USE OF EXTERNAL VISUAL REPRESENTATIONS IN PROBABILITY PROBLEM SOLVING

JAMES E. CORTER
Teachers College, Columbia University
corter@exchange.tc.columbia.edu

DORIS C. ZAHNER
Teachers College, Columbia University
dwc14@columbia.edu

ABSTRACT

We investigate the use of external visual representations in probability problem solving. Twenty-six students enrolled in an introductory statistics course for social sciences graduate students (post-baccalaureate) solved eight probability problems in a structured interview format. Results show that students spontaneously use self-generated external visual representations while solving probability problems. The types of visual representations used include: reorganization of the given information, pictures, novel schematic representations, trees, outcome listings, contingency tables, and Venn diagrams. The frequency of use of each of these different external visual representations depended on the type of probability problem being solved. We interpret these findings as showing that problem solvers attempt to select representations appropriate to the problem structure, and that the appropriateness of the representation is determined by the problem's underlying schema.

Keywords: *Statistics education research; Probability problem solving; Visual representations; Trees; Outcome listings; Venn diagrams*

1. INTRODUCTION

Consider the following probability problem:

An apartment building has four parking spaces in front (call them A, B, C, and D). There are four apartments in the building (#1, #2, #3, and #4), and each apartment has a single occupant with a single car. Every evening, all four occupants come home and park in a randomly chosen space. What is the probability that this evening they park so that the occupant of Apt #1 is in space A, the occupant of #2 is in space B, the occupant of #3 in space C, and the occupant of #4 in space D?

How would you go about solving this problem? Many people report visualizing the cars and parking spaces. After that, strategies for solving the problem tend to diverge (as do success rates). One of our points in presenting this problem (used in the present study) is that probability word problems are often simple to pose, yet difficult for many students to solve. Another point is that visualization and visual solution methods, such as self-generated external pictures and diagrams, can be very helpful in solving some probability problems.

Statistics Education Research Journal, 6(1), 22-50, http://www.stat.auckland.ac.nz/serj

Probability problem solving (PPS) can be quite difficult for students (Garfield & Ahlgren, 1988; Konold, 1989; O'Connell, 1993; Pollatsek, Well, Konold, Hardiman & Cobb, 1987), even when the mathematics involved is simple. Of course, other types of mathematics word problems are also difficult for many students, perhaps because solving them requires the problem solver to think abstractly about situations, and then model these situations using mathematical concepts. However, some researchers (Garfield & Ahlgren, 1988; Konold, 1989) have suggested that probability problem solving may be especially difficult because people have natural misconceptions about probabilistic concepts (e.g., Kahneman, Slovic, & Tversky, 1982).

Recommendations have been made for how to teach concepts in probability (e.g., Bantanero, Godino, & Roa, 2004; Gelman & Nolan, 2002; Gigerenzer, 1994; Keeler & Steinhorst, 2001; Konold, 1995, 1996; Sedlmeier & Gigerenzer, 2001). However, as pointed out by Garfield and Ahlgren (1988), only a few articles have tried to gather empirical evidence on the processes by which students solve probability problems. In one such study, O'Connell (1999; O'Connell & Corter, 1993) described a pedagogical model of recommended process steps by which students should solve probability problems. O'Connell (1993, 1999) classified student errors in probability problem solving, showing that they could be grouped into several categories: text comprehension errors, conceptual errors, procedural errors, and computational errors. Konold, Pollatsek, Well, and Lohmeier, and Lipson (1993) documented inconsistencies in probabilistic reasoning and discussed implications for probability problem solving. Due to this paucity of research on PPS, Chance and Garfield (2002) call for more research on the cognitive processes of probability problem solvers using innovative methods such as videotaped clinical interviews. The present study is intended as a step in that direction.

1.1. IS THERE A SPECIAL ROLE FOR VISUALIZATION IN PROBABILITY PROBLEM SOLVING?

In studying the cognitive processes of probability problem solvers, one issue that deserves special attention is the role of visualization. After all, anecdotal evidence suggests that visualization plays an important role in how experts solve probability problems (and mathematics problems generally). Also, informal observations of how students in statistics courses solve probability problems provide ample evidence that they sometimes spontaneously use visual devices (e.g., outcome trees) in their written work. Finally, Sedlmeier (2000) has suggested that common cognitive "fallacies" in reasoning about conditional probabilities may be ameliorated by graphical representations. Visualization may be especially important for probabilistic reasoning and probability problem solving because of the inherently abstract nature of the concepts introduced in probability.

