EFFECTS OF DIAGRAMS ON STRATEGY CHOICE
IN PROBABILITY PROBLEM SOLVING

Chenmu Xing

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

EFFECTS OF DIAGRAMS ON STRATEGY CHOICE IN PROBABILITY PROBLEM SOLVING

Chenmu Xing

The role of diagrammatic representations and visual reasoning in mathematics problem solving has been extensively studied. Prior research on visual reasoning and problem solving has provided evidence that the format of a diagram can modulate solvers' interpretations of the structure and concept of the represented problem information, and influence their problem solving outcomes. In this dissertation, two studies investigated how different types of diagrams influence solvers' choice of solution strategy and their success rate in solving probability word problems. Participants' solution strategies suggested that problem solvers tended to construct solutions that reflect the structure of a provided diagram, resulting in different representations of the mathematical structure of the problem. For the present set of problems, a binary tree or a binary table tends to steer solvers to use a sequential-sampling strategy, which defines simple or conditional probabilities for each selection stage and calculates the intersection of these probabilities as the final probability value, using the multiplication rule of probability. This strategy choice is structurally matched with the diagrammatic structure of a binary tree or a binary table, which represents unequally-likely outcomes at the event level. In contrast, an N-by-N (outcome) table steers solvers to use of an outcome-search strategy, which involves searching for the total number of target outcomes and all the possible outcomes at the equally-likely outcome level, and calculates the part-over-the-whole value as the final probability, using the classical definition of probability. This strategy is strongly cued by the N-by-N (outcome) table, because the table structure represents all equally-likely outcomes for a probability problem, and

organizes the information so that the target outcomes can be seen as a subset embedded in the whole outcome space. When an N-ary (outcome) tree was provided, choices were split between the two solutions, because the N-ary tree structure not only cues searching for equally-likely outcomes but also organizes the problem information in a sequential-sampling, stage-by-stage way. Furthermore, different diagrams seem to be associated with different patterns of characteristic errors. For example, solving a combinations problem with an N-by-N table tended to elicit erroneous solutions involving miscounting those self-repeated combinations represented by the table's diagonal cells as valid outcomes. Typical errors associated with the use of a binary tree involved incorrect value definitions of the conditional probability of the outcome of a selection. And the N-ary tree may lead to less successful coordination of all the target outcomes for the studied problems, because the target outcomes were dispersed in the outcome space depicted by the tree, thus not salient.

The findings support arguments (e.g., Tversky, Morrison, & Betrancourt, 2002) that in order to promote problem solving success, a diagrammatic representation must be carefully selected or designed so that its structure and content can be well-matched to the problem structure and content. And for computational efficiency, information should be spatially organized so that it can be processed readily and accurately. In addition to the implications for effective diagram design for problem solving activities, the findings also offer important insights for probability education. It is suggested that a variety of diagram types be utilized in the educational activities for novice learners of probability, because they tend to highlight different probability concepts and structures even for the same probability topic.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor James Corter, for his mentorship and support over the course of my study. Thank you for your invaluable and step-by-step guidance on my dissertation and other research work, and also giving me a desirable amount of independence to think and explore ideas. You have helped me to see my potential and have motivated me to work hard on realizing it. You have fostered in me an ever-growing mindset, faith and passion for carrying on my research. For your contributions to my growth and your profound impact on me, I am eternally grateful.

Thank you to the rest of my committee members, Professors Barbara Tversky, Stephen Peverly, Roger Anderson, and Doris Zahner, for your important feedback and warm support at all stages of my dissertation work.

I also gratefully acknowledge Springer's Lecture Notes in Computer Science as the original source of publication for portions of my dissertation to appear. These parts, including but not limited to the description of Study 1 in Chapter 3, are included in a version accepted for publication that can be cited as follows: Xing, C., Corter, J. E., & Zahner, D. (2016). Diagrams affect choice of strategy in probability problem solving. In M. Jamnik, Y. Uesaka, & S. Elzer Schwartz (Eds.), *Lecture Notes in Computer Science (LNCS): Diagrammatic Representation and Inference* (Vol. 9781, pp. 3-16). Heidelberg, Germany: Springer.

In addition, I would like to thank Professor Catherine Chase for being another great source of inspiration and thoughts during my doctoral study. Thank you for your valuable advice and all the opportunities you have given me to expand my research and teaching practice in this field. It is truly pleasant to work with you.

I am thankful to Professors John Black and Catherine Chase for letting me recruit research participants from their classes. I am also thankful for Myra Luna-Lucero and other IRB personnel at TC, whose hard work facilitated the timely completion of my data collection.

I am grateful for the friendship and support from many fellow doctoral students in my department and other important friends. Thank you to Lina (Na Li), Carol Lu and Laura Lu, for all the happiness, help, and encouragement you have given me throughout my graduate school study. Besides all the fun and laughter we have shared, thank you for having kept me strong, encouraged, and perseverant when I face difficult times and self-doubts. Thank you to Sehrish Jamshed, Yangyang Liu, Maria "Sweet" San Pedro, Kristen Bundy, Jie Gao, and many other friends and colleagues at Teachers College. Thank you for your friendship and the insights you have given to my study or life. I am also wholeheartedly grateful to Hong Fan and the rest of the Yu family, Aimee Tiu Wu and the rest of the Wu family, and Joanne Park. You have made the definition of "friend" so close to "family". Your thoughtfulness and hospitality over all these years have made my stay here a better and more memorable experience. And to Mingxing Wu, thank you for your love and endless support, which is beyond words to describe how important it has been for me to complete the doctoral study.

Lastly, I would like to express my deepest gratitude to my parents and grandparents. Especially to my mother, thank you for always listening to my academic experiences, reflections and research thoughts with great patience and cheering for me all the time. Thank you for always encouraging me to look on the bright side whenever a challenge or disappointment happens. Thank you for supporting me in every possible way so that I can realize my dreams. To your love, patience, support and full faith in me, I owe my accomplishments. This work is dedicated to you.

CHAPTER I: INTRODUCTION

Background of the Problem

Diagrams are essential tools for representation, communication, and reasoning. In education, diagrams have been used widely, and play an important role in STEM learning and problem solving (e.g., Hegarty & Kozhevnikov, 1999; Heiser & Tversky, 2006; Lowrie & Kay, 2001; Manalo & Uesaka, 2006; Presmeg, 1986a, 1986b; Zahner & Corter, 2010; Novick & Catley, 2013), learning and comprehension of complex systems (Heiser & Tversky, 2006), judgment (Gattis & Holyoak, 1996; Simkin & Hastie, 1987), reasoning (Tversky, Corter, Gao, Tanaka, & Nickerson, 2013), analogical transfer (Gick & Holyoak, 1983; Novick, 1990), planning (Mason, Corter, Tversky & Nickerson, 2012), and data representation and interpretation (Braithwaite & Goldstone, 2013; Zacks & Tversky, 1999).

But like any tool, a diagram must be well chosen for the task at hand, and its use affects both the process and the product of the activity. First, as an external representation of a cognitive or educational problem, salient aspects of the diagram must map to relevant aspects of the problem (Markman, 1999; Novick and Hurley, 2001). Second, structural, visuospatial, and implicit aspects of the chosen diagram can influence and alter people's perceptions, inferences, and judgments of the relations and structures of the represented information (e.g., Gattis & Holyoak, 1996; Tversky, Corter, Gao, Tanaka, & Nickerson, 2013; Zacks & Tversky, 1999). How a diagram steers people to certain inferences and judgments concerning the represented concepts and relations is not arbitrary. Rather, it stems from cognitively natural ways of mapping visuospatial elements and relations to conceptual content and relations, externally or internally, based on shared metaphorical (or analogous) similarity of abstract relational structures (e.g., Gattis, 2004; Gattis & Holyoak, 1996; Novick & Hurley, 2001; Tversky, Kugelmass, & Winter, 1991).

A rich body of research has explored how specific types of diagrams affect inferences in reasoning and judgment tasks. When asked to describe the relation of individual data points shown in statistical graphs, people given a bar graph tended to describe the relation as comparisons of discrete entities, but as trends of continuous change when given the same information depicted as a line graph (Zacks & Tversky, 1999; Shah, Mayer, & Hegarty, 1999). To describe complex mechanical systems depicted by diagrams, people reading mechanical diagrams with arrows described the functions of the systems, whereas those reading the same diagrams without arrows gave structural descriptions (Heiser & Tversky, 2006). When people tried to keep track of individuals' locations over different time points, data depicted in lines that connected individuals' locations over time led to use of people as the dominant information organizer and their movements over time as the structure of the description. On the other hand, when separate dots were entered in the location-by-time cells to represent individuals, people's attention was directed to the table cells and they were more likely to summarize information by group of people by location (Tversky et al., 2013).

People's inferences using diagrams are systematically related to the schemas that different types of diagrams convey. For example, lines connect and associate entities, indicating paths, relations, and movement (Tversky et al., 2013; Tversky, Zacks, Lee, & Heiser, 2000; Zacks & Tversky, 1999); bars and boxes suggest enclosures and separate categories (Tversky et al., 2013; Zacks & Tversky, 1999); and arrows show asymmetric directions and sequences from actions to goals and causes to effects (Heiser & Tversky, 2006; Tversky et al., 2000).

Furthermore, the degree to which diagram structures can be analogically mapped onto problem structures, achieving the correct mapping, not only affects the type of inferences people make in reasoning and judgment tasks, but it also has a great impact on problem solving success. Research evidence (e.g., Braithwaite & Goldstone, 2013; Gattis & Holyoak, 1996; Novick & Hurley, 2001; Simkin & Hastie, 1987) suggests that a high degree of visual-conceptual compatibility often leads to higher accuracy and faster speed in problem solving. For example, in Gattis and Holyoak's (1996) rate of change judgment tasks, people were given a line graph depicting the relation between two variables, and were asked to infer how the rate of one variable changes with the increase of the other variable. Their findings consistently suggest that for problem solving accuracy, it is crucial for the variable of cause (e.g., the IV) to be assigned to the x-axis and the queried variable (e.g., the DV) to the y-axis, because graph users regularly follow such a mapping convention that a steeper line on the graph indicates faster changes in the y-axis and the queried variable. Violation of the suggested visual-conceptual mapping conventions in graph use leads to significantly lower rates of correct judgment. In a statistical judgment task, Braithwaite and Goldstone (2013) found that people given a line graph depicting the interaction of data points were able more quickly to infer the interaction effect of two variables without loss of accuracy, compared to those using a tabular representation of the data. Hurley and Novick (2010) compared people's accuracy and speed in making inferences about the information that was represented by convention-following and convention-violating diagrams. With a convention-violating representation (e.g., objects assigned to the lines and relations assigned to the nodes in a network diagram), solvers took longer time to make inferences, which were also more erroneous. Zahner and Corter (2010) tested the relation of probability problem solvers' choice of diagrams and their problem solving outcomes. They found that problem solving success was promoted by diagrams, but only when a representation was appropriately matched to a problem type (e.g., Venn diagrams for joint event probability). When Simkin and Hastie (1987) compared the effects of various types of graphs in judgment tasks, bar graphs led to highest accuracy in judging comparisons, while pie charts were most facilitative for judging proportions.

Diagram theorists interpret this kind of phenomena as demonstrating interplay between graph features and problem information types on facilitating or biasing perception and judgment (e.g., Pinker, 1990; Simkin & Hastie, 1987; Zacks & Tversky, 1999). According to Pinker (1990), a diagram itself does not possess universal advantage or disadvantage for information processing. The features of a diagram interact with the types of information to be represented so that the extraction of certain information or structures may be facilitated by certain diagram types but hindered by some other diagram types. Tversky, Morrison, and Betrancourt (2002) proposed two principles for effective graphic design that facilitate information extraction, comprehension, and inference using graphs. The Congruence Principle, based on the cognitive naturalness of visual-conceptual mappings, emphasizes compatibility between the structure and content of a graph and the desired structure and content of the information to be represented. The Apprehension Principle emphasizes optimal organization and display of information so that it can be perceived readily and accurately.

Overview of the Present Research

The present research gathers evidence that can be used to evaluate theories of WHY diagrams are useful in reasoning, learning and problem solving, and WHEN diagrams are useful. More specifically, the goal is to investigate how structural and content compatibility between diagrams and problems might apply in probability problem solving. Based on the theories and evidence reviewed earlier, we argue that in mathematics problem solving, appropriate diagrams both direct solvers' attention to the underlying mathematical structures via cognitively natural correspondences, and help to organize problem information to facilitate the coordination of sub-goals and computational stages in the process of problem solving.

Elementary probability problems can sometimes be solved by more than one appropriate solution method. For example, problems are used here that can be solved using strategies based on two distinct probability concepts, the classical definition of probability and the multiplication rule of probability. However, applied probability problems describe a wide variety of real-world situations. Based on evidence from related work (as yet unpublished), I believe that the semantic aspects of these problems evoke interpreted problem structures that may or may not map naturally to one or more probability concepts and solution strategies underlying these problems. It is predicted that providing an appropriate diagrammatic representation can highlight the underlying problem concept and structure as its diagrammatic structure and components will map to the probability concept.

In the current investigation on diagram effects, we focus on tree diagrams and tables, because they are among the most commonly used diagrams in probability education and problem solving (e.g, Russell, 2000; Zahner & Corter, 2010). Research suggests that these two types of diagrams are often used by problem solvers to represent different probability situations and schemas (e.g., Corter & Zahner, 2007; Novick, 1990; Novick & Hmelo, 1994; Zahner & Corter, 2010). In probability problem solving, N-by-N tables are frequently used to represent all possible combinations for compound-events problems (e.g., Novick, 1990), while outcome trees seem especially appropriate for situations involving sequential selection (e.g., Novick & Hmelo, 1994). Zahner and Corter (2010) found that using tree diagrams was particularly useful for solving conditional probability problems, a type of probability problem that involves reasoning about sequential processes and dependent events. These observations and empirical findings suggest that the structures of trees and tables may be best for representing different problem structures and probability concepts, and may facilitate solution of different types of probability problems. However, this effect has only been demonstrated or hypothesized for a few specific problem instances. The current research aims to specify and summarize the types of probability problems where trees and tables are best suited for problem representation, and to explain these benefits in reference to specific aspects of the diagrams and visual reasoning processes in understanding probability concepts and procedures.

In this research, the aim is to test diagram effects in the domain of probability problem solving. The studies described below explore how using different types of diagrams can affect both the process (strategy choice) and the product (solution success) in probability problem solving. Specifically, the studies investigate whether and how different types of diagrams can steer probability problem solvers to choose one solution strategy or another for solving probability problems that admit of multiple types of solution strategies. Specifically, four types of generic (i.e., unlabeled) diagrams are used as potential aids in problem solving: binary trees, N-ary (outcome) trees, binary tables, and N-by-N (outcome) tables. By comparing the effect of these different diagrams on solvers' choice of solution strategy and solution correctness, the two studies of this dissertation seek to answer the following research questions:

Do different types of diagrams for representing probability problems elicit the use of different solution strategies?

Do appropriately chosen diagrams increase solution correctness rates, compared to no diagram given? Does an ill-chosen diagram hurt performance?

If such effects are found, what aspects of the diagram seem to account for the differences in strategy choices and may be instrumental to obtain the facilitative effect?

Significance of the Research

The current research has the potential to add three distinctive contributions to research on the role of diagrams in thinking, reasoning, and STEM education.

First, the current research expands the ways in which visual representations have been shown to influence the thinking process and outcomes in cognitive activities. Prior studies have predominantly focused on how diagrams can influence the outcomes of reasoning, such as inferences and judgment (e.g., Heiser & Tversky, 2006; Tversky et al., 2013; Zacks & Tversky, 1999), or the accuracy or speed of problem solving and transfer (e.g., Gattis & Holyoak, 1996; Mason et al., 2012; Sanfey & Hastie, 1998). This research aims to produce empirical evidence that diagrams also affect how people formulate or select solution strategies in the process of problem solving.

Second, this research investigates the particular aspects of tables and tree diagrams that might affect problem solving, beyond their general structure type. Previous investigations have focused on the global structures and applications of schematic diagrams such as tables and tree diagrams in their general format (e.g., Novick & Hurley, 2001; Novick, Hurley, & Francis, 1999), or on the features and applications of some specific variants of a diagram such as cladograms and polygenetic trees (e.g., Novick & Catley, 2013; 2014). The second study described below systematically manipulates the levels of information abstraction in tables and tree diagrams, to test how this type of visual representation feature interacts with the general structure type of a diagram to affect choice of solution strategy and solution correctness.

Lastly, the current research expands the study of diagram effects on reasoning specifically in the domain of statistics and probability learning and problem solving. The studies explore the issue of compatibility of diagram structures and probability problem structures, and its effect on solvers' choice of mathematical solution strategies and solution accuracy. The results should have important educational implications for statistics and probability education, because they address how effective visual representations can be designed to facilitate the conceptual understanding of probability and the procedural flexibility for solving its problems, and the factors that might affect novice problem solvers' choices of formal solution strategies.

This dissertation is divided into five chapters. Chapter One provides the background of the problems explored in this dissertation, the goal of the research, and the research questions posed. Chapter Two presents a review of the literature related to the research questions. Chapters Three and Four report the two studies conducted to address the research questions. These two chapters describe in detail the research methodologies including the design, procedure, and test materials. Analysis results and some discussions are also provided following the description of methodology. This dissertation is then concluded by Chapter Five, which provides an overall and general discussion that synthesizes the results of the two studies and the educational and cognitive implications with regard to the research problem this dissertation sets out to investigate.

CHAPTER II: LITERATURE REVIEW

External visual representations, such as diagrams, have a great impact on perception, reasoning, and problem solving, in both mathematics and more general domains. This chapter is a review of the role of external visual representations, especially diagrams, in mathematics problem solving. The primary goal of this chapter is to answer why and when diagrams are useful for problem solving, with an additional focus on how appropriate and effective diagrams may be designed for representing probability problems, highlighting their underlying concepts and structures, and facilitating problem solving.

Visual Representations in Mathematics Problem Solving

Applications of Visual Representations in Mathematics Activities

External visual representations play an important role in mathematics learning and problem solving (e.g., Arcavi, 2003; Barwise & Etchemendy, 1991; Bishop, 1989; Bruckheimer & Arcavi, 1995; Hadamard, 1945; Nemirovsky & Noble, 1997; Zahner & Corter, 2010). Use of external visual representations can be found in various types of mathematics activities. For example, visual representations are commonly seen in instructional materials and teaching activities (e.g., Banilower, Smith, Weiss, Malzahn, Campbell, & Weis, 2013; Dufour-Janiver, Bednarz, & Belanger, 1987; Fuson & Briars, 1990). Learners and problem solvers often spontaneously create many forms of visualizations such as pictures, diagrams, and graphs across all stages of problem solving, e.g., problem comprehension and representation, reasoning, solution formulation, and solution explanations (e.g., Edens & Potter, 2008; Hegarty & Kozhevnikov, 1999; Presmeg, 1986a, 1986b; Zahner & Corter, 2010).

In these mathematics activities, a wide variety of types of visualizations have been used. Their forms of representation may vary largely, from concrete (e.g., pictures, manipulatives) to abstract (e.g., symbols, graphs), and from static (e.g., diagrams) to dynamic (e.g., animations, gestures). These various types of visualizations serve multiple different purposes in mathematics. For example, external visualizations have been used to symbolize mathematical notations, concepts, meanings, and formal solutions (Arcavi, 2003). Schematic diagrams such as tables, trees, and Venn diagrams, are important external visual devices for representing problem information and cueing mathematics problem solutions (Novick, 1990; Polich & Schwartz, 1974; Schwartz, 1971; Schwartz & Fattaleh, 1972). Graphing, such as lines, bars and scatter plots, has been used as an important means for data representation, feature discovery, and pattern interpretation, especially in statistics (e.g, Arcavi, 2003; Anscombe, 1973; Gattis & Holyoak, 1996: Pearson, 1895; Salkind, 2006; Zacks & Tversky, 1999).

