Effects of Problem Difficulty and Student Expertise on the Utility of Provided Diagrams in Probability Problem Solving

Eliza J. Bobek (eb357@columbia.edu)
Teachers College, Columbia University
525 W. 125th Street, New York, NY 10027 USA

James E. Corter (jec34@columbia.edu)
Teachers College, Columbia University
525 W. 125th Street, New York, NY 10027 USA

Abstract

This study investigated the use of schema-specific diagrams in probability problem solving. Graduate students enrolled in an introductory probability and statistics course solved four probability problems, with and without instructor-provided diagram hints. Participants' solutions were examined and coded for correctness, use of provided diagrams, and use of student-generated external visual representations. Results show that provided diagram hints helped low-ability students on all but the most difficult problem, while high-ability students were aided by diagrams on the most difficult problem. Implications for the use of diagrams in the development of problem solving proficiency are discussed.

Keywords: probability problem solving; diagrams; visual representations; trees; Venn diagrams; contingency tables

Introduction

Learning probability concepts and solving probability problems can be challenging for students (Garfield & Ahlgren, 1988; Konold, 1989; O'Connell, 1999). Successful probability problem solving requires that students understand complex concepts, and also that they master how and when to use specific formulas and procedures that are particular to this domain. External visual representations are commonly used in many types of mathematical problem solving, including probability problem solving (PPS). These representations may promote solution success and student comprehension in several ways: by making abstract concepts visible and manipulable, or by organizing the subparts of a problem in a format that can be tied to solution procedures. Using structured visual representations may help problem solvers invent, retrieve, or apply formal solution schemas, increasing the rate of successful solution.

Why are external visual representations useful in the problem solving process?

Tversky (2001) suggests that external visual representations can serve many purposes, including recording information, relieving working memory, communicating to others, and facilitating inference and discovery. Drawing a diagram can reveal implicit information that is not readily available in a written description and make some pieces of information more explicit. It can also give a problem solver unique insight into the problem's structure or schema. Van Essen & Hamaker (1990) report that the "construction of a drawing might increase the chance that the problem situation is recognized and that the correct schemata is identified among other competing schemata" (p. 311). Other researchers (e.g., Larkin & Simon, 1987; Koedinger & Anderson, 1990) propose that problems solvers possess stronger schemas for diagrams than for words that contain the same information. These diagrammatic schemas may thus have an advantage in problem solving over verbal solution methods. Diagrams may also contribute to a fuller understanding through the use of multiple representations when they are used in conjunction with mental images. Several researchers have suggested that multiple representations may lead to increased "depth" of processing (e.g. Logie & Baddeley, 1990; Mayer & Gallini, 1990; Ainsworth, 2007).

However, a diagram may not help every student at every stage of expertise. Lowrie & Kay (2001) suggest that for students who already have schemas in long-term memory, using a diagram may not be particularly helpful; instead these students are able to generate their own representations. Problem difficulty may also play a role in when students choose to create external visual representations. In their study, Lowrie & Kay (2001) found that elementary age students tended to create external visual representations for especially difficult problems. Since problem difficulty is relative to student ability level, this suggests that individual differences may play a role in diagram use.

Furthermore, the types of external visualizations used in scientific reasoning and problem solving may differ, and be used in different ways (e.g., Edens & Potter, 2008; Van Meter & Garner, 2005). Some representations are relatively abstract, and are commonly used to represent schematic or abstract aspects of the problem. These diagrammatic or schematic representations include tree diagrams, and Venn diagrams used to represent part-whole relationships. These general-purpose diagrams should be distinguished from problem-specific representations that include concrete components of the problem itself.

Other external visual representations may be iconic rather than schematic, including pictures that represent the problem context and sketches that display and/or reorganize the information presented in the problem. The type of external visual representations used may be influenced by the problem goals and other specific problems features.