To better understand the literature on visualization in mathematics problem solving, it is important to distinguish between *internal* visual representations (i.e., "mental imagery") and *external* visual representations (e.g., graphs, charts, pictures, etc.). Another distinction about the way external representations may be used in problem solving concerns whether the external representations are provided to the student by an instructor or experimenter, or are spontaneously generated by the student in the course of solving the problem. Although there is an extensive literature on how instructor-provided graphics can aid in scientific problem solving (summarized below), there has been little or no research on students' spontaneous creation and use of pictures, graphics and other visual devices in the course of mathematics problem-solving activities. In the present study, we use written and think-aloud protocols to study when and why probability problem solvers spontaneously produce external visual representations in their written work (when not required to do so), and what types of visual representations they employ.

1.2. PREVIOUS RESEARCH ON EXTERNAL VISUAL REPRESENTATIONS AND PROBLEM SOLVING

Results from previous research on scientific problem solving by schoolchildren (e.g., Lehrer & Schauble, 1998; Penner, Giles, Lehrer, & Schauble, 1996) and by high school and college students (e.g., Hall, Bailey, & Tillman, 1997; Hegarty & Just, 1993; Kaufmann, 1990; Mayer, 1989; Mayer & Anderson, 1991, 1992; Mayer & Gallini, 1990; Mayer, Mautone, & Prothero, 2002; Molitor, Ballstaedt, & Mandl, 1989; Santos-Trigo, 1996; B. Tversky, 2001) suggest that experimenter-provided external visual representations can aid scientific problem solving. The visual representations investigated in these studies ranged from diagrams that accompanied text (Hall, Bailey, & Tillman, 1997; Mayer, 1989; Mayer & Anderson, 1991, 1992; Mayer & Gallini, 1990; Mayer, Mautone, & Prothero, 2002) to actual physical models of scientific systems (Lehrer & Schauble, 1998; Penner, Giles, Lehrer, & Schauble, 1996). In spite of the wide range of external visual representations used in these studies, a common finding was that experimenter-provided external visuals often facilitate problem-solving success. Many of the studies also conclude that such external visual representations can aid in the development of student understanding of physical systems and mechanisms.

Incidentally, it is likely that individuals vary in the extent to which they use and benefit from visual representations. Some researchers in this area (e.g., Hegarty & Kozhevnikov, 1999; Kozhevnikov, Hegarty, & Mayer, 2002) have taken an individual differences perspective, grouping problem solvers into one of several types: those who tend to use verbal representations, and those who primarily use visual/spatial representations. Kozhevnikov et al. (2002) suggest that the visualizer group can be further split into object visualizers and spatial visualizers, with spatial visualizers showing some advantages in scientific and mathematical tasks.

Research conducted specifically in the domain of mathematics has also shown that experimenter-provided external visual representations can be useful in mathematics problem solving (e.g., Sedlmeier & Gigerenzer, 2001; Koedinger & Anderson, 1997; Nemirovsky, 1994). In particular, a number of studies (e.g., Hollebrands, 2003; Hannafin, Burruss, & Little, 2001; Hannafin & Scott, 1998) have found that the use of Geometer's Sketchpad®, a geometry graphing computer program, can be helpful in developing students' concepts and problem solving in geometry. Schwartz and Martin (2004) investigated the use of graphical tools in statistics instruction and found that experimenter-prompted graphical "invention activities" by students led to significant gains in understanding of statistical concepts.

Previous work (e.g., Russell, 2000; Zahner & Corter, 2002) in our own lab has shown that most probability problem solvers choose to use external visual representations while solving problems (after being taught the use of such visuals in an introductory statistics course), and that a wide variety of such external visual devices are used. External visual representations used by probability problem solvers include at least these types: graphs, tree diagrams, contingency tables, Venn diagrams, and pictures. Arguably, formulas and mathematical symbols could be included in this list, because they incorporate visuospatial relationships (cf. Presmeg, 1986). However, their usefulness in solving probability problems is not in question.