Effects of Visual Representations on Mathematics Problem Solving

However, research on the effects of using visual images on mathematics problem solving has shown mixed findings (e.g., Hegarty & Kozhevnikov, 1999; Lean & Clements, 1981; van Garderen & Montague, 2003). For example, a correlational study (Lean & Clements, 1981) measured engineering students' preference for using visual representations and their mathematics test performance. The results revealed that students who preferred to process mathematical information by verbal means tended to outperform those using visual means on the mathematics tests. The authors speculated that the poorer performance associated with visual reasoning was due to reliance on concrete pictorial representation of problem information, which distracted problem processing and solving with unnecessary information. Similarly, in Hegarty and Kozhevnikov's study (1999) on the relationship between use of visual representation and students' mathematics problem solving performance, it was found that not all visual representations were associated with higher problem solving success, but only those that depicted the schematic structures of the problems. These findings, which revealed negative or mixed effects of using visual solutions on mathematics problem solving, are in conflict with the positive visual effects found by other studies (e.g., Lowrie & Kay, 2001; Webb, 1979).

Taxonomy of Visual Representations in Mathematics Problem Solving

To explain the contradictory findings of the role of visualizations in mathematics learning and problem solving, the research literature has suggested that different types of visual displays used in mathematics activities should be distinguished, and that their effects on facilitating problem solving may vary, depending on their visual types and schemas. Presmeg (1986a, 1986b) distinguished five types of visual imagery in mathematics: pictorial imagery, schematic pattern imagery (e.g., diagrams), kinesthetic imagery, dynamic imagery, and memory for formulas.

Among these five different types of visual imagery that Presmeg has distinguished, it was suggested that the most effective visual representation format should be schematic pattern imagery, because it removes concrete, mathematically irrelevant information from the problem information, and only displays the essential relations described in the problems. Use of concrete pictorial imagery, on the other hand, usually leads to a lower rate of success in mathematics problem solving (Hegarty & Kozhevnikov, 1999; van Garderen & Montague, 2003). Pictorial representations generally impair problem solving success, because the depictions often include irrelevant details for problem solvers to process, which distracts their attention from processing only the essential information and the underlying structure of the problem (Hegarty & Kozhevnikov, 1999; Lean & Clements, 1981; Presmeg, 1986a, 1986b). That schematic diagrams facilitate mathematics problem solving has been supported by previous research. For example, Hegarty and Kozhevnikov (1999) studied how sixth grade students' self-created visual representations affected mathematics problem solving. Students' visual creations were coded as either primarily schematic (or diagrammatic) or primarily pictorial, based on whether a diagram or abstract spatial relations were depicted in their gestures or sketches. The results show that while the use of schematic diagrams was associated with higher problem solving success, the use of pictorial representations was associated with more problem solving failures. Similarly, Zahner and Corter's study (2010) examined probability problem solvers' spontaneous use of visual representations, and found that problem solving was more successful when diagrams were used, such as trees and Venn diagrams, but not when pictorial images were drawn. Gattis and Holyoak (1996) manipulated the assignment of an independent variable (altitude) to the x- or the y-axis of a line graph and explored how that affected graph users' accuracy in judging the rate of change of the dependent variable (temperature) as the independent variable (altitude) changes. In one condition, they assigned the variable of altitude to the y-axis so that the graph observes a pictorial correspondence of "up" between the concept of altitude going "up" and the graphic depiction of the y-axis going "up". In another condition, they assigned altitude to the x-axis and temperature to the y-axis so that the graph observes an abstract, schematic correspondence between the concept of temperature being the effect and the diagrammatic convention of the y-axis being the outcome variable. This reversed assignment of variables to axes did not observe the pictorial correspondence depicted for the other condition. Participants' performance confirmed that accuracy was increased by the diagrammatic correspondence, but decreased by the pictorial correspondence.

In conclusion, previous research suggests that among all types of external visual representations, diagrams may be the most facilitative type of external visual representation for mathematics learning and problem solving. Therefore, it is educationally important to examine closely the use of diagrams in mathematics problem solving and their effects in problem solving, and to seek principles for designing or selecting effective diagrammatic representations as aids for problem solving.

In the following sections, the discussion will be focused on diagrams, used in the narrow sense of abstract schematic pattern representations. The role of diagrams in mathematics reasoning and problem solving activities will be reviewed. Principles for effective diagram design will also be suggested.

Diagrams in Mathematics Problem Solving

Definition of a Diagram

A diagram is defined as a type of graphic representation that depicts only abstract structures and spatial relationship without references to literal, quantitative, or context-specific information (Brasseur, 2003; Lowe, 1993). Examples of diagrams that are commonly used in mathematics (including statistics and probability) problem solving include tables, trees, Euler or Venn diagrams, and in a broad sense also graphs and charts such as pies, bars, lines, and networks such as bipartite graphs.

Functions of Diagrams in Problem Solving

As stated earlier in this paper, diagrams facilitate mathematics problem solving, compared to other types of visual representations such as pictures and icons. Then the question of interest becomes: how and why do diagrams facilitate mathematics problem solving?

This can be explained by the various functions that diagrams serve in general as well as in mathematics problem solving. Commonly acknowledged advantages of using diagrams over just the text information include that diagrams can ease the process of problem understanding by schematizing and simplifying information from problem text; and diagrams can increase computational efficiency by structuring problem information so that they can be easily grasped (e.g., Diezmann & English, 2001; Fagnant & Vlassis, 2013; Gagatsis & Shiakalli, 2004; Hegarty & Kozhevnikov, 1999; Larkin & Simon, 1987; Lowe, 1993; Presmeg, 1986b, 2006; Tversky, Morrison, & Betrancourt, 2002; Zahner & Corter, 2010).

To illustrate this point, Larkin and Simon (1987) contrasted the computational efficiency for processing information that is of the same quantity but represented in two different formats: sententially or diagrammatically. Sentential representations equalize the accessibility of information across the text, thus making certain information implicit and less noticeable, or costing readers extra time and effort to extract the necessary information for use (Larkin & Simon, 1987). On the contrary, diagrammatic representations can make implicit text information explicit and easy to grasp, by organizing information by location on a plane, and by chunking and/or highlighting necessary information for processing and computation (Larkin & Simon, 1987). In addition, Gattis and Holyoak (1996) argued that sometimes graphs can integrate or reduce the number of steps or dimensions that are otherwise involved in purely sentential or mathematical solution procedure. This is argued to be an advantage over sentential representation, because trying to coordinate information across various dimensions and stages often costs a heavy cognitive load (Gattis & Holyoak, 1996). Thus, information represented diagrammatically has often been found to be more easily recognized, coordinated, and computed (Gattis & Holyoak, 1996; Larkin & Simon, 1987).

Similar to this reason (e.g., Gattis & Holyoak, 1996; Larkin & Simon, 1987) given as to why diagrams facilitate information processing and computational efficiency, Winn (1989) argued that diagrams can simplify otherwise complex information described in narratives, and abstract necessary concepts from unnecessary or irrelevant details. For example, a food-chain diagram can metaphorically depict the roles of animals and their predator-prey relations more efficiently than sentential descriptions (Winn, 1989). A map, as another example of schematic visual representation, eliminates photographic details of roads and architectures, and highlights essential information such as directions, locations, architecture categories, and altitudes (Winn, 1989). Such a diagrammatic representation is argued to increase efficient perception and access to essential information (Winn, 1989).

In addition to easing information processing and computations, several other functions of diagrams have also been identified that may explain why they help mathematics problem comprehension and solving. Diagrams can be used for developing mathematical insights (e.g., Polya, 1957; Edens & Potter, 2008). They can provide external support to problem solvers' internal representations and be used to offload intermediate computational results from working memory (Bauer & Johnson-Laird, 1993; Tversky, 2001). And diagrams can also serve as memory cues for problem solutions (e.g., Larkin & Simon, 1987; Novick & Hmelo, 1994; Novick, 1990; Phillips, Norris & Macnab, 2010; Stylianou & Silver, 2004).

Furthermore, using diagrams has been found to be particularly important or useful for solving complex or difficult problems (e.g., Bobek & Corter, 2010; Lowrie & Kay, 2001; Manalo & Uesaka, 2006; Webb, 1979). For difficult or complex problem solving, diagrams can help to coordinate sub-goals (e.g., Zahner & Corter, 2010); to simplify complex situations in the problems and make computations easier (Larkin & Simon, 1987; Lynch, 1990; Winn, 1989); and to offload intermediate calculation results (Schreiber, 2004; Tversky, 2001).

Diagrams and Mathematics Problem Solving Success

Although both theories and empirical evidence suggest that diagrams are powerful representational formats in reasoning and problem solving activities, research has also revealed that diagrams are not invariably associated with problem solving success, nor are they effective for solving problems across all types (Zahner & Corter, 2010).

First, diagrams do not always improve problem solving performance, especially if their perceptual properties fail to conform to the problem schemas they try to represent or the graphing conventions (e.g., Gattis & Holyoak, 1996; Novick & Hurley, 2001). For example, Bauer and Johnson-Laird (1993) tested the effectiveness of different types of diagrammatic representations for syllogistic reasoning. The results showed that only those diagrams of which the spatial relations were depicted analogous to the structure and content of the problem information significantly increased answer accuracy compared to no diagram given. On the other hand, when a diagram depicted problem elements and relations in a spatially arbitrary way, it even decreased answer accuracy compared to verbal reasoning only. In Hurley and Novick's (2010) study, the results of diagram users' performance in judgment tasks showed that convention-violating diagrams not only impaired answer accuracy but also prolonged the reaction time. Thus, Hurley and Novick (2010) concluded that effectiveness of diagrams for accurate and efficient problem solving requires that a diagram is constructed so that it *facilitates* the perceptual inferences of the problem information.

Second, a diagram may be an appropriate representation for one type of problem, but not another, depending on the degree to which its structure and content can correspond to those of the problems. For example, in a study that examined the effect of various types of user-created visualizations on probability problem solving (Zahner & Corter, 2010), the results showed that the effectiveness of a diagram for problem solving was contingent on the types of probability problems it was created to represent. For example, the use of tree diagrams was found facilitative for solving problems of conditional probability and combinations, but not significantly so for compound independent events or fundamental principles of combinatorics. The tree structure features conditional and hierarchical leveling (Novick & Hurley, 2001) and typically cues a temporal ordering schema (Zahner & Corter, 2010). This may explain why a tree diagram is particularly helpful for solving conditional probability and combinations problems, because it is structurally well-matched to the problem features, which involve sequential events and finding probabilities that are dependent of the previous event outcomes. However, such a visual schema may not always be essential for highlighting the probability concepts of compound independent events or fundamental counting principles, because events are independent of each other, and sometimes the efficient search for outcomes does not follow a temporal or hierarchical order.

Given the evidence reviewed above, the next question to be addressed is: what types of diagrams are effective for mathematics problem solving? Or in other words, what types of diagrams should mathematics educators provide to facilitate problem understanding, ease information processing and search, increase computational efficiency, and thus improve student problem solving?

Principles for Effective Diagram Design

Simply put, a diagram must be an appropriate representation of the problem information so that necessary information can be highlighted and accurately grasped for use. To be specific, Tversky, Morrison, and Betrancourt (2002) have suggested two cognitive principles for effective diagram design: the Principle of Congruence and the Principle of Apprehension. The Principle of Congruence requires the content and structure of a diagram to be consistent with the content and structure of the problem that it is used to represent; the Principle of Apprehension suggests that the information and schema that a diagram conveys must be organized so that it can be easily noticed and grasped (Tversky, Morrison, & Betrancourt, 2002).

Practically speaking, that means an effective diagram must be designed or selected to share a high structural similarity with the type of problem it is to represent in order to facilitate problem solving. Why is this structural compatibility between a diagram and a problem type particularly important? I believe this can be explained by the cognitively natural correspondences between perceptual properties and nonvisual concepts, and the graphing and communicative conventions that graph authors and viewers regularly observe (Zacks & Tversky, 1999). Furthermore, research evidence has suggested that these visual-conceptual correspondences and graphic use conventions are not only consistently followed by experienced graph authors and users (Novick & Hurley, 2001), but also by people without awareness or explicit knowledge of these conventions (Zacks & Tversky, 1999).

To conclude, different types of diagrams are applicable to different situations, and are usually facilitative only for solving those problems of which they are appropriate visual representations. Different types of appropriate diagrammatic representations may lead problem solvers to extract different types of information depending on their unique perceptual structures, perceive and interpret the structure of a mathematics problem differently, and thus choose different solutions (Gattis & Holyoak, 2004; Pinker, 1990; Zacks & Tversky, 1999). The degree of shared structural similarity between a diagram and a problem type can also affect problem solving success, so that a higher structural compatibility may lead to a higher rate of solution success.

Diagrams for Probability Problem Solving for the Current Study

In this dissertation, two types of diagrams, namely trees and tables, are examined in terms of their effect on solving probability problems of different topics. Trees and tables have been commonly-used visual representations in probability education and problem solving activities

(Bobek & Corter, 2010; Russell, 2000; Zahner & Corter, 2010). However, they differ in their diagrammatic structures (Novick & Hurley, 2001), and thus may highlight different schematic information about a probability problem, and lead to different inferences and ways of using the problem information. An overview is provided that analyzes the structures of trees and tables, and the probability concepts and problem structures that they may cue respectively.

Tree versus Table: Structures and Applications

Novick and Hurley (2001) analyzed the basic structures and schematic components of trees and tables. The global structure of a table features the cross-classification of elements of two variables or sets, or elements of a single set that are selected twice. The rows and columns each represent a variable or set, with each row or column cell representing an element contained in a set. Each intersection cell represents the combination of two elements that are selected respectively from the two sets or the two selections. The diagonal cells of an N-by-N table are special intersection cells when the table represents combinations of elements of a single set, because a diagonal cell represents the combination of an item by itself (e.g., an outcome that the same element is selected for both times). The global structure of a tree diagram features hierarchical levels of events. The tree structure starts with a single node, which branches out into subsequent levels; and these subsequent levels serve as the root nodes for their following level, and branch again. Different levels of events are often dependent so that the identities of one level depend on the identities of the preceding level, although it depends on the specific situations. Items listed at the same level are mutually exclusive and identical in status, whereas items listed at different levels differ in status or sequence.

Probability Concepts and Problem Types

Three probability topics that are among the most commonly taught topics in elementary probability are permutations, combinations, and the fundamental counting principle of independent events.

All these types of problems ask for the probability that an intersection (or combination) of the favorable events will occur. However, two underlying concepts distinguish these three types of problems: whether joint events are independent of each other; and whether the order of the combination s is important or not. Table 1 shows the probability principles underlying these types of problems.

In permutations, objects are selected or arranged with regard to the order of an object position. In combinations, objects are selected without regard to the order of an object position. And in both permutations and combinations, events of selecting or arranging an object are dependent of each other, which is that the probability that a selection occurs affects the probability of the next selection. For example, the probability of the outcome of selecting a second ball from an urn of five balls will be affected by the probability of the outcome of the first selection, because the number of balls that remain selectable is reduced. On the other hand, for independent events, the selections of the favorable events are independent of each other, which is that the probability of the outcome of the first selection does not affect the second selection. For example, the probability of getting a tail from flipping a coin is not affected by flipping another coin, because all the possible outcomes for such an event to occur remain the same, regardless of the number of the trials. To put it in a more practical way, the same population element cannot be selected at two different “places” in permutations or combinations problems, but can keep occurring in some types of independent events.

Table 1
Key Probability Concepts (and Distinctions) Underlying the Probability Topics

Probability Topics	**Joint Events/Outcomes Independent?**	**Order Important?**
Permutations	Dependent (sampling without replacement)	Yes
Combinations	Dependent (sampling without replacement)	No
Independent Events	Independent (sampling with replacement)	Yes

Conceptually (as well as procedurally), probability for these situations can be defined and computed in two ways: the classical definition of probability; and the multiplication rule of probability.

Following the classical definition of probability, measuring the probability that an event will occur is based on *outcomes* that are equally likely to happen. Thus, in a classic sense, the probability of an event is determined by the number of all possible outcomes of the target event over the number of all possible outcomes that are equally likely to happen. For example, to find the probability two girls will be selected from a group of three girls and two boys, the classical definition will define the probability as 3/10, by dividing 3 (the number of all possible outcomes that two girls will be selected) into 10 (the number of all possible outcomes that two students will be selected). A strategy like this is termed the "outcome-search" strategy.

In a different sense, probability can be defined as the chance of the intersection of *events* by the multiplication rule of probability. According to the multiplication rule of probability, the probability is determined as the chance that all the target events will happen. For example, to find the probability that both events, A and B, will occur, the multiplication rule defines the probability as $P(A \cap B) = P(A) \times P(B|A)$, meaning the probability that both A and B will occur is the probability that A will occur multiplied by the probability that B will also occur given A occurred. Using the same probability problem above, the multiplication rule will define the probability as 3/10 using a different approach – by multiplying 3/5 (the probability that the first selection is a girl) by 2/4 (the probability that the second selection is a girl, too, given the first girl has been selected). A strategy like this is termed the “sequential-sampling” strategy.

Structural Compatibility between Diagrams and Probability Problems

Given this conceptual and procedural understanding of probability in permutations, combinations, and fundamental counting principle of independent events, tables and trees may naturally cue students with different underlying concepts and structures for probability problems, according to the diagrammatic structures they each depict.

An N-by-N table (shown in Figure 1 and Figure 2) is used to illustrate the table structure in relation to the kind of probability concepts and structures that it may highlight. An N-by-N table is the prototype of the table structure. As Novick and Hurley’s (2001) diagram analyses demonstrate, the cells of such a table represents all possible equally-likely outcomes of the target event and in the whole outcome space. Furthermore, as the example shows, it depicts the target outcomes as a subset of outcomes embedded in the whole outcome space. Therefore, it is reasonable to speculate that the structure of an N-by-N table cues problem solvers to process a probability problem in the sense of classical definition, and to search for all possible and all target outcomes that are equally likely to happen.

Although an N-by-N table should cue the classical definition of probability and an outcome-search solution strategy in general, it has different levels of structural compatibility with problems of different topics, depending on whether events are dependent or independent. To represent the equally-likely outcomes for an independent events problem, every intersection cell represents a valid combination of the selections from two sets, or from a single set twice (represented by the row and the column). However, to represent the outcomes for a permutations or combinations problem, only the non-diagonal intersection cells are valid representations. In permutations and combinations, events are dependent, which means the selection of an object cannot be repeated. Therefore, as all the diagonal cells represent those self-repeated combinations, they are invalid and should not be included in the counting of equally-likely outcomes. This difference in the visual-conceptual compatibilities suggests that the structure and components of an independent events problem can be better mapped onto the visual structure and components of an N-by-N table, compared to a dependent events problem (i.e., permutations, combinations).

Student A	Student B				
	BOY	BOY	GIRL	GIRL	GIRL
BOY		--	--	--	--
BOY	BB		--	--	--
GIRL	GB	GB		--	--
GIRL	GB	GB	GG		--
GIRL	GB	GB	GG	GG	

Figure 1. An annotated N-by-N table for a combinations problem. The intersection cells represent all possible outcomes that two students may be selected from a group of five students (represented by the table column and row). The stricken diagonal cells indicate invalid outcomes. The dashes indicate non-essential repetitive outcomes.

Spinner A	Spinner B		
	RED	BLUE	WHITE
RED	R-R	R-B	R-W
BLUE	B-R	B-B	B-W
WHITE	W-R	W-B	W-W

Figure 2. An annotated N-by-N table for an independent events problem. The intersection cells represent all possible color combinations for the two spinners (each with three equal-size color sections, represented by the table column and row).

A binary tree (shown in Figure 3) is used to illustrate the tree structure and how it may highlight certain probability concepts and structures. A binary tree features a hierarchical structure where events of unequally-likely probabilities are depicted to happen sequentially from left to right (or top to bottom if rotated by 90 degrees clockwise). Therefore, I believe that with this kind of a diagram structure, a binary tree will cue problem solvers to solve probability problems with the multiplication rule, which defines the probability of the outcomes of each selection event sequentially, and then multiplies them to find the probability that all target events will occur.