Use of external visual representations in probability problem solving (PPS)

In the specific area of probability problem solving, Zahner and Corter (2010) provide evidence that particular types of external visual representations are more often used in particular stages in the problem solving process. Some representations, such as reorganizing the problem information or drawing sketches, are more often used early in the solution process when a problem solver is trying to build a mental model of the problem text. Schematic diagrams, such as an outcome tree, can facilitate abstraction of the text, building a mathematical representation of the problem, and planning of the solution process. In general, different external visual representations may come into play in different stages of the problem solving process because their specific structures afford particular functions.

Quite often, the external inscriptions created by students solving probability problems provide evidence of spontaneously created diagrams and other visual representations. Three types of external visual representations often depicted in textbooks and used in PPS are Venn diagrams, outcome trees, and contingency tables. The structure of these diagrams allows for elements of the problem to be represented externally in an organized way. In an outcome tree, for example, each branch can be used to represent individual probabilities of events. The combination of events can be calculated by multiplying the values assigned to each branch. A Venn diagram is used to organize problem information, typically with overlapping circles to show the union and intersection of events. A contingency table is a matrix structure that shows the frequencies or probabilities of events, to show combinations of events.

Previous studies have pointed to the use of specific visual representations appropriate to specific problem types. Russell (2000) found that the use of outcome trees was correlated with improved performance specifically on conditional probability problems. He also found that students in a probability course used outcome trees more often than contingency tables or Venn diagrams. Interestingly, instructing students to draw an outcome tree did not affect performance. However, students who did draw outcome trees outperformed students who did not.

Zahner & Corter (2010) found that particular external visual representations were associated with specific probability topics, and that particular representations were associated with higher rates of solution correctness for some problem types. Their study suggests that using correct external visual representations may generally be facilitative in problem solving, but this facilitation is difficult to detect because students must first choose the correct diagrammatic representation. Other researchers too have noted this challenge (Novick 1990; 2001, Novick & Hmelo, 1994).

The evidence above shows that spontaneous student use of diagrams is associated with higher rates of solution success. But the causal direction is not entirely clear. Evidence that asking students to create diagrams may have a facilitative effect is provided by Schwartz & Martin (2004), who found that student understanding of statistical concepts was influenced by experimenter-prompted “invention” activities (activities in which students created representations.)

The Present Study

While probability problems can be solved using formulas, we hypothesize that using external visuals may help students overcome comprehension difficulties and may lead to greater problem solving success. From an educational standpoint, we would like to better understand the positive correlation between use of diagrams and problem-solving success. The question is whether drawing correct diagrams leads to better understanding, which facilitates problem solution, or if better understanding enables both the creation of correct diagram and problem solving success. Previous work in this area has shown that a major barrier to success in PPS is problem comprehension and representation.

Choosing an appropriate representation is a significant factor in problem solving success, and should be viewed as a skill unto itself (Novick & Hmelo, 1994; Edens & Potter, 2007; Uesaka, Manalo, & Ichikawa, 2007). We hypothesize that cuing or providing diagram “hints” appropriate to the problem type may aid students in the problem comprehension phase because the diagram provides a structure upon which problem components can be mapped. They may also help students to recognize the structure of the problem.

In particular, we are interested in the use of schema-specific external visual representations (distinguished from other representations, such as drawing a picture and reorganizing the given information), because we believe they have a special role in PPS. Thus, three common diagrams used in PPS were selected for use in the study: contingency tables, outcome trees, and Venn diagrams. The study attempted to investigate the role of correct external visual representations by providing appropriate but “generic” diagrams (“diagram hints”) directly to students. Each problem was chosen as a prototype of a specific problem topic/type and matched to diagram hints that are commonly used in probability curricula. The problems in this study were typical of those presented in the curricula and students had prior exposure to using specific diagrams for specific problem topics. For example, problem 4 is a conditional probability problem, for which outcome trees are an appropriate representation.

We have three main research questions:

1. Do instructor-provided diagram hints (e.g. a correct but unlabeled Venn diagram) increase the probability of problem solving success on specific problems?

2. Do students actually use the diagram "hints," or are they ignored, or are different external visual representations spontaneously created by students?