1.3. WHY ARE EXTERNAL VISUAL REPRESENTATIONS USEFUL IN PROBLEM SOLVING?

In order to use research results on visualization and problem solving to improve mathematics teaching and learning, it is important to ask *why* external visual representations are useful in mathematics problem solving. One possible answer to this question is that external visual representations help to augment cognitive capabilities in certain ways (e.g., Lowrie & Kay, 2001; Novick, 2001; Qin & Simon, 1995), for example by aiding memory. Tversky (2001) lists a number of possible functions of external diagrams and visual devices, including attracting attention, recording information and supporting memory, communication, providing models, and facilitating inference and discovery. Another possibility is that using multiple representations of a problem (including visual ones) leads to a fuller understanding of the problem and an increased "depth of processing" (Logie & Baddeley, 1990; Mayer, 1989, 2001; Mayer & Gallini, 1990). Other potential explanations for the use of external visual representations include the possibility that such representations can help problem solvers build a mental model of the described problem situation (Schwartz & Black, 1996). Finally, for certain problems the graphical devices may be used as a solution tool in a more specific way: for example, reading a value from a graph, or counting outcomes in an outcome tree. Alternatively, it might be that there is no benefit in using external visual representations, rather their use is just an epiphenomenon, a reaction to training from classroom instruction.

Of course, these accounts of why visualizations might be useful are not all mutually exclusive or contradictory. But only fragmentary data exist that might support or discredit any of these explanations. Some hints might come from studies examining *when* problem solvers choose to use external visual representations. For example, there is some evidence that both internal (Hampson & Morris, 1990) and external (Lowrie & Kay, 2001; Zahner & Corter, 2002) visual representations tend to be used more for unfamiliar or more difficult problems. This observation seems to support certain explanations (e.g., visuals as supporting memory, or facilitating inference and discovery) more than others.

1.4. THE PRESENT STUDY

This study focuses on the use of external visual representations in probability problem solving (PPS). We are interested in what types of problems tend to elicit use of visual representations, how and when external visual representations are used in PPS, and finally, if external visual representations facilitate correct solution of the problems. We used a variety of problem types, in order to investigate if the usefulness of visuals and the type of visual device chosen by the problem solver depends on specific aspects of the problem being solved. Specifically, we ask: Are particular types of representations used with particular problem topics (for example, problems dealing with permutations)? Also, we investigate if external visual representations are used more often with unfamiliar types of problems, because the student may have a higher cognitive load in these cases, or because the elicitation of a familiar problem-solving schema may be less likely.

As background to the present work, we assume that the process of solving a probability word problem can be broken down into roughly sequential stages (cf. Kintsch & Greeno, 1985; Mayer, 1992; O'Connell, 1993, 1999; O'Connell & Corter, 1993; Reusser, 1996). These stages are assumed to be:

i) initial problem understanding (text comprehension),

ii) formulating the mathematical problem,

iii) finding a solution method or schema,

iv) computing the answer.

Novick and Hmelo (1994) make a more gross distinction between *problem representation* and *solution procedure* phases of problem solving. Consistent with this simpler classification, our coding scheme for written protocols of students did not attempt to code use of visuals separately for stages i-iii, because we do not believe this can be done reliably with the present data. Rather, our scheme coded two types of uses of graphs or other external visual representations: a) for problem understanding, mathematical formulation, or for selection of a solution schema (i.e., any such use in the first three stages above), and b) for any use in the final stage, that of actually computing a numerical answer. We refer to the latter type of use of visual devices under the term "computational method." As an example of the first type of use, consider the use of a picture of a spinner or a Venn diagram to depict aspects of the probability word problem. An example of the second type of use of external visualizations, using them to compute an answer, would be counting the number of outcomes (leaves) in an outcome tree to find the denominator for a probability calculation.

Finally, we are interested in knowing if the use of external visual representations is associated with solution success for these probability problems. If external visual devices are used because they are helpful, then we ought to be able to find evidence of that. However, there are several factors that complicate this relationship, including the student's prior knowledge of the visual devices used, student spatial and mathematical ability, student cognitive style and the difficulty of the problem. Alternatively, it might be that the use of external visual representations is associated with solution failure, because participants might be more likely to use visual representations when they find a problem confusing or difficult (cf. Hegarty & Kozhevnikov, 1999; Lowrie & Kay, 2001).

2. METHOD

2.1. PARTICIPANTS

Twenty-six students were recruited from introductory probability and statistical inference classes during the Fall semester of 2002 from an urban college of education and psychology in the U.S.A. All participants were graduate students (post-baccalaureate) in education and social sciences, with widely varying math backgrounds. Participants were each paid ten dollars. They were informed that they were going to participate in a study of probability problem solving, and that the primary focus of the study was on the methods by which students solve problems. Because all the participants were enrolled in the same introductory statistics class at the college, their recent curricular background in probability problem solving was well-controlled and known, though the degree to which each participant mastered the material in that course was not measured. This course included approximately six lectures in probability. Topics included: events and outcome spaces, definition of probability for equally-likely and unequally-likely events, combinatorics, compound events, conditional probability, independence of events, and Bayes' Rule.