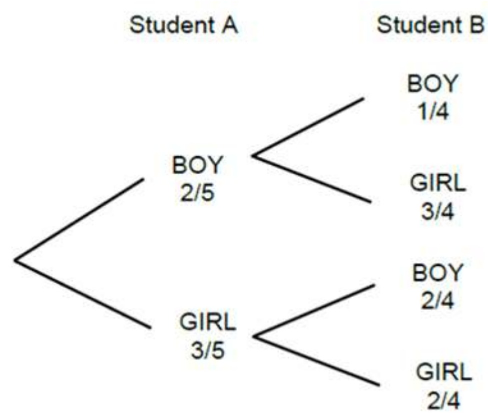


Figure 3. An annotated binary tree for a combinations problem. The nodes of a level represent the outcomes and their probabilities from a certain selection, as two students are selected from a group of five students (two boys and three girls).

Student A	Student B	
	Boy	Girl
Boy	(2/5)*(1/4)	(2/5)*(3/4)
Girl	(3/5)*(2/4)	(3/5)*(2/4)

Figure 4. An annotated binary table for a combinations problem. Each intersection cell represents the joint probability of the outcomes of two selections, as two students are selected from a group of five students (two boys and three girls).

The diagram effect may become more complicated (or flexible) when variation formats of trees and tables (i.e., binary tables and N-ary trees) are involved in problem representation. For example, although a binary table (shown in Figure 4) has the global structure of a table, it represents the intersections of events that are unequally likely to happen, rather than equally-likely outcomes. Therefore, although the global structure of a binary table visually depicts the target event (i.e., selecting two girls) as a subset of all possible event intersections, it may still cue the sequential-sampling strategy that defines probability by the multiplication rule.

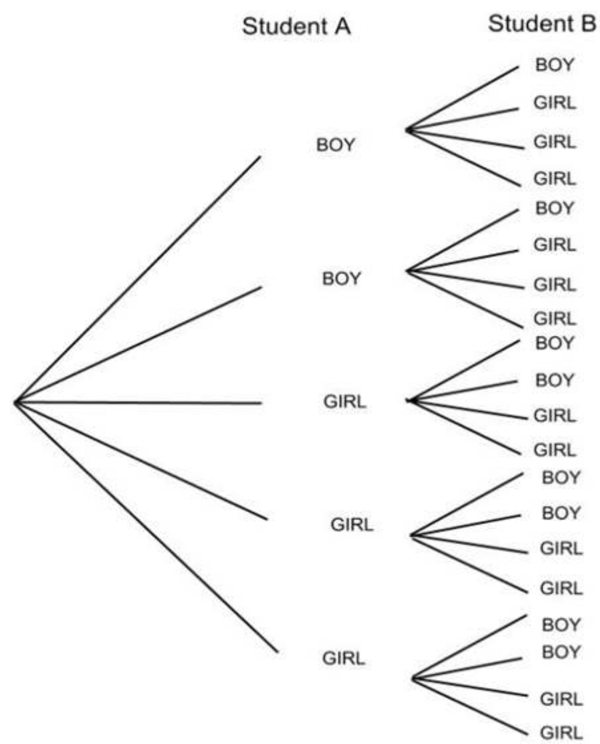


Figure 5. An annotated N-ary tree for a combinations problem. The nodes of a level represent all possible outcomes from a certain selection, as two students are selected from a group of five students.

On the other hand, although an N-ary (outcome) tree (shown in Figure 5) has the global feature of hierarchy and sequence, it enumerates all equally-like outcomes. Therefore, users may be cued to either principle. N-ary (outcome) tree users may be cued to use an outcome-search strategy if their attention is directed to the outcomes that the diagram enumerates. Or, they may be cued to use a sequential-sampling strategy, if their attention is directed to the selection events that the diagram depicts in a sequential order. However, even if the N-ary (outcome) tree structures may cue the same solution strategy as the N-by-N (outcome) table structures, their effectiveness may be different. Each has its own advantage and disadvantage in terms of facilitating problem solving accuracy. For solvers to search for outcomes for an independent events problem, information may be less efficiently organized by an N-ary (outcome) tree as the tree branches are likely to distance one target outcome from another, and make the search of all target outcomes more difficult than an N-by-N (outcome) table. On the other hand, this information is chunked by an N-by-N (outcome) table that integrates all the equally-likely outcomes within the table space for computational efficiency. However, for representing permutations and combinations problems, N-ary (outcome) trees may facilitate problem solving accuracy more, because the branches of an N-ary (outcome) tree do not include any false self-repeated combinations, but the diagonal cells on an N-by-N (outcome) table can easily trap solvers to count them as valid outcomes.

The goal here is to investigate whether and how trees and tables at different levels of outcome abstraction (i.e., binary tree; N-ary tree; binary table, N-by-N table) lead to different results in interpreting mathematical structures and selecting solutions for probability problems. The diagrams were selected to manipulate two factors: the type of diagram structure (tree versus table) and the abstraction level of the represented outcome space (either a large space of equally-likely outcomes, or a smaller space of unequally-likely outcomes). To test the diagram effects on choice of solution strategy and solution correctness, problems representing two probability topics were used: combinations and fundamental principles of independent events. These specific types of problems were chosen because they admit of multiple types of solution strategies (i.e., the outcome-search strategy and the sequential-sampling strategy), and each can be represented by both trees and tables.

It is hypothesized that the choice of solution strategy is influenced by the diagrammatic representation of a problem. The idea is that different types of diagrams should bias people to formulate different mathematical solution strategies, due to the similarity correspondences between a diagram's structure and the selected strategy's mathematical structure. That is, the problem solver tends to choose a solution with procedural or mathematical structure that can be easily aligned with the structure of a diagram. Specifically, it is predicted that the N-by-N table will lead to more frequent use of an outcome-search strategy for the problems, because it represents the problem outcomes at an equally-likely outcome basis, and its N-by-N matrix structure integrates these combination outcomes so that the target outcomes and all possible outcomes may be perceived as a simultaneously sampled subset of a full set. The binary tree is predicted to cue the sequential-sampling strategy, because it features a hierarchical and sequential leveling structure with the nodes at each selection (branch) level representing unequally-likely outcomes of a selection. The binary table is also predicted to cue the sequential-sampling strategy, because the four cells of the table represent selection events with unequally-likely outcomes, and the overall structure does not offer a high compatibility with the outcome-search solution structure (i.e., equally-likely target outcomes as a subset of all possible outcomes). Finally, for the N-ary (outcome) tree, it is hypothesized that both strategies may be cued, because it not only features a sequential and hierarchical structure that may cue the sequential-sampling strategy, but also cues searching for all possible equally-likely outcomes.

With regard to the influence that different types of diagrams have on problem solving, it is hypothesized that when the content and structure of a diagram is well-matched to those of a problem type, and when information is organized in an efficient way that minimizes extra information processing or coordination of sub-stages of solution procedure, problem solving success will be high.

CHAPTER III: STUDY 1

This chapter describes the design and results of the first study (Study 1). The design of the study will be described in detail, followed by the data analysis results, and a discussion of the study findings and implications with regard to the research questions this dissertation sets out to explore. The effects of three types of generic diagrams for solving two probability word problems were tested via a self-paced problem solving task. The primary goal was to examine: first, whether different types of diagrams can steer probability problem solvers to choose different solution strategies for problems that admit of multiple strategies; and second, whether they affect problem solution correctness.

Method

Participants

The participants were 48 students (39 or 81.3% female) recruited from a university in New York City. Their average age was 25.60 years (SD = 4.03 years). To be qualified for the experiment, a participant had to have taken at least one undergraduate- or graduate-level statistics course prior to participation. On average, participants reported to have taken 2.33 statistics courses. However, their levels of training varied: 19 (or 38.6%) participants reported having taken one statistics course, 11 (or 22.9%) participants reported two, 10 (or 20.8%) reported three, and 8 (or 16.7%) participants reported four or more such courses. Their undergraduate major background also varied: 15 (or 31.3%) reported non-STEM majors (e.g., literature, music), 21 (43.8%) in social science (e.g., psychology), 5 (10.4%) in mathematics or statistics, and 7 (14.6%) in other STEM domains (e.g., natural science, engineering, economics).

Materials

Each participant solved three elementary probability word problems, representing three different probability topics: combinations, independent events, and conditional probability. The first two problems/topics were the target materials for this study, because they admit of two distinct salient solution strategies: the outcome-search strategy and the sequential-sampling strategy (described in the literature review and below). The third problem, involving conditional probabilities (and referred to below as the Weather problem), did not invoke alternative solution strategies, and was treated merely as a filler problem for purposes of this investigation.

The problem text for the independent events problem (also called the Spinner problem) was:

> *Two spinners are constructed. Each spinner has 3 color sections of equal size: red, white, and blue. The two spinners are spun at the same time, and the result of each spinner is recorded. What is the probability of getting the same color on both spinners?*

In the diagram conditions, either a tree (Figure 6) or a table (Figure 7) was provided to the problem solver in addition to the problem text. These diagrams were unlabeled, but the number of branches (or rows and columns) was appropriate to the problem.

The problem text for the combinations problem (also called the Work-Group problem) was as follows. The generic diagrams for this problem are presented in Figure 8 and Figure 9.

> *Five students are in a work group. The teacher randomly selects two of them to present the group work. If there are 2 boys and 3 girls in this group, what is the probability that the teacher selects 2 girls?*

Although these two problems represent two different probability topics, the same two broad strategies can be used to solve each one. For the combinations (Work-Group) problem, the outcome-search strategy defines the probability as 3/10: 3 possible ways of selecting two girls over 10 possible ways of selecting any two students, from a group of two boys and three girls. The sequential-sampling strategy finds the probability as $P(G_1 \cap G_2) = P(G_1) \times P(G_2|G_1) = (3/5) \times (2/4) = 3/10$. Similarly, for the independent events (Spinner) problem, the probability defined by the outcome-search strategy is 3/9 (or 1/3): 3 total outcomes that both spinners land on the same color over 9 possible color combinations between the two spinners. Using the sequential-sampling strategy, the probability that both spinners land on a particular color (e.g., red) is $P(R_1 \cap R_2) = P(R_1) \times P(R_2) = (1/3) \times (1/3) = 1/9$. And because the spinners can land on any of the three color sections, the total probability is three times the probability of obtaining a particular color twice, equal to $3 \times (1/3) \times (1/3) = 3 \times 1/9 = 3/9$ (or 1/3).

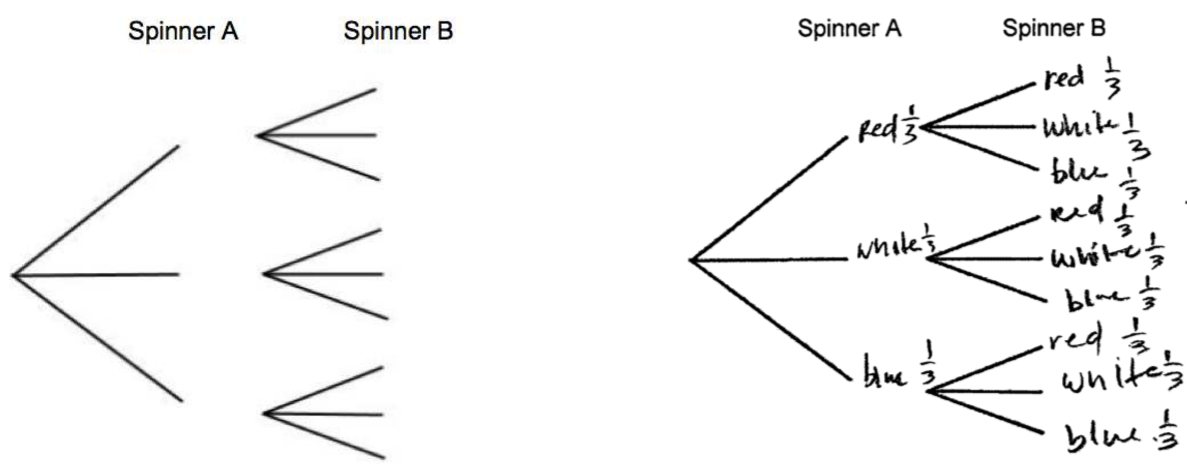


Figure 6. Tree for the Spinner problem. The panel on left shows the unlabeled tree provided with the Spinner problem in Study 1. The panel on right shows how it was annotated by one participant.

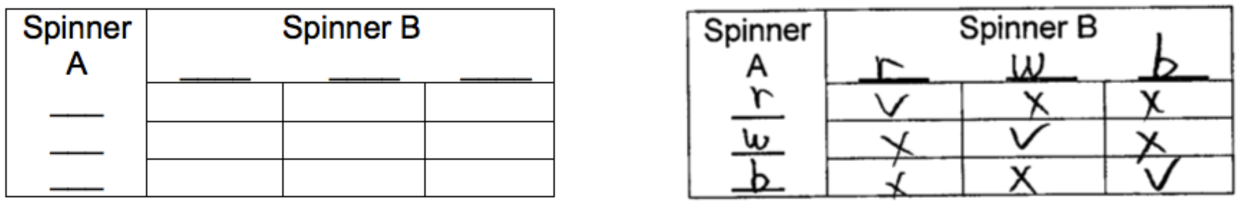


Figure 7. Table for the Spinner problem. The panel on left shows the unlabeled table provided with the Spinner problem in Study 1. The panel on right shows how it was annotated by one participant.

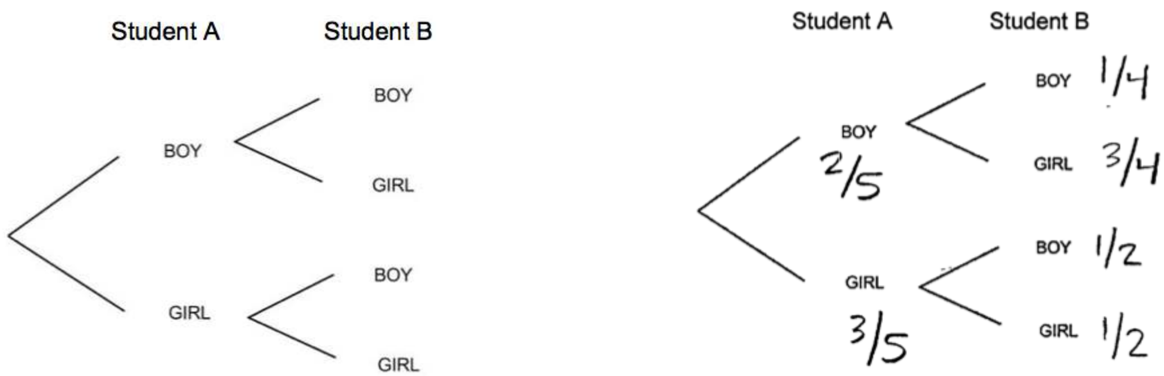


Figure 8. Tree for the Work-Group problem. The panel on left shows the unlabeled tree provided with the Work-Group problem in Study 1. The panel on right shows how it was annotated by one participant.

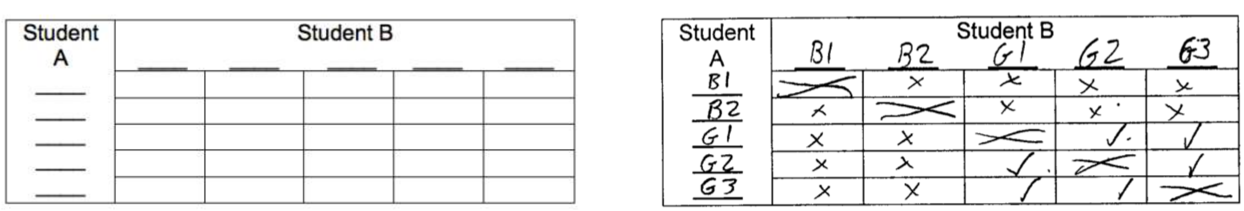


Figure 9. Table for the Work-Group problem. The panel on left shows the unlabeled table provided with the Work-Group problem. The panel on right shows how it was annotated by one participant.

Design and Procedure

Three test forms were used, each presenting the problems in a different order (Table 2). The first problem for each test form was given in text only, with no provided diagram. For the second and the third problems, one was given with a generic tree diagram, and the other one was given with a generic table diagram. Problems were presented in different orders, counterbalanced to equate possible carry-over effects. Thus, each problem was attempted by three independent groups of participants, with one third of them solving it with no provided diagram, one-third with a tree diagram, and one-third with a table diagram. As Table 2 shows, three types of provided diagrams were used: an N-by-N (outcome) table and a binary tree for the Work-Group (combinations) problem; and an N-by-N (outcome) table and an N-ary (outcome) tree for the Spinner (independent events) problem.

Table 2
Study Design and Test Forms for Study 1

	Problem 1	Problem 2	Problem 3
Form A	Spinner (no diagram)	Weather (binary tree)	Work-Group (N-by-N table)
Form B	Weather (no diagram)	Work-Group (binary tree)	Spinner (N-by-N table)
Form C	Work-Group (no diagram)	Weather (binary table)	Spinner (N-ary tree)

Participants (N=48) were randomly assigned to the three test forms, with 16 participants in each test form. Participants were tested individually in a laboratory setting, with no interaction with an on-site experimenter, although they may ask the experimenter questions for clarifying the task instructions during the experiment. In the task, each participant was given a booklet in which each problem was presented on a separate page. To prevent participants from seeing more than one diagram at a time, they were instructed not to look at any other problems when they were working on a problem. Participants were also asked to show their step-by-step solution procedure on the worksheet. Participants were explicitly instructed to use the provided graph for problem solving when a problem was accompanied with a diagram. A probability formula sheet was also provided, although participants were told that the formula sheet was optional for them to use. Participants also filled out a brief survey about their basic demographic and statistics training experience. At the end of the experiment, most participants were each paid eight dollars, and the rest of the participants received course credits as an optional way to fulfill a course requirement.

Results

Coding

Two sets of outcome variables were of focal interest: choice of a solution strategy, and problem solving correctness. Strategy types were coded based on whether an outcome-search strategy or a sequential-sampling strategy was used for problem solving. Use of each strategy was coded independently and dichotomously, with a value of 1 if the strategy was used (including a partial or informal one as long as its key feature was substantially reflected in the strategy, e.g., outcome counts or a stage-wise process), and 0 otherwise. This strategy determination was made without reference to correctness, which was coded independently. Examples of participants' work using each of the two solution strategies for the two problems are shown in Figure 10 (the Work-Group problem) and Figure 11 (the Spinner problem).

Solution success was assessed with two measures. *Answer correctness* was coded dichotomously, based on whether the final answer has the correct value, regardless of whether all the solution steps were correct. *Procedural correctness* was coded to measure if the solution steps a solver followed were appropriate, regardless of final answer correctness. For example, if a solver's solution procedure was correct, and the only error was a computational error that led to an incorrect final answer, his/her *procedure* would be coded as correct, but *answer correctness* would be coded as incorrect. On the other hand, one participant obtained a correct final answer by an incorrect procedure. For these reasons, procedural correctness is analyzed as our main measure of problem solving success (cf. Gugga & Corter, 2014).

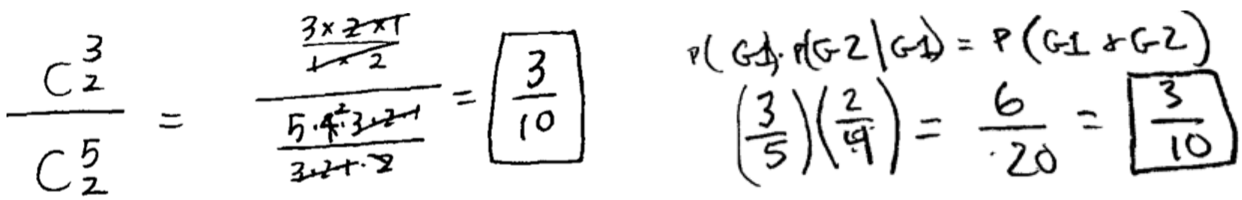


Figure 10. Two solution strategies for the Work-Group problem. The panel on left shows an example of a participant using the outcome-search strategy for the Work-Group problem. The panel on right shows an example of a participant using the sequential-sampling strategy for the Work-Group problem.