3. Does student ability mediate the effectiveness of the use of diagrams as a solution strategy?

Method

Participants. Participants were 129 students recruited from introductory probability and statistics classes at Teachers College, Columbia University. Participants were graduate students in education and social sciences, with a broad range of experience in mathematics.

Materials. Each participant was given four probability problems to solve as a problem set (Figure 1). Half of the participants received blank diagrams for problems 1 and 3 (Version A, n=64); the other half of the participants received diagrams for problems 2 and 4 (Version B, n=65). The diagrams were an outcome tree for problem 1, a contingency table for problem 2, a Venn diagram for problem 3, and an outcome tree for problem 4.

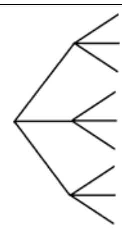
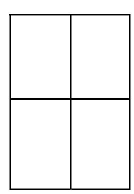
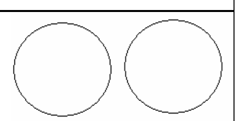
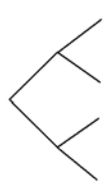
1. A bag of candy contains a mix of jelly beans that includes lime, cherry, and orange flavors. Five jelly beans are cherry, three are orange, and two are lime. Two jelly beans are randomly selected from the bag. What is the probability that the two selected jelly beans include exactly one cherry and one orange?	
2. A survey is conducted on attitudes towards handgun control. 42% of respondents to the survey are urban residents and the rest are rural residents. The results show that 33% of survey respondents are urban residents who support strict handgun controls, while 30% of survey respondents are rural residents who support strict handgun controls. What is the probability that a randomly chosen respondent is a rural resident, given that they support strict handgun control?	
3. A and B are mutually exclusive events. The probability of event A is .3, and the probability of event B is .25. What is $P(A^c \cap B^c)$*?*	
4. The weather forecast says that the probability of having good weather tomorrow is .60. If the weather is good, the probability that Eva will go out biking is .80. If it is not good weather, the probability is .20 that she will go out biking. What is the probability that Eva goes out biking tomorrow?	

Figure 1: Probability problems and provided diagrams

Procedure. Participants were allowed to use their class notes to solve the problems, which is standard practice in the course for completing homework assignments and exams. They were given approximately 20 minutes to solve the problems. This time limit was based on a pilot study and was imposed to discourage participants from either quickly scanning the problems or taking an inordinate amount of time.

Coding of participant solutions. Written solutions were coded for several features. First, we coded whether or not the participant gave a correct answer to the problem. Problems were given a score of "0" if incorrect, and a score of "1" if correct. We also totaled student scores for the four problems. Second, we coded whether or not the participant used the instructor-provided diagram hint. Next, we coded for any other type of external visual representation created by the student. The following categories, developed through previous research in our lab, were used to code for the different types of external visual representations: pictures, outcome listings, outcome trees, contingency tables, Venn diagrams, reorganization of given information in the problem, and novel schematic representations (Corter & Zahner, 2007; Zahner & Corter, 2010).

Results

An initial analysis found that over 80% of the participants made use of the instructor-provided diagram for each of the four problems. Analyzing student responses found that the four problems varied in difficulty. Comparing student performance on Version A and Version B allowed us to examine the effect of a diagram hint on the proportion of participants who correctly solved each problem. Table 1 shows the means and standard deviations of participants who correctly solved each problem. Any differences in performance between problems with and without a provided diagram are not statistically significant.

Table 1: Means and standard deviations of correct responses

Problem	Total		Diagram		No Diagram	
	M	SD	M	SD	M	SD
1	0.539	0.500	0.619	0.489	0.462	0.502
2	0.398	0.492	0.339	0.477	0.460	0.502
3	0.305	0.462	0.365	0.502	0.246	0.434
4	0.773	0.420	0.769	0.425	0.778	0.419

We also examined participants' self-generated external visual representations, since we were interested in whether the specific diagram hint we chose to provide was also spontaneously used by students who were not provided with a diagram hint. The problems in this study were typical of those presented in the curricula and students had prior exposure to using specific diagrams for specific problem topics. For example, problem 4 is a conditional probability problem, for which outcome trees are an appropriate representation. Table 2 shows that for all four problems, participants reorganized the given problem information

Table 2: Percentage of participants generating each type of external visual representation for each problem. Dashed lines indicate a cell with fewer than 3 participant uses.