2.2. MATERIALS

Each respondent was asked to solve eight probability problems. This set of eight problems (see Appendix) was designed to include four different probability topics ("problem types") each represented at two different levels of typicality for that topic. The four different problem types were labeled: "*Combinations,*" "*Sequential,*" "*Permutations*," and "*Conditional probability*." The problems representing each of these four topics were thought to have distinct "deep structures" corresponding to four distinct problem schemas, tapping somewhat different sets of knowledge and solution skills. This factor will be referred to as "problem type." As an example, consider Problem P1, an example of a *Combinations* problem:

P1. There are 10 books on Mary's bookcase. She randomly grabs 2 books to read on the bus. What is the probability that the 2 books are "Little Women" and "War & Peace"? (Both these books are on her bookshelf.)

In the curriculum to which these participants had recently been exposed, this problem typically would have been solved using the formula for the number of *combinations of n objects selected k at a time.* That formula gives the number of possible outcomes in the sample space:

$$ {}_{n}C_{k} = \frac{n!}{k!(n-k)!} = \frac{10!}{2!(8!)} = 45 \, . $$

Then the probability of Mary selecting one particular combination (two specific books) can easily be calculated to be 1/45. Any problem requiring use of this formula (or a variant of it) is therefore assumed to share the same problem schema, and is said to be of the same basic Problem Type (*Combinations* in this case).

For each problem type, there was one typical variant and one atypical variant. The typical version was a problem that could be solved using a straightforward application of a standard probability formula known to have been taught in the participants' introductory statistics class. The atypical version was a problem that was very unlikely to be isomorphic to any problem encountered in their course, and that could *not* be solved using a single application of a standard probability formula. This manipulation of typicality may be clarified by Table 1, which presents summaries of what we judge to be appropriate *formula-based* solutions for the typical and atypical variants of each problem type. These solutions are presented to show the basic underlying structure of each problem type, without reference to surface content, and to illustrate how each problem might be solved by application of one or more standard probability formulas.

Table 1 also makes clear the types of specific problem manipulations that were used to create the atypical variant of each problem type. For the *Combinations* problem, the predicted solution for the typical version requires the problem solver to use the standard formula for the number of combinations of *n* things selected *k* at a time to calculate the number of possible outcomes. Our predicted solution for the atypical *Combinations* problem requires using this formula twice, once in the numerator and once in the denominator. For the *Sequential* problems, problem solvers must use the formula for calculating the probability of three independent events. In the typical variant the three events are identical, whereas in the atypical variant they are different events with differing probabilities. The typical variant of the *Permutations* problem asks how many different ways four items can be matched up with four "slots." The atypical version of this problem asks the same question, but orders the objects only with respect to the first two slots. We consider this an atypical problem because the computational method does not correspond to straightforward application of the formula for number of permutations of *n* objects, which is known to have been taught in the participants' introductory statistics course. Note that this atypical variant is actually simpler computationally than its typical version. For the *Conditional Probability* problems, the typical version closely resembles examples used in the students' introductory statistics course, and requires the problem solver to use the formula for conditional probability (twice). The atypical variant adds a final step, in which the formula must be used a third time.

Table 1. Example formula-based solutions for typical and atypical variants of each problem type

Problem Type	Typical	Atypical
Combinations	$_{n}C_{k} = \dfrac{n!}{k!(n-k)!}$ where n = total number of books k = number selected	$P(A) = \dfrac{_{m}C_{k}}{_{n}C_{k}} = \dfrac{\frac{m!}{k!(m-k)!}}{\frac{n!}{k!(n-k)!}}$ where n = total number of books m = number of novels k = number selected
Sequential	$P(A_1 \cap A_2 \cap A_3) = P(A)P(A)P(A)$	$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$, A_1, A_2, A_3 not necessarily equal
Permutations	# outcomes = $n!$ = $(n)(n\text{-}1)\ldots(2)(1)$ where n = number of objects	# outcomes = $(n)(n\text{-}1)$ where n = number of objects
Conditional Probability	$P(B) = P(A \cap B) + P(A^c \cap B)$ $P(A \cap B) = P(B \mid A) \times P(A)$ $P(A^c \cap B) = P(B \mid A^c) \times P(A^c)$	$P(B) = P(A \cap B) + P(A^c \cap B)$ $P(A \cap B) = P(B \mid A) \times P(A)$ $P(A^c \cap B) = P(B \mid A^c) \times P(A^c)$ $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