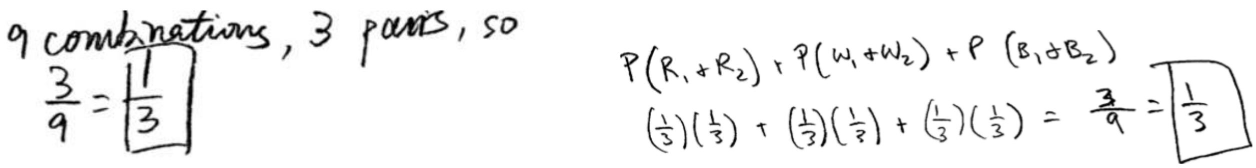


Figure 11. Two solution strategies for the Spinner problem. The panel on left shows an example of a participant using the outcome-search strategy for the Spinner problem. The panel on right shows an example of a participant using the sequential-sampling strategy for the Spinner problem.

To check reliability, a second coder independently coded 24 (or 25%) of the problem solutions in Study 1. For solution strategy coding, the inter-rater reliability was well established, as percent agreement between the two coders was 0.96, with $r = 0.92$. For solution procedural correctness coding, the inter-rater agreement was perfectly met (=1.0) between the two coders.

Statistics Training Experience by Test Form

The numbers of statistics courses taken by participants across the three test forms were compared. No statistical difference was found across the forms (Form A: M=2.25, SD=1.571; Form B: M=2.88, SD=2.029; Form C: M=1.88, SD=0.885). However, participants taking Form C seemed to have had less statistics training experience than participants taking Form B, $t(30) = 1.807$, $p = 0.081$ (without Bonferroni's correction in order to detect any possible group difference in prior statistics training experience).

Analysis of the Work-Group (Combinations) Problem: Effects of the N-by-N Table versus the Binary Tree

Strategy choice. As hypothesized, solvers' choices of the solution strategies were strongly biased by diagram types. Table 3 presents the frequency distributions of solvers' strategy choices. Figure 12 shows the difference across diagram types. Because use of each strategy was coded independently, neither or both of the two strategies might be employed on a given problem by a participant. Therefore, a participant's solution may be coded as using both the outcome-search strategy and the sequential-sampling strategy. For that reason, the choices of the solutions were not mutually exclusive, and the total number of strategies used by a diagram group (shown in Table 3) may exceed the number of its participants.

Participants given the N-by-N table more frequently used the outcome-search strategy (93.8%) compared to participants given no diagram (62.5%), $\chi^2(1, N=32) = 4.571$, $p = 0.083$ (by Fisher's Exact Test); and less frequently used the sequential-sampling strategy compared to no diagram (6.3% vs. 43.8%), $\chi^2(1, N=32) = 6.000$, $p = 0.037$ (by Fisher's Exact Test). On the other hand, participants given the binary tree showed significantly more use of the sequential-sampling strategy than those given no diagram (81.3% vs. 43.8%), $\chi^2(1, N=32) = 4.800$, $p = 0.028$; but significantly less use of the outcome-search strategy than those given no diagram (25% vs. 62.5%), $\chi^2(1, N=32) = 4.571$, $p = 0.033$.

Table 3
Strategy Choices and Correctness for the Work-Group Problem by Diagram Type

Diagram Condition	Strategy: Outcome search	Strategy: Sequential sampling	Correctness: Procedural correctness	Correctness: Answer correctness
No diagram (N=16)	10 (62.5%)	7 (43.8%)	13 (81.3%)	12 (75.0%)
Binary tree (N=16)	4 (25.0%)	13 (81.3%)	13 (81.3%)	11 (68.8%)
N-by-N table (N=16)	15 (93.8 %)	1 (6.3%)	5 (31.3%)	5 (31.3%)
Overall (N=48)	29 (60.4%)	21 (43.8%)	31 (64.6%)	28 (58.3%)

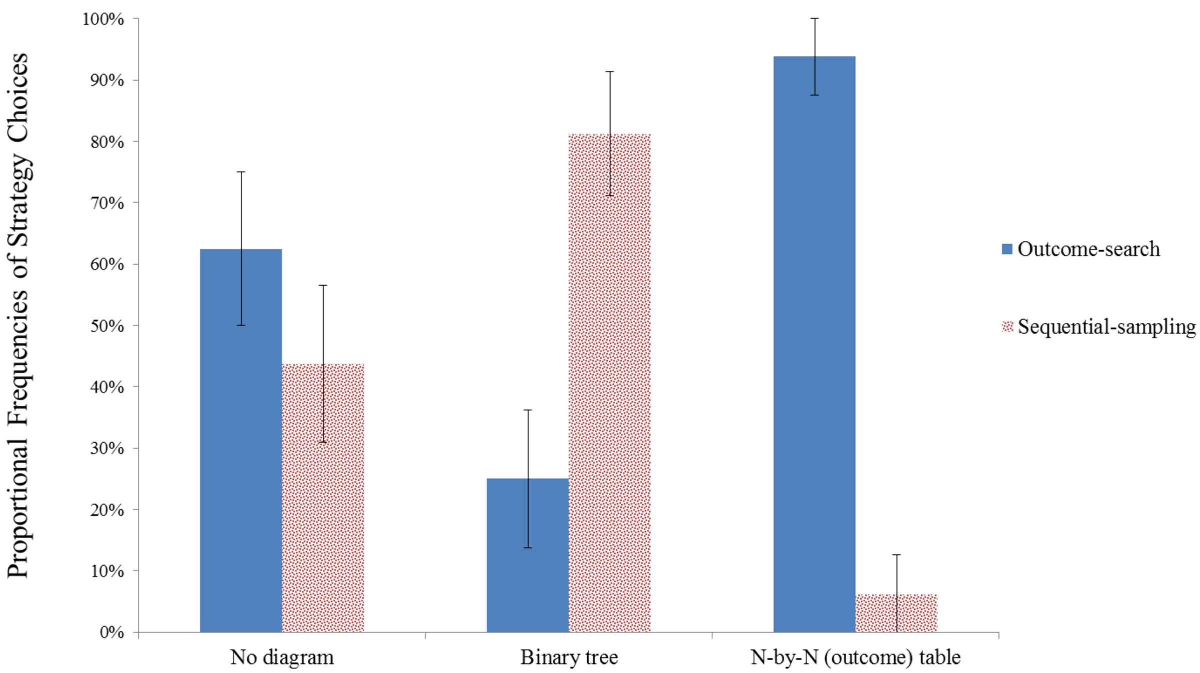


Figure 12. Frequencies of strategy choices for the Work-Group problem by diagram type. Error bars represent standard errors.

Solution Success. Table 3 also shows participants' rates of procedural correctness and answer correctness. Figure 13 shows the difference across diagram types. The procedural correctness rates for the no-diagram condition (81.3%) and the binary tree condition (81.3%) were high and identical. However, procedural correctness for the N-by-N table condition was significantly lower (31.3%) than the no-diagram condition (81.3%), $\chi^2(1, N=32) = 8.127$, $p = 0.004$.

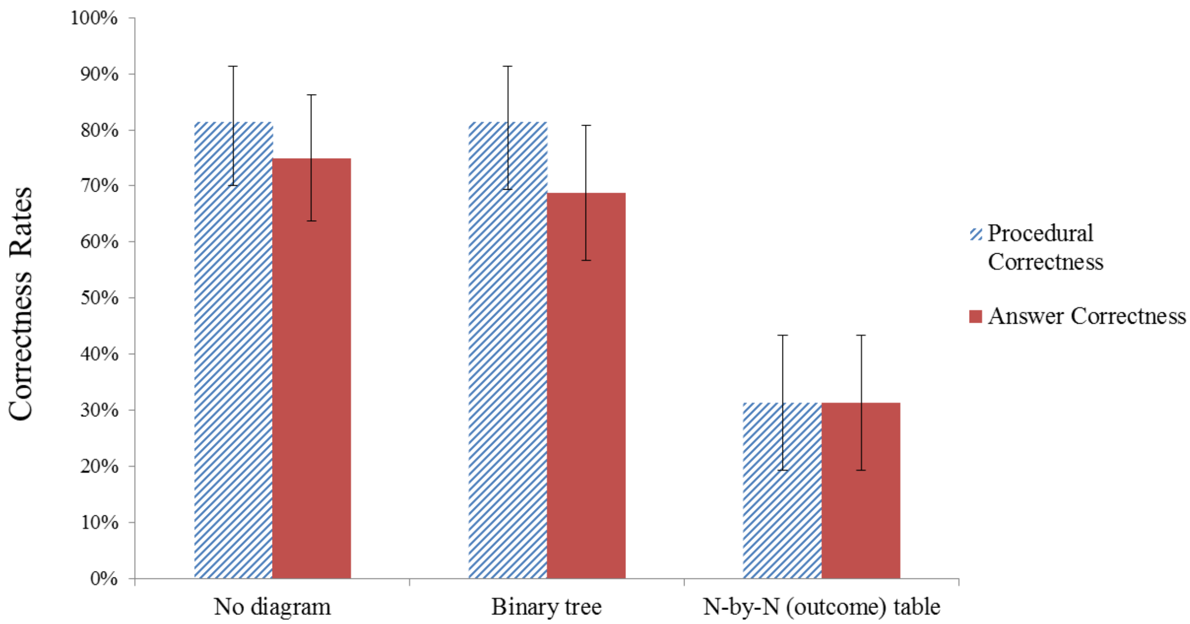


Figure 13. Correctness rates for the Work-Group problem by diagram type. Error bars represent standard errors.

Error Analysis. Unsuccessful solvers' error patterns were analyzed to better understand why the table decreased problem solving success for this problem. When participants were given the N-by-N table for the Work-Group problem, 6 of the 11 erroneous solutions resulted from incorrectly defining the total number of equally-likely outcomes, and 5 of these cases resulted directly from counting repeated selection of a single student (represented by the diagonal cells) as valid combinations. Thus, it was a common error to count all 25 cells as valid outcomes by using the 5-by-5 table (Figure 14). This corresponds to treating the selection of two (distinct) students as sampling with replacement. Another 3 errors involved incorrect use of the combinations formula, or failure to convert the count of outcomes to probability.

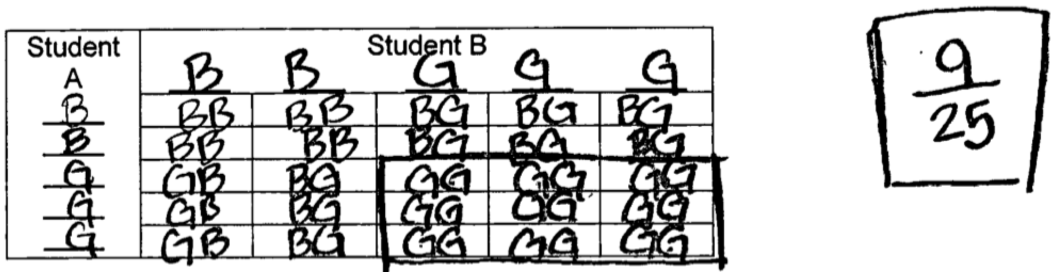


Figure 14. A common misuse of the N-by-N table. Counting diagonal cells (=self-repeated combinations) as valid outcomes of a combinations problem was a common misuse of the table that led to erroneous solutions for the Work-Group problem.

Thus, more than 80% of the procedural errors made with the N-by-N table involved the solver attempting to use the outcome-search strategy but being led astray by the structure of the N-by-N table. Specifically, in order to use the table correctly, problem solvers must recognize that the diagonal cells should not be used, because the self-repeated combinations that these cells represent are impossible outcomes when sampling without replacement. Put another way, the structure of the table does not map in a one-to-one way with the structure of the combinations problem.

The sequential-sampling strategy evoked by the binary tree led to distinctively different error types, so that three out of five incorrect answers involved incorrectly defining stage-wise probabilities. For example, the correct probabilities for the two sequential selections should be $P(G_1) = 3/5$ and $P(G_2|G_1) = 2/4$. However, the erroneous solutions involved incorrect probability values for these two events.

Discussion. For the Work-Group problem, the tree and the table diagrams altered the frequency of using particular strategies, with the N-by-N table leading solvers to select an outcome-based strategy (the "outcome-search" strategy), and the binary tree leading them to select a sequential strategy based on an event-level representation (the "sequential-sampling" strategy). These differences in turn led to differential error rates for the diagram conditions, and also to characteristic error patterns that are highly distinctive. In particular, the N-by-N table led to more use of an outcome-based strategy, and more errors in identifying the correct equally-likely outcome space; and the binary tree led to more use of a sequential strategy, and typically errors in identifying the correct stage-wise probabilities for unequally-likely outcomes of the selections.

These effects are explained in terms of the compatibility between the diagram and the relevant problem characteristics. Combinations problems (e.g., "How many ways can *N* objects be selected *k* at a time?") are typically interpreted as involving the simultaneous sampling of *k* entities from a larger set of *N* entities (e.g., by application of the formula for the number of combinations of *N* objects selected *k* at a time), but can also be formulated and solved as involving sequential sampling (*k* draws, a single object at a time, without replacement). However, in the latter case, order of selection is implied to be relevant, so answers may require appropriate adjustment.

The N-by-N table displays the outcomes simultaneously in an outcome space, by integrating all outcome cells into a single matrix structure. Therefore, it cues solvers to search for all possible outcomes in the whole outcome space and for the target event as a subset embedded in the whole outcome space on the table. Note specifically that an unlabeled N-by-N table implicitly cues solvers to consider all outcome cells as relevant to the problem, including those diagonal cells that represent self-repeated selections, although these are impossible outcomes for the combinations problem.

Thus, the table led to more erroneous identifications of the correct outcome space, because its representational format violates the Principle of Congruence articulated by Tversky, Morrison and Betrancourt (2002), in that there is poor fit between the structure of the diagram and the structure of the problem. Tversky, Morrison and Betrancourt (2002) suggest that in order to facilitate information processing, the structure and content of a diagram must be compatible with the structure and content of the represented information.

In contrast, the binary tree steers problem solvers towards use of an outcome space with only four unequally-likely outcomes: S = {BB, BG, GB, GG}. In addition, the tree diagram's left-to-right hierarchical structure may implicitly cue viewing the problem as involving sequential sampling without replacement. Therefore, the binary tree steers problem solvers towards a sequential-sampling strategy. Correct execution of this strategy requires correctly specifying conditional probabilities for the second student selected, as in Figure 10. Here, errors made by the binary tree condition tended to involve incorrect specification of these probabilities.

Analysis of the Spinner (Independent Events) Problem: Effects of the N-by-N Table versus the N-ary Tree

The two diagrams contrasted for the Work-Group problem actually vary by two aspects of the diagram at once: the general structure type (tree versus table), and the abstraction level of outcome space (an equally-likely outcome space based on the specific students selected versus an unequally-likely outcome space based only on sex of the two selected students). For the Spinner problem, exemplifying the use of the fundamental principle of combinatorics, the study chose to control for the abstraction level of outcome space and vary only the diagram type: tree versus table. This problem is referred to as the "independent events" problem, because the outcome of the first spinner is independent of the outcome of the second spinner.

Strategy Choice. As for the Work-Group problem, use of each strategy was coded independently for the Spinner problem. Therefore, neither or both of the strategies might be used in a participant's solution, so the choices of the strategies by a diagram group were not mutually exclusive (shown in Table 4). As shown in Table 4 and Figure 15, the N-by-N table diagram led to more frequent use of the outcome-search strategy (to 68.8%), compared to 31.3% when no diagram was provided, $\chi^2(1, N=32) = 4.500$, $p = 0.034$. However, the N-ary tree, which cues both use of the equally-likely outcome space and a sequential order of defining stage-wise probabilities, led to mixed choices of strategies: 50% of participants in the N-ary tree condition used the outcome-search strategy, whereas 50% of them used the sequential-sampling approach.

Table 4
Strategy Choices and Correctness for the Spinner Problem by Diagram Type

Diagram Condition	Strategy:		Correctness:	
	Outcome search	Sequential sampling	Procedural correctness	Answer correctness
No diagram (N=16)	5 (31.3%)	9 (56.3%)	8 (50%)	8 (50%)
N-ary tree (N=16)	8 (50%)	8 (50%)	10 (62.5%)	11 (68.8%)
N-by-N table (N=16)	11 (68.8%)	5 (31.3%)	14 (87.5%)	14 (87.5%)
Overall (N=48)	24 (50%)	22 (45.8%)	32 (66.7%)	33 (68.8%)

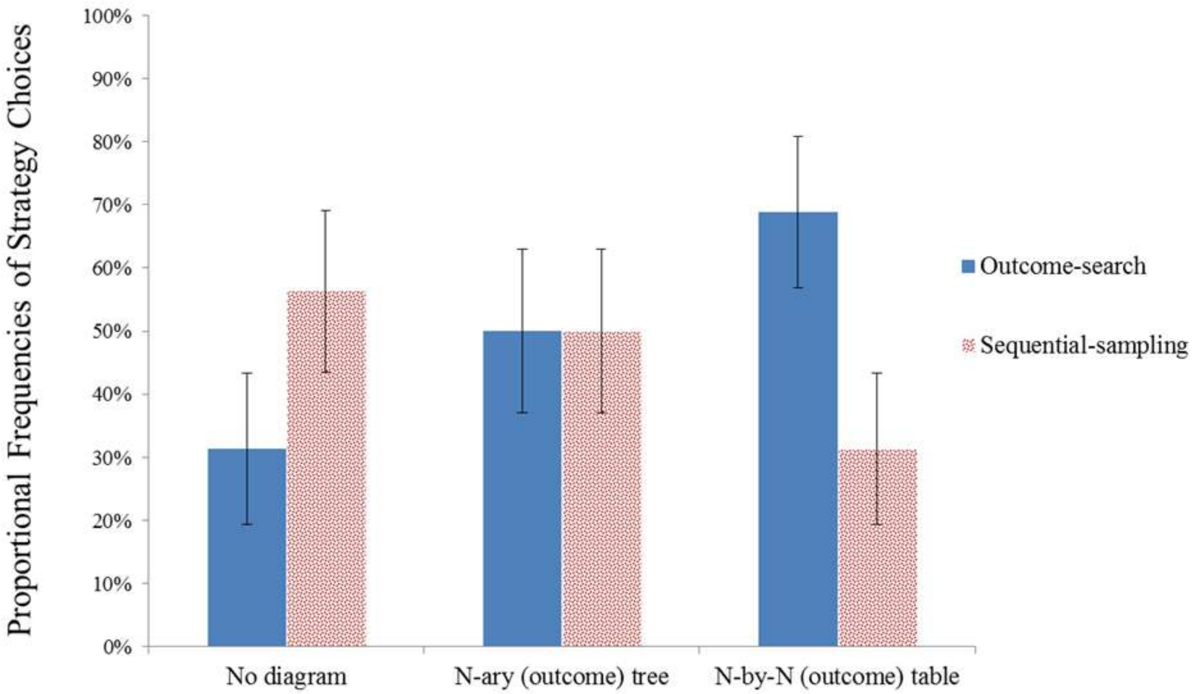


Figure 15. Frequencies of strategy choices for the Spinner problem by diagram type. Error bars represent standard errors.

Solution Success. Table 4 also shows the frequency and percentage of participants in each condition who successfully solved the Spinner problem. The difference is also shown in Figure 16. Compared to the no-diagram condition (50% procedural correctness rate), both diagrams increased the percentage of procedurally correct solutions. The increase of the procedural correctness rate, to 62.5%, was not significant for the N-ary tree, $\chi^2(1, N=32) = 0.508$, $p = 0.476$. However, the increase to 87.5% was significant for the N-by-N table, $\chi^2(1, N=32) = 5.236$, $p = 0.022$.