Representation	Problem 1		Problem 2		Problem 3		Problem 4	
	Diagram	No Diagram	Diagram	No Diagram	Diagram	No Diagram	Diagram	No Diagram
Reorganization	43.8	38.4	18.5	75.0	43.8	18.5	35.4	53.1
Pictures	18.8	32.3	--	--	--	--	--	--
Outcome Trees	--	18.5	7.69	6.25	--	--	3.07	45.3
Contingency Tables	--	--	3.07	78.1	--	--	4.61	7.19
Venn diagrams	--	--	--	--	--	--	--	--
Outcome Listings	15.6	15.4	--	--	--	--	--	--
Novel schematic	--	--	--	--	--	--	--	--

more often than any other representation. For problem 1, the provided diagram hint was an outcome tree. Although 18.5% of participants without the diagram also created an outcome tree, 15% of participants given each version also generated outcome listings. For problem 2, the provided diagram was a contingency table, and 78.1% of participants not provided with this diagram chose to create one in solving the problem. A Venn diagram was provided for problem 3; although only 2 participants spontaneously created one, no other representations were used. Finally, problem 4 was accompanied by an outcome tree. 45.3% of students not given an outcome tree created their own in solving the problem.

In order to investigate the role of student ability / problem difficulty on diagram use, we performed a median split on participants' total scores on the four problems, defining two groups of students, low-ability and high-ability. An ANOVA analyzing the effect of provided diagrams showed different effects for these two groups. We hypothesized that the diagram hints might show a facilitative effect only for problems that are hard, but not too hard. Indeed, the pattern of results shows that for both the low-ability and high-ability groups, problems of moderate difficulty were aided by diagrams (Figure 2). "Moderate difficulty" was defined operationally as any diagram showing an overall proportion correct between .3 and .7 for a given ability group. For the below-median group, the problems of moderate difficulty were problems 1 and 4. As seen in Figure 2, problem solving was aided by provided diagrams in these problems, but not for problems 2 and 3. A different pattern emerges for the above-median group. For this group of participants, a facilitative effect is shown for only problem 3, the most difficult problem. An interesting finding is shown for problem 1 in the above-median group. For this problem, providing a diagram (outcome tree) resulted in lower performance than not. It is possible that the outcome tree was not recognized by participants as an appropriate diagram for this problem; indeed outcome listings were spontaneously generated by students, both in the presence and absence of an outcome tree.

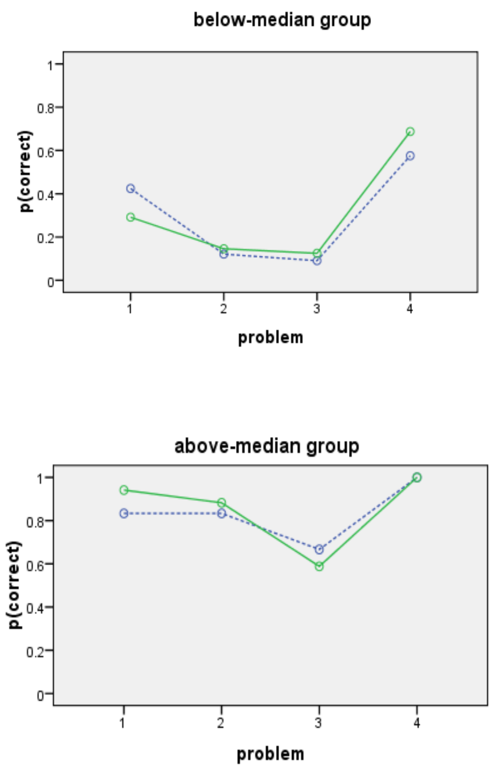


Figure 2. Dotted line shows results for Form A (diagram hints given for Problems 1 and 3); solid line for Form B (diagram hints for Problems 2 and 4).