Across two different forms (A and B) of the test booklet, each of the eight problems was formulated with two different cover stories, or *surface content.* Surface content was counterbalanced with problem typicality across test forms. For example, the typical *Combinations* problem P1 given above involved books on a bookshelf, so a participant who saw that problem would see an atypical *Combinations* problem involving cookies in a cookie jar. For another participant who saw the second test form, the atypical *Combinations* problem would involve books on a bookshelf, and the typical version would involve cookies in a cookie jar. The Appendix shows only test form A.

2.3. PROCEDURE

A structured interviewing protocol (cf. Ginsburg, 1997) was developed for use in the interviews, and was designed mainly to elicit a reasonable level of detail in the participant protocols. The same interviewer worked with all of the participants, interviewing only one participant at a time. Participants were asked to think aloud while solving the problems, and also to show their written work with provided pen and paper. The task was not timed. However, most participants finished in less than an hour. A probability formula sheet was available to them at all times (but left face down), though no participant was observed to use it. A videotape recorder was used to capture the participants' work and student/interviewer comments. The present analyses mainly focus on the participants' written work, though the verbal transcripts were analyzed as well.

The interviewer stepped in with verbal prompts in any of four circumstances, to elicit continuation of the work or more detail about the participant's solution process. The first circumstance was if the participant could not see any way to begin solving the problem. In this situation the script called for the interviewer to ask, “In general, what would be a good first step in solving this problem?” with other follow-up questions (“How would you apply it in this case?”) if the first prompt did not elicit useful work. The second type of circumstance in which the interviewer stepped in was when the participant paused for a long time (more than 30 seconds) without thinking aloud or writing. This could indicate either that the participant was thinking silently or was at an impasse, and was responded to with “What are you thinking?” and other follow-up prompts (“Let’s back up and look at this again. How else could you solve this?”). The third situation in which prompting occurred was when the participant’s verbal or written process explanations lacked sufficient detail, for example, consisting of only a few calculations with no explanation (“Can you explain how you arrived at this?”). The fourth situation eliciting interviewer intervention was when the participant indicated that he or she was finished with the problem. In most cases, this occurred when the participant had arrived at what he or she believed was the correct answer. In other cases, this was because the participant gave up on solving the problem. In either case, the interviewer then asked the participant to explain in detail all of the steps used in the solution attempt.

Coding of the written protocols The focus of the present study is on use of external visual representations in problem solving and on the methods used to solve problems. Thus, the analyses reported in the present study focus mainly on coding of the participants’ written work. Three particular aspects of the written problem solutions were coded, based upon a scheme developed in previous research (Russell, 2000). The first coded aspect was whether or not the participant gave the correct answer to the problem. The second aspect coded, described in more detail below, was the type of external visual representation used (if any) by the participant. The third aspect coded, also described in more detail in the next section, was the type of general computational method used to solve the problem. Here the identified types were: formula, graphical, or procedural.

External visual representations Written protocols for each problem solution were coded for use of different types of external visual representations. The coded types of external visual representation included *pictures, outcome listings, trees, contingency tables, Venn diagrams, novel schematic representations*, and *spatial reorganization of the given information*.

An external visual representation was coded as a *picture* if it attempted to represent the real-world situation conveyed in the problem in a non-symbolic, pictorial way. For example, in a problem about use of a spinner with separate areas marked “red,” “blue,” and so on, any picture of a spinner type device would count as a *picture* (see Figure 1 for an example). A visual device was coded as an *outcome listing* if it gave a list of outcomes in some relevant outcome space, for example: {HH, HT, TH, TT} as the outcomes space for the experiment of flipping a coin twice. A visual representation was coded as a *tree* diagram if the participant attempted to organize the information from the problem in either a complete or a partial outcome tree. An example of the use of a tree diagram is shown in Figure 2. A visual representation was coded as a *contingency table* if the participant presented the information from the problem as probabilities or frequencies in a two-way table. A visual representation was coded as a *Venn diagram* if the participant used a Venn diagram to represent set relationships.