Error Analysis. For the Spinner problem, the most common error was to calculate the probability of obtaining one particular color twice, instead of *any* of the three colors. Specifically, this error involved finding the probability for the two spinners to land on the same color to be $(1/3) \times (1/3) = 1/9$ for only one color, and stopping there. However, the correct solution should be three times this probability (=1/3), because there are three colors on each spinner, and hence three same-color outcomes. In other words, solvers making this error failed to solve the problem completely due to failure to integrate intermediate results and all possible outcomes. This type of error and other types of errors were most likely to occur with no diagram, or with the tree. On the other hand, the N-by-N table, by displaying all possible outcomes for the independent events problem in a visually efficient way, facilitated the search for complete outcome information, and the coordination of the sub-stages of the problem solving, and thus improved the success rate.

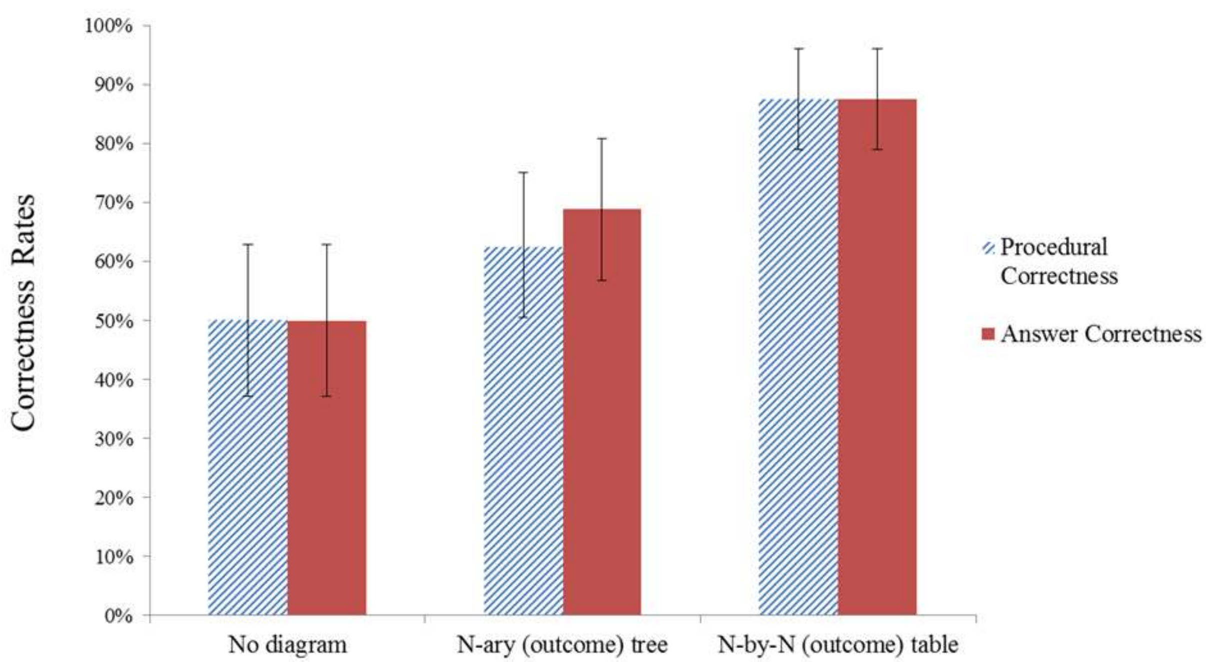


Figure 16. Correctness rates for the Spinner problem by diagram type. Error bars represent standard errors.

Discussion. For the Spinner problem, the N-by-N table was the most effective representation, improving procedural correctness over the no-diagram condition. This is not surprising, since the table represents the N×N = 3×3 = 9 equally likely outcomes in a simple and direct way, even allowing space for labeling the 9 outcomes. Furthermore, the table representation naturally suggests the semantic aspects of the problem: that there are two different spinners that are of equal status or priority (corresponding to the row and column of the table) (cf. Novick & Hurley, 2001). These findings confirm that the more compatible the structure and content of a diagram is to that of the represented problem, the more facilitative it is for solving the problem.

The N-ary tree would also seem to offer advantages for this Spinner problem: it too displays the nine equally-likely outcomes with roughly equal salience, allowing space for convenient labeling. However, the tree's hierarchical structure suggests a sequential process, and here it is not explicitly stated whether the spinners are spun simultaneously or sequentially. Also, the tree does not distinguish the target same-color outcomes for this problem to the same degree as the table; the table places these same-color outcomes on the main diagonal, where they are particularly prominent and grouped, after a fashion. In this way, the N-by-N table was able to provide extra external visual support for solvers to manage all the possible outcomes by chunking the essential information for computational efficiency (Gattis & Holyoak, 1996; Larkin & Simon, 1987).

CHAPTER IV: STUDY 2

Study 1 tested the effect of three different types of tree and table diagrams on solvers' choice of solution strategy and solution correctness for solving two problems involving different probability topics. The study found evidence to support the hypothesis that different types of diagrams direct solvers to notice different aspects of a problem structure, in some cases invoking different formal probability concepts and thus steering them to construct different solutions. The solution accuracy and error analyses for each task problem in Study 1 confirmed the hypothesis that the degree of the structural and content compatibility between a diagrammatic representation and its represented problem affects success in problem solving. When the entities and relations described in a probability word problem can be accurately matched to the visuospatial components of a diagram, and when an appropriate diagram organizes the problem information in an efficient way, the use of a diagram improves problem solving accuracy or maintains it if it is already high. On the other hand, a lower degree of visual-conceptual mapping correspondence may lead to misinterpretation of the underlying concepts and structure of a probability problem, and thus fails to improve or even impairs problem solving correctness.

However, there are two limitations in the design of Study 1. First, the two diagrammatic features, basic diagram structure types (tree versus table) and levels of information abstraction (outcome- versus event-based) were not systematically manipulated in Study 1. A 2-by-2 factorial design is needed to pinpoint which feature(s) of these diagrams accounts for the observed differences in solvers' strategy choices and solution correctness rates. And additional investigation is needed to understand how exactly the four types of diagrams that can be generated by crossing these two factors differ in their impact on probability problem solving.

Second, it remains unclear whether the impact of these different diagram types on strategy choice and solution correctness can be generalized across probability topics. For example, the effect of an N-ary (outcome) tree was only tested on an independent events problem (the Spinner problem). With no diagram, the Spinner problem was solved with the outcome-search strategy by 31.3% of the solvers and with the sequential-sampling strategy by 56.3% of the solvers. This may raise alternative explanations to such a diagram led to evenly split choices of the two strategies. For example, it might be argued that the sequential-sampling strategy was chosen by 50% of the participants with an N-ary tree just due to a random effect. If it is indeed the diagram effects that account for solvers' strategy choices, the kinds of strategy choice patterns that were found on the Work-Group and the Spinner problems should be found similar on other probability problems, too. To confirm the diagram effects found in Study 1, a replication with a wider variety of problems is needed.

Study 2 was designed to resolve these two issues. This study aimed to address two main research questions:

First, how do different features of the table and tree diagrams affect solvers' choice of solution strategy and their solution correctness for probability problem solving? Specifically, is a certain type of strategy more likely to be cued because of a diagram's general structure (tree or table), or its level of information abstraction (outcome- or event-based), or the combination of both features?

Second, do the diagram effects found on these individual problems in Study 1 hold up in the face of varying probability topics and problem types? In other words, is a certain type of solution strategy for a probability problem more likely to be chosen because of its diagrammatic representation, regardless of the probability topic or the semantic schema that a problem may evoke?

In Study 2, a set of four problems that represent two probability topics, combinations and independent events, were given to each participant. Five independent conditions were created so that each participant solved all the task problems with only one type of diagram, or no diagram, exclusively. This design of Study 2 disentangled confounding diagram features, and also made possible testing the generalizability of the effect of different diagrams across a wider variety of problems.

The design and results of Study 2 are presented in this chapter, along with a discussion of the findings and implications.

Method

Participants

One hundred and ten college or graduate school students (91 or 82.7% female) were recruited from a university in New York City. Their average age was 22.90 years (*SD* = 3.76 years). To be qualified for this study, a participant had to have taken one or more statistics or mathematics courses covering probability materials at the high school level or above prior to participation (including AP courses). Participants were randomly assigned to one of the five conditions: no diagram (N=22), binary tree (N=24), N-ary (outcome) tree (N=20), binary table (N=21), and N-by-N (outcome) table (N=23).

Materials

Two test forms (A and B) were used; each one presented four problems that represent two different probability topics and four semantic schemas (two for each probability topic): combinations (semantic schemas: simultaneous sampling; sequential sampling without replacement); and independent events (semantic schemas: simultaneous sampling (or matching); sequential sampling with replacement).

The semantic schema of a problem is not a formal problem schema. A semantic schema is the interpreted structure that solvers induce based on the inferred relation of the specific entities described in the problem content (Bassok, Wu, & Olseth, 1995). Different semantic schemas can influence probability problem solvers to interpret probability problem structures differently and choose different solutions for problems that share the same underlying probability topic and solution (Bassok, Wu, & Olseth, 1995). Therefore, two different semantic schemas were included for each probability topic in the test materials, because they allowed testing whether the diagram effects on probability problem solving hold up in the face of content variations that may also potentially influence problem solution strategy choices.

For combinations, a problem with a *simultaneous sampling* schema describes combinatorics as selecting a subset of k entities from N objects at a time. A problem with a *sequential sampling without replacement* schema describes combinatorics as selecting *one* entity at a time from N objects, for k times without replacement.

For the fundamental principles of independent events, a problem with a *simultaneous sampling (or matching)* schema describes combinatorics as matching up entities from two or N independent sets. A problem with a *sequential sampling with replacement* schema describes combinatorics as selecting *one* entity at a time from N objects, for k times with replacement.

The problems differed by cover content between the test forms (A and B), but were parallel by formal principles and solutions (shown in Appendix B). For example, the combinations problem with a simultaneous sampling schema in Form A asks for the probability of selecting two nickels out of five coins; and its counterpart in Form B asks for the probability of selecting two girls out of five students.

Design and Procedure

Five conditions were created: the binary-tree condition, the N-ary (outcome) tree condition, the binary-table condition, the N-by-N (outcome) table condition, and the no-diagram control condition. In the no-diagram condition, problems were given in text only. For each of the four diagram conditions, a particular type of diagram was provided exclusively for every task problem. For example, every problem in the binary-tree condition was accompanied with a generic binary tree diagram. Participants were randomly assigned to a condition and a test form, and asked to solve all four probability word problems that were presented in a booklet.

To minimize any memory or carry-over effects of one task problem on another, three filler problems were included in the test booklet to be interspersed with the task problems. The fillers represent other probability topics and irrelevant solutions such as joint probability. In between every two task problems, a filler problem was placed to distance them, so that participants did not solve any task problems back to back.

The task problems were randomly permuted into positional slots 1, 3, 5, and 7, with the constraint that each problem occurred approximately equally often in the first position. The filler problems were randomly permuted into positional slots 2, 4, and 6. Participants were randomly assigned to the five conditions and tested individually in a laboratory setting. The same task procedure and instructions were followed as in Study 1. Following the problem solving task, a brief demographic survey was administered to collect basic background information of the participants, such as undergraduate majors, and post-secondary statistics and mathematics training experience. As compensation for their participation in this approximately 30-minute long experiment, participants were paid $8 or given course credit (for those registered in eligible courses) at their choice.

Results

Coding and Scoring

For each problem, the solution strategy type a participant chose, the solution procedural correctness, and the final answer correctness were coded following the same coding scheme used in Study 1. Because inter-rater reliability for coding solution strategies and procedural correctness had already been established as adequate in Study 1, the problem solutions in Study 2 were coded by a single coder. For each participant, a total score was computed for the use of a strategy type as the measure of how frequently a certain type of solution strategy was used for overall problem solving. Following the coding scheme in Study 1, use of the two strategies was coded independently, so that a participant might use both strategies. Thus, for a group the proportional frequencies of use of the two strategies do not necessarily add up to 100%. A total score across problems was also computed for a participant’s overall solution procedural correctness and overall final answer correctness. Because each participant solved four problems, the total score for each measure (i.e., use of a certain strategy; procedural correctness; answer correctness) can possibly range from 0 to 4.

Mathematics Training Experience by Condition

Participants’ prior mathematics and statistics training experience and relevant academic training experience was measured.

To check whether participants in the five conditions had roughly equivalent amounts of post-secondary mathematics training experience, the total numbers of statistics and other mathematics courses taken at the college level or above were reported by the participants and compared across conditions (shown in Table 5). An ANOVA test suggests that there was no significant overall difference in the number of statistics or mathematics courses taken across the five conditions, $F(4, 105) = 1.830$, $p = 0.129$, for statistics courses; $F(4, 105) = 0.743$, $p = 0.565$, for other mathematics courses; and $F(4, 105) = 0.944$, $p = 0.442$, for all mathematics courses (including statistics). However, pair-wise comparison tests (conducted at the .05 level) showed that the average number of statistics courses taken by participants in the N-by-N (outcome) table condition ($M = 1.61$, $SD = 1.34$) was significantly lower than the no-diagram condition ($M = 3.091$, $SD = 2.81$), $t(43) = 2.27$, $p = 0.030$, and also lower than the binary-tree condition ($M = 3.00$, $SD = 2.78$), $t(45) = 2.17$, $p = 0.037$, suggesting some prior knowledge difference among the conditions. Please note that although multiple t tests were conducted, the pair-wise comparison p values were unadjusted for these tests for the purpose of detecting any possible group difference in prior mathematics training experience.

In addition to the total number of post-secondary mathematics courses taken, participants' academic background in terms of its relevance to mathematics training was measured by asking them to report their undergraduate majors. Overall, 22.7% of the participants reported to have non-STEM majors (e.g., philosophy, creative writing) or undeclared majors; 20% in social science (e.g., psychology); 18.2% in mathematics and/or statistics; and 39.1% in other STEM domains (e.g., natural science, engineering, economics). Table 5 also presents the distribution of undergraduate majors by condition. A chi-squared test of homogeneity showed no significant group differences in the distribution of undergraduate major background.

Another chi-squared test of homogeneity tested whether participants in Study 2 had equivalent relevant academic experience as in Study 1. Natural science, engineering, and mathematics majors may be exposed to more graphic and/or formal science and mathematics materials in their regular academic curriculum (e.g., Novick & Hurley, 2001). Participants in Study 2 showed a significantly different undergraduate major distribution (57.3% mathematics and other STEM majors, 22.7% non-STEM majors and 20% social science majors) from participants in Study 1 (25% mathematics and other STEM majors, 31.25% non-STEM majors and 43.75% social science majors), $\chi^2(2, 158) = 15.217$, $p < 0.001$. The results suggest that participants in Study 2 have a relatively high level of mathematics expertise for the probability problem solving task.

Table 5
Undergraduate Majors and Post-Secondary Mathematics Courses Taken by Participants in Study 2

	Undergraduate Major				Math courses taken	
	Non-STEM	Social Science	Math or Statistics	Other STEM	Statistics	Statistics & other math
No Hint (N=22)	7 (35%)	1 (5%)	5 (25%)	7 (35%)	3.091	6.227
Binary Tree (N=24)	2 (9.5%)	3 (14.3%)	7 (33.3%)	9 (42.9%)	3.000	7.292
N-ary Tree (N=20)	3 (15%)	6 (30%)	1 (5%)	10 (50%)	1.975	5.825
Binary Table (N=21)	5 (23.8%)	4 (19%)	4 (19%)	8 (38.1%)	2.357	6.667
N-by-N Table (N=23)	4 (17.4%)	7 (30.4%)	3 (13%)	9 (39.1%)	1.609	4.174
Overall (N=110)	25 (22.7%)	22 (20%)	20 (18.2%)	43 (39.1%)	2.418	6.041

Strategy Choices by Diagram

One of the main research questions was how different diagram types affect overall probability problem solving. To answer this question, the average frequencies of using a certain solution strategy (outcome-search or sequential-sampling) were compared across the conditions.

Table 6 presents the average frequencies of using a certain solution strategy across the conditions, and Figure 17 shows the group difference.

Table 6
Frequencies of Strategy Choices and Correctness for the Four Problems by Diagram Condition (Standard Deviations in Parentheses)

	Strategy:		Correctness:	
Diagram Condition	Outcome search	Sequential sampling	Procedural correctness	Answer correctness
No Diagram (N=22)	0.68 (0.84)	3.41 (0.73)	3.05 (1.05)	3.00 (1.07)
Binary Tree (N=24)	0.88 (1.23)	3.42 (1.14)	3.29 (0.96)	3.21 (1.02)
N-ary (outcome) Tree (N=20)	2.20 (1.67)	2.70 (1.59)	2.90 (1.41)	2.80 (1.51)
Binary Table (N=21)	0.86 (1.01)	3.38 (0.92)	3.29 (0.78)	3.19 (0.87)
N-by-N (outcome) Table (N=23)	2.83 (1.53)	1.52 (1.56)	3.00 (1.04)	2.83 (1.11)
Overall (N=110)	1.48 (1.54)	2.88 (1.43)	3.11 (1.05)	3.01 (1.12)

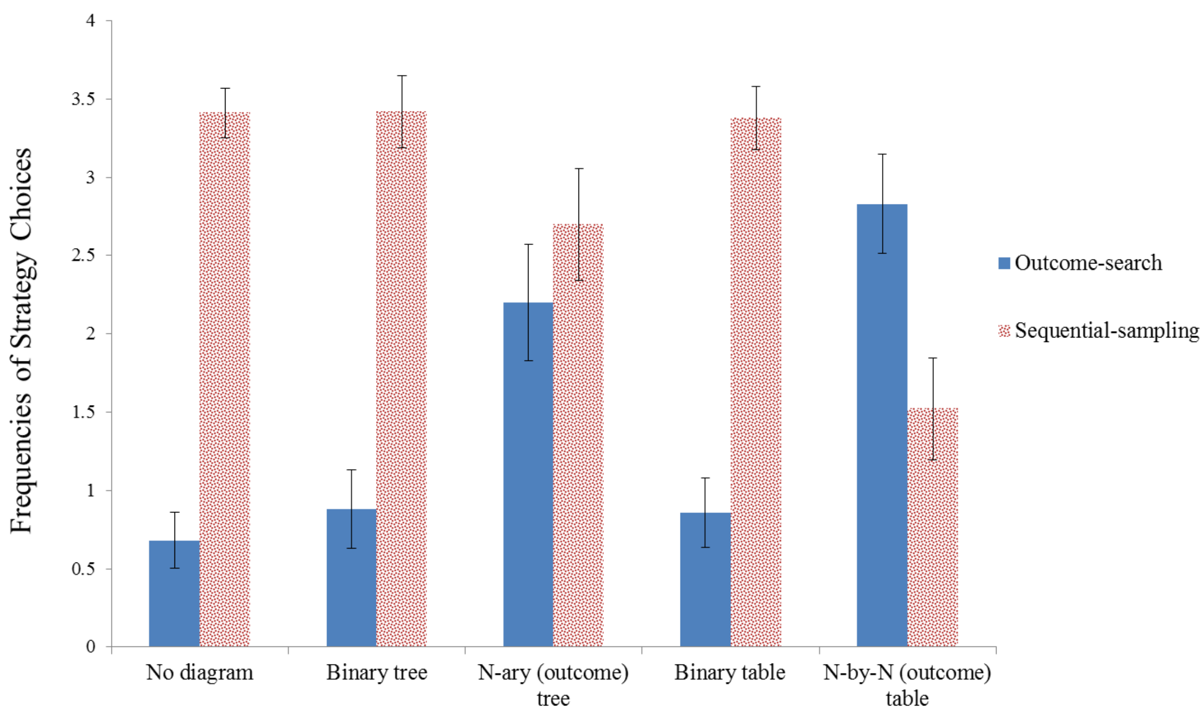


Figure 17. Frequencies of strategy choices for the four problems by condition in Study 2. Error bars represent standard errors.