Discussion

Successful problem solving in mathematics, and especially in PPS, depends on the construction of appropriate representations. External visual representations, including diagrams, are often used to aid in the comprehension and representation of problem information. Diagrams and other external visual devices that are used to comprehend and solve problems are commonplace in the field of mathematics (e.g. Mayer, 1992). Previous research has revealed that using these schema-specific diagrams can facilitate successful probability problem solving (Russell, 2000; Zahner & Corter, 2010). In this study, we investigated whether the presentation of a diagram was related to problem solving success for each problem.

The participants in this study were novice probability problem solvers; they had only received instruction in PPS for a portion of the semester. We hypothesized that providing them with a schema-specific diagram would influence their success with the comprehension and representation phases of problem solving. The vast majority of participants interacted with the provided diagram in some way. Some participants made marks on the diagram, and also drew a diagram of their own, and some participants filled in the diagram with appropriate numbers and calculations. Many of the students used the diagram to organize and rewrite information. However, the students that used the diagram did not necessarily progress to the stage of comprehending the problem sufficiently to plan a solution. We do not have sufficient evidence to conclude that providing a device particular to solving the problem is necessarily an aid to students at all ability levels. Overall, our results show that providing diagrams does not necessarily help students solve a problem successfully.

Our results show that provided diagrams are able to help low-ability and high-ability students differently (cf. Lowrie & Kay, 2001; Uesaka et al., 2007). High-ability participants may not have been helped by a diagram hint because they already possessed a schematic understanding of the problem. They may have generated their own diagram or used a mathematical formula to solve the problem. Low-ability participants, on the other hand, were helped on the less difficult problems only. We posit that providing a diagram hint helped them form a more complete schematic understanding of the problem and helped them achieve a correct solution. For problems beyond the participants' grasp, however, providing a diagram hint did not help. Students must still know how and when to use the diagram in order for it to be an effective tool. Low-ability students may not have been able to associate the diagrams with a schema appropriate for the problem. We interpret these results in the context of Vygotsky's zone of proximal development (Vygotsky, 1978). Providing a diagram hint on problems of moderate difficulty may be sufficient in helping students relate their current schematic understanding of specific types of probability problems to a solution schema, while more assistance may be needed on problems of greater difficulty.

Many students generated their own diagrams when attempting to solve the problems, even when they were given a correct diagram. In their study of mathematical problem solving, van Essen & Hamaker (1990) found that many students generated external visual representations when solving problems. They posit that "generating a drawing does not guarantee that one finds the correct solution, but merely increases the chance that a problem will be conceptualized correctly" (p. 309). Clearly, the nature of instruction contributes significantly to how students use formulas or diagrams. Students must learn how to use diagrams correctly to solve problems and also receive sufficient practice in order to apply the use of those diagrams to new problems. Previous researchers (e.g., Lewis , 1989; Van Meter & Garner, 2005) argue that learning how to represent a problem is essential, and that it can be taught. This study found that when unsuccessful problem solvers were taught how to represent word problems, their scores on a post-test improved significantly.

Understanding the problem schemas and choosing an appropriate representation is a major barrier to successful problem solving. In other words, diagrams must be understood in order to be helpful. A majority of the participants who were not given a diagram and solved the problem correctly generated their own diagram. Future studies need to examine why people drew fewer diagrams on problems that the data indicate to be more difficult, when research suggests they should do the opposite (e.g. van Essen & Hamaker, 1990).

