, CA: EyeTracking, Inc.

FOVIO [Apparatus].(2014).Canberra, Australia: Seeing Machines.

McDonald, R. P. (1999). *Test Theory: A Unified Treatment.* Mahwah, NJ: Lawrence Erlbaum Associates.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National*

Mathematics Advisory Panel. Washington, DC: U.S. Department of Education.

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.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*, 380-400.