Because the analyses of the participants' mathematics training background suggested that there might be some possible group differences in prior knowledge, which might affect the observed rates of solution success, the total number of the statistics and other mathematics courses that participants had taken was controlled as a covariate in the following analyses.

An ANCOVA test with participants' total number of mathematics courses taken as the covariate was used to test whether diagram conditions differed in their frequencies of using the outcome-search strategy for solving the task problems. The results showed significant differences in the frequencies of the outcome-search strategy use across the five conditions, $F(4, 104) = 13.432$, $p < 0.001$ (presented in Table 7).

Post-hoc tests with Bonferroni's correction were used to examine how exactly different types of diagrams led to this difference, compared to the no-diagram condition. Participants given an outcome-level representation (i.e., N-ary tree; N-by-N table) both used the outcome-search strategy more frequently than solvers who spontaneously solved the problems with no diagram, both p's < 0.002 (shown in Table 8). The two event-level representation conditions (i.e., binary tree; binary table) did not differ statistically in their frequencies of using this strategy than the no-diagram condition (shown in Table 8).

Table 7

Omnibus Test of Diagram Effects on the Use of the Outcome-Search Strategy

Source	SS	MS	*df*	*F*	*p*	η_p^2
Total statistics & mathematics courses taken (covariate)	5.048	5.048	1	3.099	0.081	0.029
Diagram conditions	87.528	21.882	4	13.432	< 0.001	0.341
Error	169.425	1.629	104			

Table 8
Specific Comparisons between the No-Diagram Condition and Each Diagram Condition on Use of the Outcome-Search Strategy

Condition	Adjusted *M* (*SD*)	*t*	*df*	*p*
No Diagram (N=22)	0.675 (1.276)	-	-	-
Binary Tree (N=24)	0.828 (1.283)	0.405	44	= 1.00
N-ary (outcome) Tree (N=20)	2.208 (1.277)	3.888	40	< 0.002
Binary Table (N=21)	0.834 (1.278)	0.407	41	= 1.00
N-by-N (outcome) Table (N=23)	2.896 (1.291)	5.803	43	< 0.001

Note. Multiple *t*-tests (with Bonferroni's correction) were used to compare the frequencies of using the outcome-search strategy between the no-diagram condition and each of the four diagram conditions. The frequencies for comparisons were the adjusted group means, adjusted for the covariate "total number of post-secondary statistics and mathematics courses taken".

Similarly, an ANCOVA test with participant's total number of mathematics courses taken as the covariate was conducted to test whether diagram conditions differed in their frequencies of using the sequential-sampling strategy for solving these task problems. Figure 17 shows the group differences in the average frequencies of using this solution strategy. The results (presented in Table 9) showed that the five diagram conditions differed significantly in their frequencies of using the sequential-sampling strategy, $F(4, 104) = 10.693$, $p < 0.001$.

Post-hoc tests with Bonferroni's correction suggested that although the two outcome-level representation conditions (N-ary tree; N-by-N table) both reduced the frequencies of using the sequential-sampling strategy compared to the no-diagram condition, only the N-by-N table condition led to a significant reduction of the use of this strategy compared to the no-diagram condition, $t(43) = 5.292$, $p < 0.001$ (shown in Table 10). The N-ary tree condition did not differ significantly from the no-diagram condition in the use of the sequential-sampling strategy, $t(40) = 1.901$, $p = 0.32$ (shown in Table 10). Furthermore, there was a significant difference in the frequencies of using the sequential-sampling strategy between the two outcome-level representation conditions (N-ary tree vs. N-by-N table). The N-ary tree condition (adjusted *M* =

2.693, *SD* = 1.228) demonstrated significantly more frequent use of the sequential-sampling strategy than the N-by-N table (adjusted $M = 1.465$, $SD = 1.242$), $t(41) = 3.251$, $p = 0.01$. The two event-level representation conditions (binary tree; binary table) had very high frequencies of using the sequential-sampling strategy, but these levels did not differ from the control condition (no diagram), perhaps due to a ceiling effect.

Table 9
Omnibus Test of Diagram Effects on the Use of the Sequential-Sampling Strategy

Source	SS	MS	*df*	*F*	*p*	η_p^2
Total statistics & mathematics courses taken (covariate)	3.222	3.222	1	2.136	0.147	0.020
Diagram conditions	64.499	16.125	4	10.693	< 0.001	0.291
Error	156.822	1.508	104			

Table 10
Specific Comparisons between the No-Diagram Condition and Each Diagram Condition on Use of the Sequential-Sampling Strategy

Condition	Adjusted *M* (*SD*)	*t*	*df*	*p*
No Diagram (N=22)	3.415 (1.228)	-	-	-
Binary Tree (N=24)	3.454 (1.234)	0.109	44	= 1.00
N-ary (outcome) Tree (N=20)	2.693 (1.228)	1.901	40	= 0.32
Binary Table (N=21)	3.400 (1.229)	0.040	41	= 1.00
N-by-N (outcome) Table (N=23)	1.465 (1.242)	5.292	43	< 0.001

Note. Multiple *t*-tests (with Bonferroni's correction) were used to compare the frequencies of using the sequential-sampling strategy between the no-diagram condition and each of the four diagram conditions. The frequencies for comparisons were the adjusted group means, adjusted for the covariate "total number of post-secondary statistics and mathematics courses taken".

Solution correctness by diagram

Table 6 presents the average frequencies of the solution procedural correctness and the final answer correctness for solving the problems by each condition (possible scores ranging from 0 to 4), and Figure 18 shows the group differences in these measures.

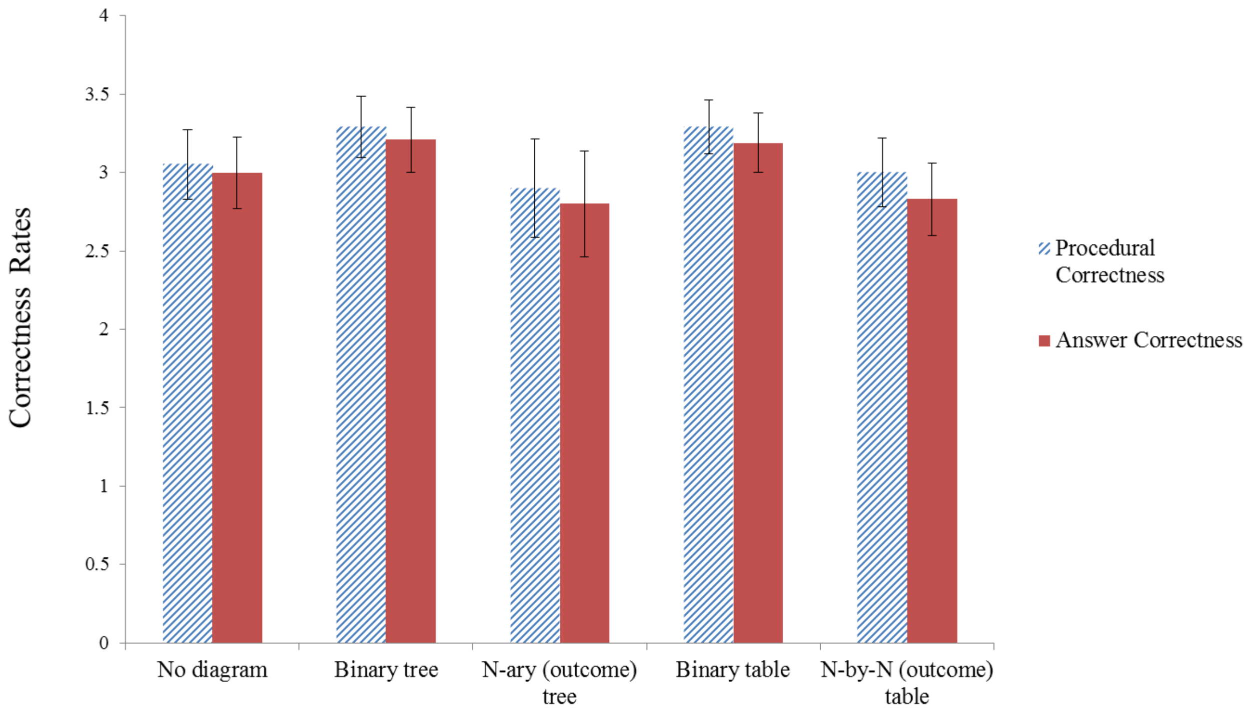


Figure 18. Correctness rates for the four problems by condition in Study 2. Error bars represent standard errors.

Solution procedural correctness was compared across the conditions, using an ANCOVA test with participant's total number of mathematics courses taken as the covariate. The results (presented in Table 11) showed little overall difference in solution procedural correctness frequency by condition, $F(4, 104) = 0.434$, $p = 0.784$, confirmed by the post-hoc test results (presented in Table 12) that found no difference in procedural correctness between any of the diagram conditions and the no-diagram condition. However, participants' prior statistics and mathematics training experience was significantly predictive of their solution procedural correctness, $F(1, 104) = 4.191$, $p = 0.043$. Thus, although some diagram conditions showed slightly higher solution correctness rates than some others (shown in Table 6), the difference may be due to differences in participants' prior mathematics knowledge (shown in Table 5). For example, the binary tree condition had a higher correctness frequency for problem solving (M = 3.29, SD = 0.96) than the N-ary tree condition (M = 2.90, SD = 1.41). However, participants in the binary tree condition also had taken more post-secondary statistics and mathematics courses (M = 7.292, SD = 6.779) than the N-ary tree condition (M = 5.825, SD = 6.003).

Table 11
Omnibus Test of Diagram Effects on the Procedural Correctness Rates

Source	SS	MS	df	F	p	η_p^2
Total statistics & mathematics courses taken (covariate)	4.571	4.571	1	4.191	0.043	0.039
Diagram conditions	1.895	0.474	4	0.434	0.784	0.016
Error	113.427	1.091	104			

Table 12
Specific Comparisons between the No-Diagram Condition and Each Diagram Condition on the Procedural Correctness Rates

Condition	Adjusted M (SD)	t	df	p
No Diagram (N=22)	3.039 (1.044)	-	-	-
Binary Tree (N=24)	3.247 (1.050)	0.673	44	= 1.00
N-ary (outcome) Tree (N=20)	2.908 (1.044)	0.406	40	= 1.00
Binary Table (N=21)	3.263 (1.046)	0.704	41	= 1.00
N-by-N (outcome) Table (N=23)	3.067 (1.056)	0.090	43	= 1.00

Note. Multiple t-tests (with Bonferroni's correction) were used to compare the procedural correctness rates between the no-diagram condition and each of the four diagram conditions. The procedural correctness rates for comparisons were the adjusted group means, adjusted for the covariate "total number of post-secondary statistics and mathematics courses taken".

Discussion

The results of Study 2 addressed two unresolved issues in Study 1, by pinpointing the effect of different diagrammatic features on probability problem solving, and confirming the effect on a wide variety of probability problems.

Study 2 examined systematically how diagram features (general structure types and levels of information abstraction) of the table and tree diagrams steered solvers' interpretations of probability problem structures and choice of solution strategy. The analyses of participants' strategy choices by condition suggested that both the basic structure type of a diagram (tree or table) and its level of information abstraction (outcome- or event-based) can bias solvers' choice of solution strategy for probability problems.

First, levels of information abstraction have an effect on the problem solution strategy choices, across different diagram structure types. As the results showed, when a diagram represents the problem outcomes at the more concrete equally-likely outcome level, it tends to cue the use of the outcome-search strategy. Compared to the no-diagram condition, an outcome-level representation, either an N-ary tree or an N-by-N table, significantly increased the use of the outcome-search strategy. These two diagrams also led to lower frequencies of using the sequential-sampling strategy, slightly by the N-ary tree and significantly by the N-by-N table. In contrast, when the problem information is represented at the more abstract unequally-likely event level (i.e., binary tree or binary table), solvers were almost always cued to use a sequential-sampling strategy and rarely an outcome-search strategy, regardless of a diagram's general structure type. No statistical difference was found in the use of the sequential-sampling solution strategy between these two diagram conditions and the no-diagram condition, perhaps due to a ceiling effect: The no-diagram condition already demonstrated a very high frequency of using the sequential-sampling strategy and little use of the outcome-search strategy. However, the drastic difference in solvers' strategy choices between the outcome-level representations and the event-level representations offered strong evidence that levels of information abstraction have a great impact on how a probability problem structure is interpreted and how the strategy choice is made accordingly.

Second, with an outcome-level representation, the tree structure influences problem structure interpretation and solution strategy choices in a remarkably different way from the table structure. With an N-by-N (outcome) table, an outcome-search strategy was most likely to be cued (M = 2.83 out of 4 in maximum, SD = 1.53). The N-ary (outcome) tree can also significantly increase the use of an outcome-search strategy compared to problem solving with no diagram (M = 2.20 out of 4 in maximum, SD = 1.67). However, its strength as a cue to the outcome-search strategy was not as strong as the N-by-N table. Even more interestingly, while an N-by-N table significantly reduced the use of the sequential-sampling strategy (M = 1.52 out of 4 in maximum, SD = 1.56), the N-ary tree still remains a relatively strong cue for the sequential-sampling strategy (M = 2.70 out of 4 in maximum, SD = 1.59). This difference between the N-ary tree and the N-by-N table was statistically significant, $t(41) = 3.251$ for the adjusted group mean difference, $p = 0.01$. The findings suggested that while both outcome-level diagrams can cue the search for equally-likely outcomes, a tree structure can additionally evoke sequential thinking, thus cueing an alternative interpretation of the probability problem structure, resulting in an almost equally-likely preference for choosing either solution strategy.

This difference can be explained by an in-depth structural comparison between the two diagrams. With an N-by-N (outcome) table, the diagrammatic structure not only represents all possible equally-likely outcomes of a problem, it also represents these outcomes in a spatially integrated fashion. Therefore, such a visual representation depicts the target outcomes as a grouped subset embedded in the whole outcome space, and can naturally cue solvers to perceive the target probability as part of the whole and thus to choose an outcome-search strategy for finding the probability. Therefore, an N-by-N table representation should cue an outcome-search strategy most naturally and strongly, compared to other diagram formats. On the other hand, an N-ary (outcome) tree structure leads to mixed strategy choices. It not only cues solvers to list all the equally-likely outcomes for a problem, it can also cue solvers to perceive the problem probability as the result of a sequential sampling procedure because of how it spatially organizes the outcomes. Specifically, its hierarchical structure naturally serves as a cue for solvers to define simple and conditional probabilities stage by stage, and to process the calculations towards the final probability in a sequential order.

Why does an N-ary (outcome) tree cue the outcome-search strategy less strongly than an N-by-N table, and the sequential-sampling strategy less strongly than the event-level representations, if its diagrammatic features are appropriate for cueing both types of probability problem structures and solution strategies?

Perhaps this can be explained by two reasons. The first and perhaps foremost reason is that most problem solvers chose to provide only one solution as long as it can solve the problem successfully. Therefore, if both strategies are cued, solvers might have randomly chosen one strategy or the other, leading to fewer demonstrations of either solution strategy compared to other representations that typically cue only one strategy.

An additional speculation is that although the N-ary tree is an appropriate visualization for both solution strategy types, the way it spatially organizes the problem information does not provide the information as efficiently for either strategy's computational procedure as the other diagram types. As reviewed, a tree diagram features hierarchical leveling in a sequential order, and the subsequent branches of each node represent all the possible outcomes conditional on the particular outcome from a previous selection. Therefore, although the target outcomes of these particular probability problems can be shown as a part embedded in the whole outcome space at the end level of the tree diagram, these outcomes are dispersed over the whole outcome space, not grouped. In other words, it is less efficient for solvers to complete the search of all the target outcomes represented at the end level of a tree than on an N-by-N (outcome) table. This may explain why an N-ary tree is less likely to cue the outcome-search strategy than an N-by-N table, although they both represent all equally-likely outcomes of the whole outcome space. Next, compared to the event-level representations, the N-ary tree is also less efficient for organizing the information so that it is ready for computations for a sequential-sampling strategy. A binary tree cues solvers to define binomial probabilities at each selection event, thus making the information immediately ready for computation. The N-ary tree cues solvers to first list out all equally-likely outcomes at each selection event. In order to use a sequential-sampling strategy, solvers need to transform this information into a probability value for each selection. In other words, the event-level representations represent stage-wise probabilities immediately at a symbolic level, while the N-ary tree involves an extra step of the iconic-to-symbolic mathematical transformation. Thus, this might explain why the sequential-sampling strategy is also less often chosen when an N-ary tree is provided, compared to the binary tree and the binary table.

CHAPTER V: GENERAL DISCUSSION

Summary

This dissertation explored the effect of diagrams on probability problem solving. Two studies were conducted to investigate how different types of diagrammatic representations of the problem information affect the process (choice of solution) and the product (solution correctness) in probability problem solving. Four different types of diagrams that varied by two diagrammatic features, the general diagrammatic types (tree versus table) and the levels of information abstraction (outcome-based versus event-based) were provided to different groups of problem solvers as the visual aids for solving a set of probability word problems. It was of particular interest to understand how different types of diagrams and their diagrammatic features influence solvers' interpretations of the problem structures and their choice of solution strategy for probability problems that admit of multiple types of solution strategies, and how that may affect solution correctness.

The results of Study 1 showed that different diagrams steer people to choose different solution strategies for probability word problems. Specifically, a binary tree steers people to choose a sequential-sampling strategy, an N-by-N (outcome) table is more likely to cue an outcome-search strategy, and with an N-ary (outcome) tree, the strategy choices tend to be mixed, with the two strategy types being chosen approximately equally often.

In addition, these diagrams affected problem solving success differently. For the combinations (Work-Group) problem, solvers had a higher rate of solution correctness when a binary tree diagram was provided, compared to when the N-by-N (outcome) table was provided. An error analysis revealed that different diagram types accounted for different patterns of characteristic errors. Given the N-by-N table, more than 80% of the solution errors involved incorrect implementation of the outcome-search strategy, such as an unsuccessful attempt at instantiating the combinations formula. Many such errors occurred because the solvers included all the self-repeated combinations (represented by the diagonal cells on the table) as the valid outcomes for the combinations problem. The N-by-N table led to a higher error rate and this type of solution error in particular because its diagrammatic components failed to map one-to-one with the problem it represented. For a combinations problem, it is impossible for an entity to be selected more than once, because sampling is inherently *without* replacement. Therefore, although an N-by-N table uses its row and column to represent all possible entities to be selected for the two times of sampling, all the diagonal cells represent those self-repeated combination outcomes that are impossible for a combinations problem.

Problem solving errors with the binary tree diagram showed a different characteristic error pattern. Here, 60% of the incorrect answers were due to incorrectly defining the stage-wise probabilities, which was a typical type of procedural error involved in the use of the sequential-sampling strategy. Solvers' solution success rate remained high with the binary tree, perhaps because of the high structural compatibility between the tree structure and the problem structure. The tree structure was appropriate for representing the procedure of sequential sampling for a combinations problem. Furthermore, its hierarchical structure may appropriately suggest that the probability of a selection outcome is dependent and conditional upon its previous selection.

For the independent events (Spinner) problem, a common type of solution error made by solvers with no diagram provided was to only calculate the probability of a particular outcome and stop there (e.g., when both spinners land on the red color). Both the N-by-N table and the N-ary tree were able to increase the correctness rates of problem solving. But only the N-by-N table significantly increased the correctness rate compared to no diagram provided. Again, the structural compatibility between a diagram and the problem can explain this difference. For the independent events problem, sampling is *with* replacement. Therefore, every intersection cell of the table represents a possible combination outcome for the problem. Furthermore, the table is able to spatially organize the problem information in a more efficient way, in that all the diagonal cells represent the preferred outcomes and all the other outcomes are represented off the diagonal on the table. Therefore, the N-by-N table was found to be more facilitative for solving the independent events problems.