The participants were all enrolled in a probability course which taught problem solving using the types of diagrams chosen for this study, and these problems were typical of those presented in the course. As students become more proficient in PPS, their associations between problem topic and appropriate diagram use likely become stronger. A notable limitation of this study is that the number of problems studied leaves an alternative explanation of the results, namely that certain diagrams may be easier for students to learn to use and associate with problem structure. For example, problems 1 and 4 could both be solved using an outcome tree; these were also the only problems in which a diagram helped the low-ability group. Thus it may be that outcome trees in particular are helpful to low-ability students. Thus the learnability of the diagrams must be considered as a factor when using them to support students. Future studies examining one type of diagram at a time could help provide information about the properties of the diagrams that make them more or less useful for particular problems. To further support our findings, further research examining the use of diagrams and problem difficulty within a given probability topic is warranted.

Novick (2001) argues that spatial diagrams are "tools for thinking" and that successful construction of these diagrams can lead students to see deep similarities among problem situations. These similarities might otherwise not be prominent. It is important to understand the possible advantages of using external visual representations, as they may help problem solvers to build a mental model and to formulate problem schemas. Further research could explain the choices problem solvers make when solving probability problems. Additionally, this research should explore the role of problem difficulty, problem type, and background knowledge.

Acknowledgments

Support for the investigators' research and dissemination activities has been provided by grants IIS-0725223 and IIS-0855995 from the National Science Foundation.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16(3),* 183-198.

Corter, J. E., & Zahner, D. C. (2007). Use of external visual representations in probability problem solving. *Statistics Education Research Journal, 6(1),* 22-50, http://www.stat.aukland.ac.nz/serj.

Edens, K., & Potter, K. (2008). How students "unpack" the structure of a word problem: Graphic representations and problem solving. *School Science and Mathematics, 108(5),* 184-196.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal of Research in Mathematics Education, 19(1),* 44-63.

Koedinger, K., & Anderson, J. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science, 14,* 511-550.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*, 59-98.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65-99.

Lewis, A. B. (1989). Training students to represent word problems. *Journal of Educational Psychology, 81(4),* 521-531.

Logie, R., & Baddeley, A. (1990). Imagery and working memory. In P. J. Hampson, D. E. Marks, & J. T. E. Richardson (Eds). *Imagery: Current Developments* (pp. 103-128). New York, NY: Routledge.

Lowrie, T., & Kay, R. (2001). Relationship between visual and nonvisual solution methods and difficulty in elementary mathematics. *Journal of Educational Research, 94(4),* 248-255.

Mayer, R. (1992). Mathematical problem solving: Thinking as based on domain specific knowledge. In R. E. Mayer (Ed), *Thinking, Problem Solving, Cognition.* (pp. 455-489). New York, NY: W. H. Freeman & Co.

Mayer, R., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82,* 715-726.

Novick, L. (1990). Representational transfer in problem solving. *Psychological Science, 1*, 128-132.

Novick, L. (2001). Spatial diagrams: Key instruments in the toolbox for thought. In D. L. Merlin (Ed.), *The psychology of learning and motivation, 40,* 279-325.

Novick, L., & Hmelo, C. (1994). Transferring symbolic representations across nonisomorphic problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,*1296-1321.

O'Connell, A. A. (1999). Understanding the nature of errors in probability problem solving. *Educational Research and Evaluation, 5,* 1-21.

Russell, W. E. (2000). *The use of visual devices in probability problem solving.* (Doctoral Dissertation, Columbia University, 2000). New York: Academic Press.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficacy of encouraging original student production in statistics instruction. *Cognition and Instruction, 22(2),* 129-184.

Tomlinson, S. & Quinn, R. (1997). Understanding conditional probability. *Teaching Statistics, 19(1),* 2-7.

Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought.* Cambridge, MA: MIT Press.

Uesaka, Y., Manalo, E., & Ichikawa, S. (2007) What kinds of perceptions and daily learning behaviors promote students' use of diagrams in mathematics problem solving? *Learning and Instruction, 17,* 322-335.

Van Essen, G. & Hamaker, C. (1990). Using self-generated drawings to solve arithmetic problems. *The Journal of Educational Research, 83(6),* 301-312.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review, 17(4),* 285-325.

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in Society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Zahner, D., & Corter, J. E. (2010). The process of probability problem solving: Use of external visual representations. *Mathematical Thinking and Learning, 12(2),* 177-204.