However, Study 1 did not manipulate the two diagrammatic features (general diagram structure types and levels of information abstraction) systematically. In addition, different sets of diagrams were compared on different types of probability problems. Thus, a second study was conducted to resolve these issues: to systematically compare the diagram types in order to pinpoint what diagrammatic features account for solvers' choice of strategy and solution correctness; and to test whether the diagram effects hold up across a variety of probability problem types.

The results of Study 2 suggested that when a diagram represents probability at the unequally-likely event level, it steers people to choose the sequential-sampling strategy. This effect has been found with both a binary tree and a binary table. However, when a diagram represents probability at the equally-likely outcome level, the basic structure type of a diagram leads to different choices of solution strategies. Given an N-by-N (outcome) table, solvers showed a strong tendency of using the outcome-search strategy, and a significantly decreased frequency of using the sequential-sampling strategy. But given an N-ary (outcome) table, solvers showed relatively frequent use of both strategies.

Study 2 did not find large differences in the problem solving correctness rates across conditions. A ceiling effect might explain this. With an average of four to seven post-secondary mathematics courses taken by each condition, all conditions' procedural correctness scores were similarly high, from 2.90 to 3.29 out of 4 points as the maximum.

Implications

The study findings offer three important implications with regard to visual reasoning and diagram design in STEM education and problem solving.

First, the findings of both studies consistently suggest that diagrammatic representations influence people's perception and interpretation of probability problem structures, and their choices of solutions. Specifically, the results confirmed prior research findings (e.g., Gattis & Holyoak, 1996; Zacks & Tversky, 1999) that use of diagrams follows a cognitively natural way of mapping corresponding visuospatial relations and components to conceptual relations and problem content. Therefore, when a probability problem admits of two distinctive solution strategies, different types of diagrams highlight and direct attention to different underlying structures of the problem, and thus steer solvers to choose solutions accordingly.

Second, these findings also point out the importance of the shared compatibility between a diagram and its represented problem type if a diagrammatic representation is to facilitate problem solving. Analyses of problem solving performance in Study 1 found that a diagram can increase or maintain high problem solving success only when the structure and content of the diagram is well-matched with those of the problem that it represents. In contrast, incompatibility between a diagram and a problem type can impair problem solving because the components represented by the diagram may lead to misinterpretation of the problem structure. The findings of Study 1 also suggest that to optimize the facilitative effect of diagrams on problem solving, information should be spatially organized in an efficient, ready-to-grasp way. The N-by-N table and its applicability in solving the combinations problem and the independent events problem in Study 1 offered supporting evidence to this argument. The generic N-by-N table significantly improved the correctness rate of solving an independent events problem, because the table structure is not only appropriate for representing independent factorial combinations of this problem, but also organizes the problem information in a highly efficient way that eases the search of outcomes and computation. On the other hand, it shares a lower structural compatibility with the combinations problem and impaired its problem solving, because all the diagonal cells on an N-by-N table represent self-repeated combination outcomes that are invalid in the topic of combinations. Thus, to design or select effective and efficient visual representations for STEM education and problem solving, it is useful to follow what Tversky , Morrison and Betrancourt (2002) have suggested as principles for effective diagram design: the Congruence Principle, which emphasizes the structural and content compatibility of a diagram with the problem it represents, and the Apprehension Principle, which emphasizes efficient information organization in the perceptual space to ease information search and computation.

Lastly, the study findings offer important and specific suggestions for how to design useful visual representations for novice probability learners. The results indicate that both the general structure type of a diagram and the level of information abstraction to represent the problem probability play a crucial role in influencing learners' and solvers' perception of a probability problem structure and solution. The findings suggest important insights for probability education. As observed, instructional materials of probability (e.g., online tutorials; lecture notes) often rely on the use of only one or two types of diagrams for demonstrating the reasoning and solving of a certain probability problem type. With the current findings, different diagrams may all be chosen for the same probability topic, based on the objectives of the lesson. For example, if the multiplication rule of probability and definition of stage-wise probabilities for unequally-likely selection outcomes are the focus of the lesson, it is suggested that an event-level representation, binary tree or binary table, be used as a visual aid. If the classical principle of probability and search for equally-like outcomes for the whole outcome space are the emphasis, it is suggested that an N-by-N table be used. For more flexible reasoning about the probability concepts and solution procedure, an N-ary tree may be most appropriate.

Limitations and Future Directions

One limitation of this dissertation is the high statistics and mathematics expertise of the participants in Study 2, which added confounds to investigation of the diagram effects on probability problem solving. Study 2 did not find facilitative effects of diagrams on problem solving correctness, perhaps because all the conditions, including the control condition, showed high solution correctness rates. Because the majority of the Study 2 participants came from a strong science, engineering, mathematics, or statistics background, they may have been over-qualified for elementary-level probability problem solving, or may have their own preferred ways of using the diagrams (or of moving directly to the relevant formula). Perhaps for this reason, some participants skipped using the provided diagrams when they solved the problems, making it less clear whether a solver's strategy choice was due to the diagram effects, or his/her own prior knowledge or problem solving preference. In the future, additional evidence is needed of the effect of these diagram types on more novice learners of probability, so that the results will offer more clear suggestions as how diagrams affect probability concept learning and problem solving.

Another limitation is that the problems and diagrammatic representations used in this investigation represent only a special case of the relevant probability principles: the selection or combination of *two* entities, not any arbitrary value "*N*". There is also a limitation in the diagram design itself. While a tree structure allows more than two layers, it will be technically difficult to design a table that represents a combination of more than two or three entities. Therefore, it is unclear to what extent the use of these diagrams and problem examples in an educational setting can lead to the abstraction of the general formulas and solutions for probability problems of any variable size. However, because diagrams may serve a "scaffolding" function, facilitating solution only for probability problems of appropriate difficulty (cf. Bobek & Corter, 2010), this limitation may not be of large practical significance in educational contexts. A future investigation may use near- and far-transfer tasks to further understand the effectiveness of these diagram types on learners' probability conceptual understanding and general solution strategy selection and application.

REFERENCES

Anscombe, F. J. (1973). Graphs in statistical analysis. *American Statistician, 27*, 17-21.

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics, 52*(3), 215-241.

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education*. Horizon Research, Inc.

Barwise, J., & Etchemendy, J. (1991). Visual information and valid reasoning. In W. Zimmermann and S. Cunningham (Eds.), *Visualization in Teaching and Learning Mathematics* (pp. 9-24). Mathematical Association of America. Washington, DC.

Bassok, M., Wu, L.-L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory and Cognition, 23*(3), 354-367.

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science, 6*, 372-378.

Bishop, A. (1989). Review of research on visualization in mathematics education. *Focus on Learning Problems in Mathematics, 11*(1), 7-16.

Braithwaite, D. W., & Goldstone, R. L. (2013). Flexibility in data interpretation: Effects of representational format. *Frontiers in Psychology, 4,*1-16.

Brasseur, L. E. (2003). *Visualizing technical information: A cultural critique.* Amityville, NY: Baywood Publishing.

Bruckheimer, M., & Arcavi, A. (1995). A visual approach to some elementary number theory, *Mathematical Gazette, 79*(486), 471-478.

Bobek, E. J., & Corter, J. E. (2010). Effects of problem difficulty and student expertise on the utility of provided diagrams in probability problem solving. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2650-2655). Austin, TX: Cognitive Science Society.

Corter, J. E., & Zahner, D. C. (2007). Use of external visual representations in probability problem solving. *Statistics Education Research Journal, 6*(1), 22-50.

Diezmann, C. M., & English, L. D. (2001). Promoting the use of diagrams as tools for thinking. *2001 National Council of Teachers of Mathematics Yearbook: The Role of Representation in School Mathematics*, 77-89.

Dufour-Janiver, B., Bednarz, N., & Belanger, M. (1987). Pedagogical considerations concerning the problem of representation. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 110-120). Hillsdale, NJ: Erlbaum.

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

Fagnant, A., & Vlassis, J. (2013). Schematic representations in arithmetical problem solving: Analysis of their impact on grade 4 students. *Educational Studies in Mathematics, 84*(1), 149-168.

Fuson, K. C., & Briars, D. J. (1990). Base-ten blocks as a first-and second-grade learning/teaching approach for multidigit addition and subtraction and place-value concepts. *Journal for Research in Mathematics Education, 21*, 180-206.

Gagatsis, A., & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology, 24*(5), 645-657.

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 231-239.

Gattis, M. (2004). Mapping relational structure in spatial reasoning. *Cognitive Science, 28*, 589-610.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Gugga, S. S., & Corter, J. E. (2014). Effects of temporal and causal schemas on probability problem solving. In P. Bello, M. Guarini, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2650-2655). Austin, TX: Cognitive Science Society.

Hadamard, J. (1945). *The psychology of invention in the mathematical field.* New York: Oxford University Press.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4), 684-689.

Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science, 30*, 581-592.

Hurley, S. M., & Novick, L. R. (2010). Solving problems using matrix, network, and hierarchy diagrams: The consequences of violating construction conventions. *The Quarterly Journal of Experimental Psychology, 63*(2), 275-290.

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

Lean, G., & Clements, M. K. (1981). Spatial ability, visual imagery, and mathematical performance. *Educational Studies in Mathematics, 12*(3), 267-299.

Lowe, R. K. (1993). Diagrammatic information: Techniques for exploring its mental representation and processing. *Information Design Journal, 7*(1), 3-18.

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

Lynch, M. (1990). The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 153–186). Cambridge, MA: MIT Press.

Manalo, E., & Uesaka, Y. (2006, December). *Quantity and quality of diagrams used in math word problem solving: A comparison between New Zealand and Japanese students.* Paper presented at the New Zealand Association for Research in Education (NZARE) National Conference. Rotorua, New Zealand.

Markman, A.B. (1999). *Knowledge representation.* Mahwah, NJ: Lawrence Erlbaum Associates.

Mason, L. M., Corter, J. E., Tversky, B., & Nickerson, J. V. (2012). Structure, space and time: Some ways that diagrams affect inferences in a planning task. In *Diagrammatic representation and inference* (pp.277-290). Springer Berlin Heidelberg.

Nemirovsky, R., & Noble, T. (1997). On mathematical visualization and the place where we live, *Educational Studies in Mathematics, 33*(2), 99-131.

Novick, L. R., & Catley, K. M. (2014). When relationships depicted diagrammatically conflict with prior knowledge: An investigation of students' interpretations of evolutionary trees. *Science Education, 98*, 269-304.

Novick, L. R., & Catley, K. M. (2013). Reasoning about evolution's grand patterns: College students' understanding of the tree of life. *American Educational Research Journal, 50*, 138-177.

Novick, L. R., & Hurley, S. M. (2001). To matrix, network, or hierarchy: That is the question. *Cognitive Psychology, 42*(2), 158-216.

Novick, L. R., Hurley, S. M., & Francis, M. (1999). Evidence for abstract, schematic knowledge of three spatial diagram representations. *Memory and Cognition, 27*(2), 288-308.

Novick, L. R., & Hmelo, C. E. (1994). Transferring symbolic representations across nonisomorphic problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1296-1321.

Novick, L. R. (1990). Representational transfer in problem solving. *American Psychological Society, 1*(2), 128-132.

Pearson, K. (1895). Contributions to the mathematical theory of evolution. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 186*, 343-414.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Erlbaum.

Phillips, L. M., Norris, S. P., & Macnab, J. S. (2010). *Visualization in mathematics, reading and science education.* Dordrecht, The Netherlands: Springer.

Polich, J. M., & Schwartz, S. H. (1974). The effect of problem size on representation in deductive problem solving. *Memory and Cognition, 2*, 683-686.

Polya, G. (1957). *How to solve it*. Princeton, NJ: Princeton University Press.

Presmeg, N. C. (1986a). Visualization in high school mathematics. *For Learning of Mathematics, 63*, 42-46.

Presmeg, N. C. (1986b). Visualization and mathematical giftedness. *Educational Studies in Mathematics, 17*, 297-311.

Presmeg, N. C. (2006). Research on visualization in learning teaching mathematics. *Handbook of Research on the Psychology of Mathematics Education*, 205-235.

Russell, W. E. (2000). The use of visual devices in probability problem solving. Doctoral dissertation, Columbia University, *Dissertation Abstracts International, 61*, 1333.

Salkind, N. J. (2006). *Statistics for people who (think they) hate statistics: The excel edition*. Thousand Oaks, CA: SAGE publications.

Schreiber, C. (2004). The interactive development of mathematical inscriptions – A semiotic perspective on pupils' externalisation in an internet chat about mathematical problems. *ZDM, 36*(6), 185–194.

Schwartz, S. H. (1971). Modes of representation and problem solving: Well evolved is half solved. *Journal of Experimental Psychology, 91*, 347-350.

Schwartz, S. H., & Fattaleh, D. L. (1972). Representation in deductive problem solving: The matrix. *Journal of Experimental Psychology, 95*, 343-348.

Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology, 91*(4), 680-702.

Simkin, D., & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association, 82*(398), 454-465.

Stylianou, D. A., & Silver, E. A. (2004). The role of visual representations in advanced mathematical problem solving: An examination of expert-novice similarities and differences. *Mathematical Thinking and Learning, 6*(4), 353–387.

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

Tversky, B. (2005). Prolegomenon to scientific visualization. In Visualization in Science Education (pp. 29-42). Springer: the Netherlands.

Tversky, B., Corter, J. E., Gao, J., Tanaka, Y., & Nickerson, J. (2013). People, place, and time: Inferences from diagrams. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3593-3597). Austin, TX: Cognitive Science Society.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23*(4), 515-557.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies, 57*(4), 247-262.

Tversky, B., Zacks, J., Lee, P. U., & Heiser, J. (2000). Lines, blobs, crosses, and arrows: Diagrammatic communication with schematic figures. In M. Anderson, P. Cheng, & V. Haarslev (Eds.), *Theory and application of diagram*s (pp. 221-230). Berlin, Germany: Springer.

van Garderen, D., & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research and Practice, 18*(4), 246-254.

Webb, N. L. (1979). Processes, conceptual knowledge and mathematical problem-solving ability. *Journal for Research in Mathematics Education, 10*, 83-93.

Winn, W. (1989). The design and use of instructional graphics. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp.125-144). Amsterdam, The Netherlands: Elsevier.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory and Cognition, 27*(6), 1073-1079.

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.

APPENDIX A: STUDY 1 TASK MATERIALS

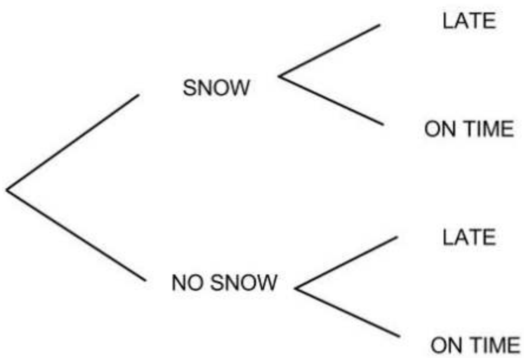
Form A: Problems, Instructions, and Diagrams

Q1. The Independent Events (Spinner) Problem with no diagram
Two spinners are constructed. Each spinner has 3 color sections of equal size: red, white, and blue. The two spinners are spun at the same time, and the result of each spinner is recorded. What is the probability of getting the same color on both spinners?

Please show all your work and box your final answer.

Q2. The Filler (Weather) Problem with a binary tree
The weather forecast says the probability that it will snow is 0.60 tomorrow. If it snows, the probability that Eva will be late for school is 0.70. If it doesn't snow, the probability that she will be late for school is 0.25. What is the probability that Eva will be late for school tomorrow?

Please use the graph to help you solve the problems.
Please show all your work and box your final answer.



Q3. The Combinations (Work-Group) Problem with an N-by-N table
Five students are in a work group. The teacher randomly selects two of them to present the group work. If there are 2 boys and 3 girls in this group, what is the probability that the teacher selects 2 girls?

Please use the graph to help you solve the problems.
Please show all your work and box your final answer.

Student A	Student B				
	____	____	____	____	____
____					
____					
____					
____					
____					

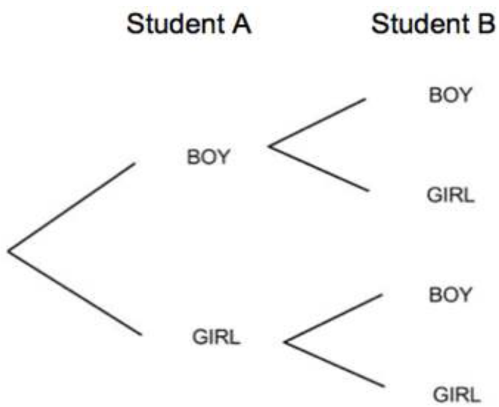
Form B: Problems, Instructions, and Diagrams

Q1. The Filler (Weather) Problem with no diagram
The weather forecast says the probability that it will snow is 0.60 tomorrow. If it snows, the probability that Eva will be late for school is 0.70. If it doesn't snow, the probability that she will be late for school is 0.25. What is the probability that Eva will be late for school tomorrow?

Please show all your work and box your final answer.

Q2. The Combinations (Work-Group) Problem with a binary tree
Five students are in a work group. The teacher randomly selects two of them to present the group work. If there are 2 boys and 3 girls in this group, what is the probability that the teacher selects 2 girls?

Please use the graph to help you solve the problems.
Please show all your work and box your final answer.



Q3. The Independent Events (Spinner) Problem with an N-by-N table
Two spinners are constructed. Each spinner has 3 color sections of equal size: red, white, and blue. The two spinners are spun at the same time, and the result of each spinner is recorded. What is the probability of getting the same color on both spinners?

Please use the graph to help you solve the problems.
Please show all your work and box your final answer.

Spinner A	Spinner B		
	____	____	____
____			
____			
____			

Form C: Problems, Instructions, and Diagrams

Q1. The Combinations (Work-Group) Problem with no diagram
Five students are in a work group. The teacher randomly selects two of them to present the group work. If there are 2 boys and 3 girls in this group, what is the probability that the teacher selects 2 girls?

Please show all your work and box your final answer.

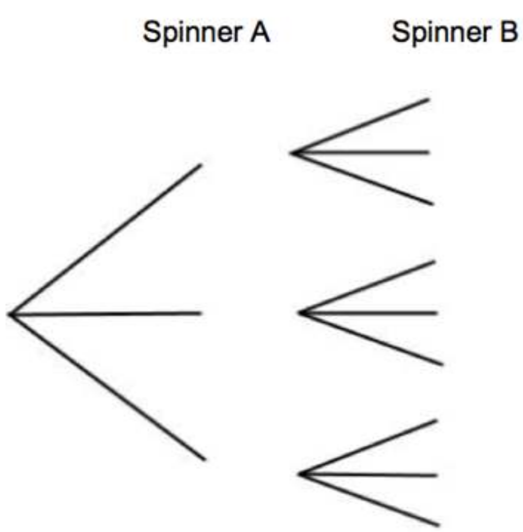
Q2. The Filler (Weather) Problem with a binary table
The weather forecast says the probability that it will snow is 0.60 tomorrow. If it snows, the probability that Eva will be late for school is 0.70. If it doesn't snow, the probability that she will be late for school is 0.25. What is the probability that Eva will be late for school tomorrow?

Please use the graph to help you solve the problems.
Please show all your work and box your final answer.

	Late	On Time
Snow		
No snow		

Q3. The Independent Events (Spinner) Problem with an N-ary tree
Two spinners are constructed. Each spinner has 3 color sections of equal size: red, white, and blue. The two spinners are spun at the same time, and the result of each spinner is recorded. What is the probability of getting the same color on both spinners?

Please use the graph to help you solve the problems.
Please show all your work and box your final answer.



Demographics Survey

Q1. What is your gender?
□ Male
□ Female

Q2. How old are you?
Age: _____________

Q3. What is the highest educational level you have completed?
□ Some college or below
□ Bachelor's degree
□ Some graduate school
□ Master's degree
□ Doctoral degree

Q4. What is your undergraduate major?
Major: ______________________

Q5. How many statistics courses have you taken, including any current ones?
Number of statistics courses taken: ___________________________

Q6. Are you currently registered in one of these two courses at TC?
□ HUDM4120: Basic Concepts in Statistics
□ HUDM4122: Probability and Statistical Inference
□ Other. Please specify which course: ___________________________________
□ None of the above

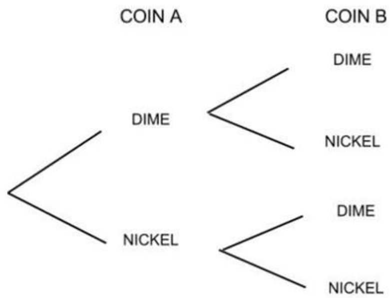
APPENDIX B: STUDY 2 TASK MATERIALS

Form A: Problems and Diagrams

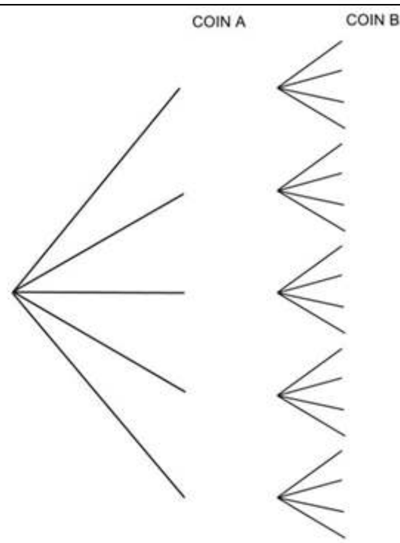
Probability topic: Combinations. Semantic schema: Simultaneous sampling.

Problem text:
Mary reaches into her wallet (which contains two dimes and three nickels) and randomly grabs two coins at once. What is the probability that the coins she selects are both nickels?

Binary tree:



N-ary (outcome) tree:



Binary table:

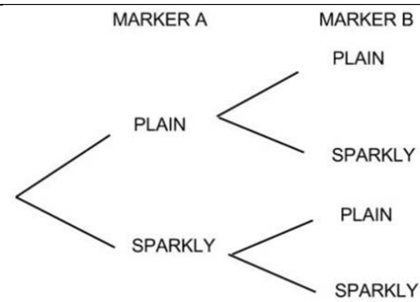
COIN A	COIN B	
	DIME	NICKEL
DIME		
NICKEL		

N-by-N (outcome) table:

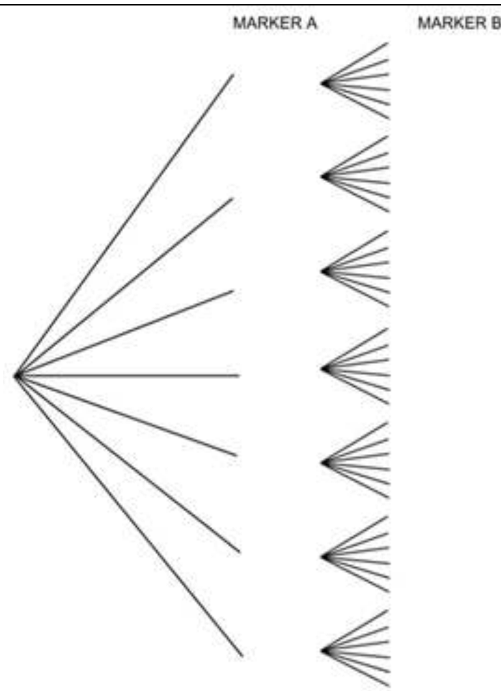
COIN A	COIN B				
	______	______	______	______	______
______					
______					
______					
______					
______					

Probability topic: Combinations. Semantic schema: Sequential sampling w/o replacement.
Problem text: Sam and his brother Andy are drawing a picture together. They are sharing a box of markers. The box contains seven markers, three of them plain colors and four with sparkles. Sam randomly selects a marker from the box. Then he passes the box to Andy and Andy randomly picks a marker from those remaining. What is the probability that Sam and Andy both select a sparkly marker?

Binary tree:



N-ary (outcome) tree:



Binary table:

MARKER A	MARKER B PLAIN	 SPARKLY
PLAIN		
SPARKLY		

N-by-N (outcome) table:

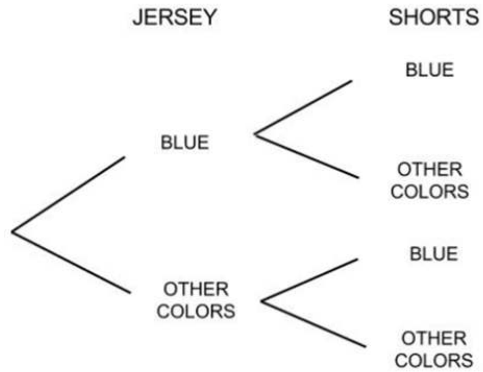
MARKER A	MARKER B _____	_____	_____	_____	_____	_____	_____
_____							
_____							
_____							
_____							
_____							
_____							
_____							

Probability topic: Independent events. Semantic schema: Simultaneous sampling (Matching).

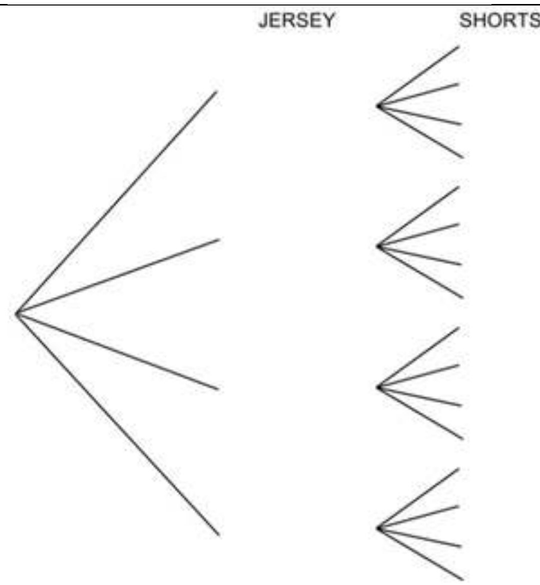
Problem text:

One day, Adam is getting dressed for playing basketball. He decides to randomly choose a basketball jersey and a pair of shorts from his closet. He has four basketball jerseys (two are orange, and two blue) and four styles of basketball shorts (three blue, and one white). What is the probability that the jersey and the shorts he chooses are both blue?

Binary tree:



N-ary (outcome) tree:



Binary table:

JERSEY	SHORTS	
	BLUE	OTHER COLORS
BLUE		
OTHER COLORS		

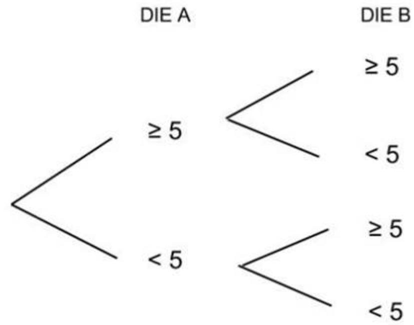
N-by-N (outcome) table:

JERSEY	SHORTS			
	______	______	______	______
______				
______				
______				
______				

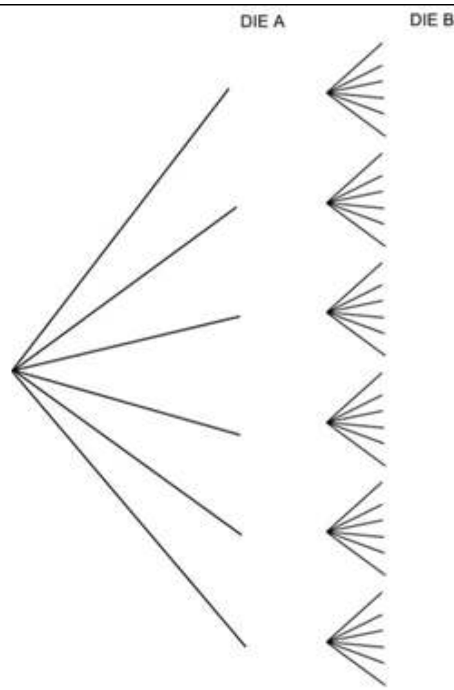
Probability topic: Independent events. Semantic schema: Sequential sampling with replacement.

Problem text:
In a dice game, Billy rolls a six-sided die and records his result. Then he hands the die to another player, Carol. Carol rolls the die and records her result. What is the probability that both of them obtain a value of 5 or greater?

Binary tree:



N-ary (outcome) tree:



Binary table:

DIE A	DIE B	
	≥ 5	< 5
≥ 5		
< 5		

N-by-N (outcome) table:

DIE A	DIE B					
	_____	_____	_____	_____	_____	_____
_____						
_____						
_____						
_____						
_____						
_____						

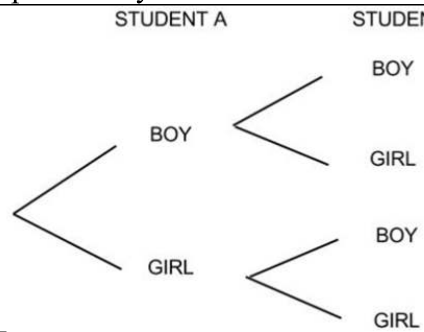
Form B: Problems and Diagrams

Probability topic: Combinations. Semantic schema: Simultaneous sampling.

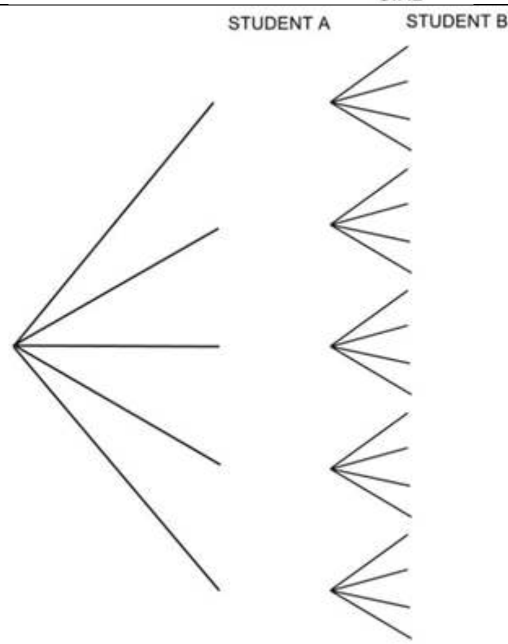
Problem text:

Five students are in a work group working on a joint project. The teacher randomly selects two of them to present the group's work to the rest of the class. If there are two boys and three girls in the work group, what is the probability that the teacher selects two girls?

Binary tree:



N-ary (outcome) tree:



Binary table:

STUDENT A	STUDENT B	
	BOY	GIRL
BOY		
GIRL		

N-by-N (outcome) table:

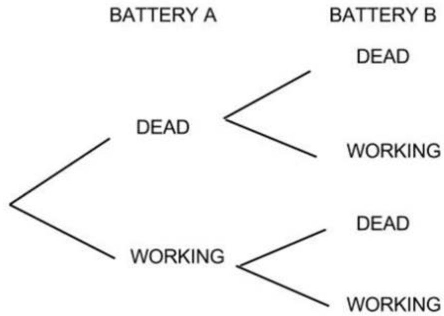
STUDENT A	STUDENT B				
	______	______	______	______	______
______					
______					
______					
______					
______					

Probability topic: Combinations. Semantic schema: Sequential sampling w/o replacement.

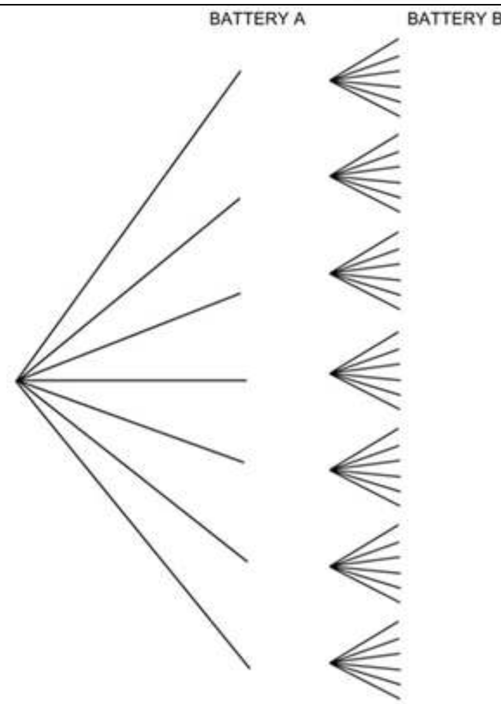
Problem text:

Bob is cleaning his kitchen drawer. There are seven loose batteries in the drawer, so he pulls them out and tests them, one by one, using a single-battery tester he has. If there are three dead batteries and four working ones in the drawer, what is the probability that the first two batteries he tests are working?

Binary tree:



N-ary (outcome) tree:



Binary table:

BATTERY A	BATTERY B	
	DEAD	WORKING
DEAD		
WORKING		

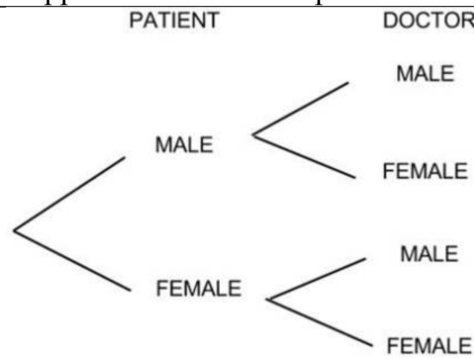
N-by-N (outcome) table:

BATTERY A	BATTERY B						
	______	______	______	______	______	______	______
______							
______							
______							
______							
______							
______							
______							

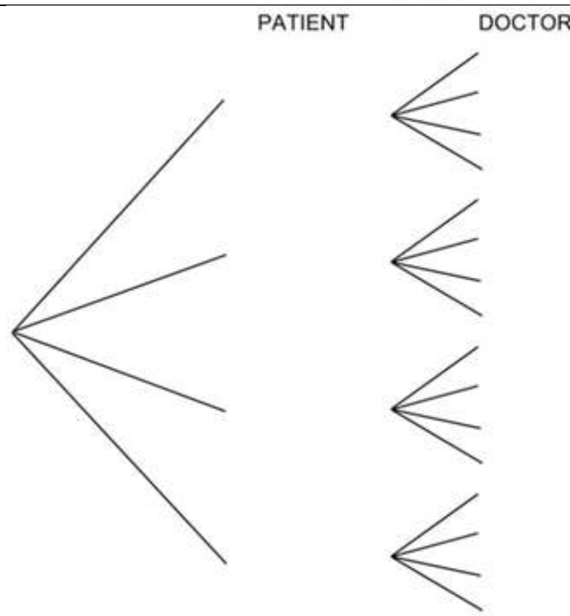
Probability topic: Independent events. Semantic schema: Simultaneous sampling (Matching).

Problem text:
At a health clinic, there are two female doctors and two male doctors. When the clinic opens on a Wednesday, there are three female patients and one male patient waiting to see doctors. If one of the four patients is randomly called, and a doctor is randomly assigned to see that patient, what is the probability that in this first appointment a female patient sees a female doctor?

Binary tree:



N-ary (outcome) tree:



Binary table:

PATIENT	DOCTOR	
	MALE	FEMALE
MALE		
FEMALE		

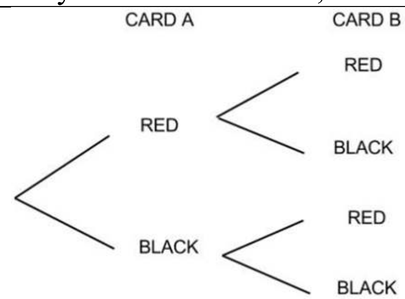
N-by-N (outcome) table:

PATIENT	DOCTOR			
	______	______	______	______
______				
______				
______				
______				

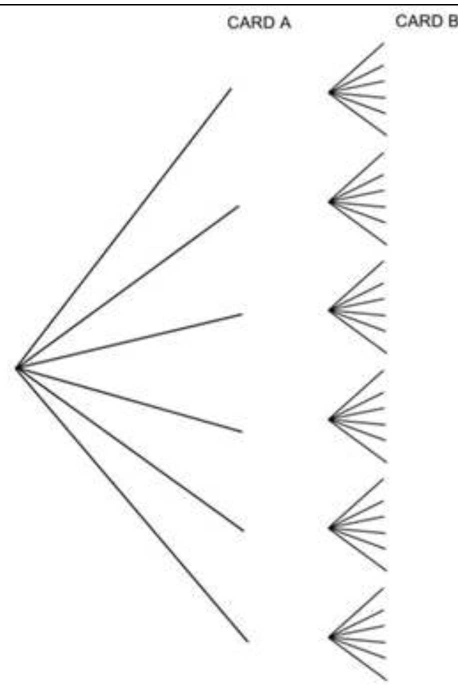
Probability topic: Independent events. Semantic schema: Sequential sampling with replacement.

Problem text:
Nate places six cards of different numbers (two in red and four in black) face down on the table. He randomly draws a card and records its number. Then he puts the card back, shuffles them until they are completely randomized, places them face down on the table again, and blindly draws a card. What is the probability that for both times, he draws a card in black?

Binary tree:



N-ary (outcome) tree:



Binary table:

CARD A	CARD B	
	RED	BLACK
RED		
BLACK		

N-by-N (outcome) table:

CARD A	CARD B					
	____	____	____	____	____	____
____						
____						
____						
____						
____						
____						

Filler Problems for Both Forms

Filler 1.

Diagram provided to the four diagram conditions:

Problem text:

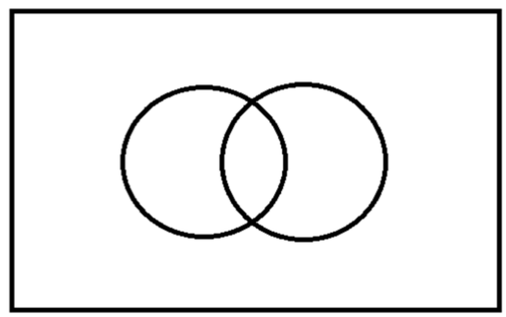
A college course has 50 students, including 8 History majors, 16 Physics majors, 12 Chemistry majors, and 14 Philosophy majors. If one student is randomly selected from the class, what is the probability that the student is not a Chemistry major?

Filler 2.

Diagram provided to the four diagram conditions:

Problem text:

There are 12 passengers on a bus. Eight of them are female and the rest male. Exactly 3 of the females and 1 male are elderly. Are the events (male) and (elderly) mutually exclusive?

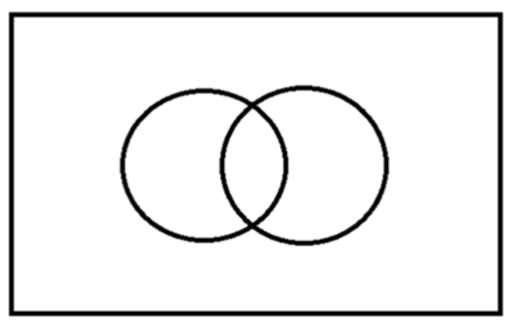


Filler 3.

Diagram provided to the four diagram conditions:

Problem text:

Twenty slips of paper are numbered 1-20, then mixed up in a hat. Jane draws one slip of paper out of the hat. What is the probability that she draws a number greater than 15 OR an odd number?



Demographics Survey

Q1. What is your gender?
□ Male
□ Female

Q2. How old are you?

Age: _____________

Q3. What is the highest educational level you have completed?
□ High school or below
□ Some college
□ Bachelor's degree
□ Some graduate school
□ Master's degree
□ Doctoral degree

Q4. What is your undergraduate major?

Major: ______________________

Q5. How many college- or graduate-level statistics courses have you taken, including any current ones?

Number of statistics courses taken: ___________________________

Q6. How many college- or graduate-level mathematics (not including statistics) courses have you taken, including any current ones?
Number of mathematics courses taken: ___________________________