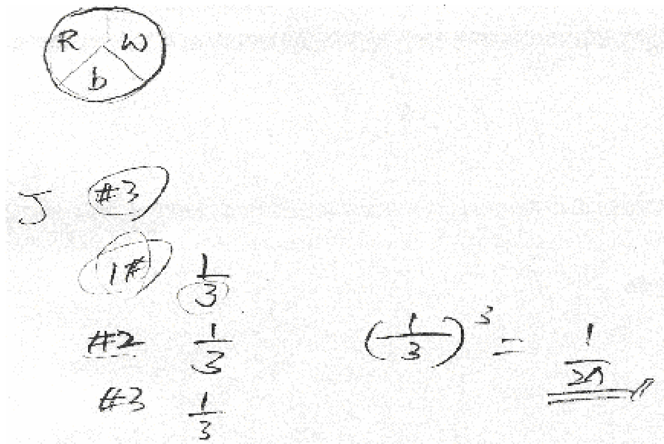

Figure 1. A participant's written work for the typical version of the Sequential problem, illustrating use of a picture

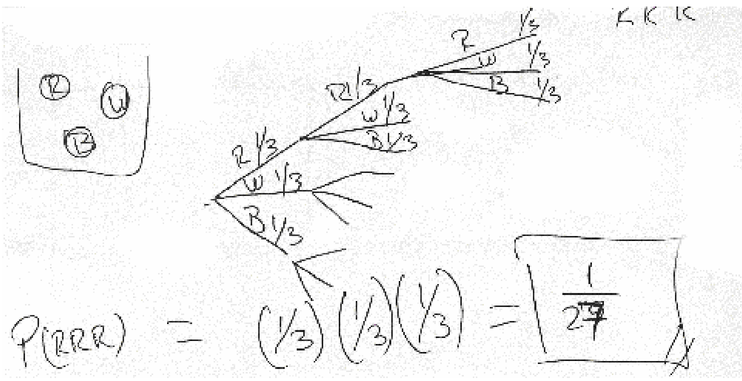

Figure 2. A participant's written work for the typical version of the Sequential problem, illustrating use of an outcome tree (and a picture)

Besides these standard pictorial and schematic representations used in previous studies (Russell, 2000; Zahner & Corter, 2002), we created two additional coding categories to cover cases not handled by the above classes. The first is a code indicating any attempt to invent and use what we termed a "*novel schematic representation*." Use of the term "novel" is meant to denote a schematic visual device that was not taught in the introductory class the participants were taking or had taken, nor used in standard probability texts. It is not meant to imply that the student invented and used a previously unknown type of visual device. An example of this category is a graphic used by several subjects for the *Permutations* problems: a list of four names whose elements are connected by lines or arrows to elements in a list of four numbers (see Figure 3). This type of representation (that we would classify as a directed graph) is apparently an attempt to develop or discover the correct outcome space for the problem. This type of representation uses spatial information and schematic elements (lines or arcs) to represent relational aspects of the problem, and is thus different from a simple outcome listing. The second additional coding category was defined to include any *spatial reorganization of the given information*. Use of a spatial organization scheme for information is not a formal graphical representation nor a purely pictorial representation. However, we have included this coding category because we have observed frequent use of spatially-organized rewriting of information to aid in problem solving. In the present study, many participants were observed to line up corresponding given probabilities or conditional probabilities (see Figure 4). This practice may make it easier for novice problem solvers to check for needed or missing information, to break down problem solution into subparts, or to make visual associations to relevant formulas.

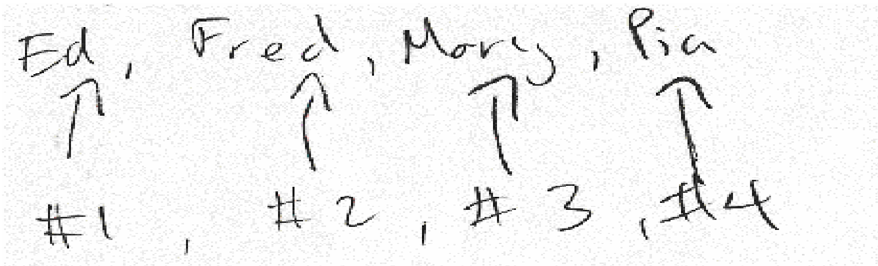

Figure 3. A participant's written work for typical version of the Permutation problem, illustrating use of a novel schematic representation

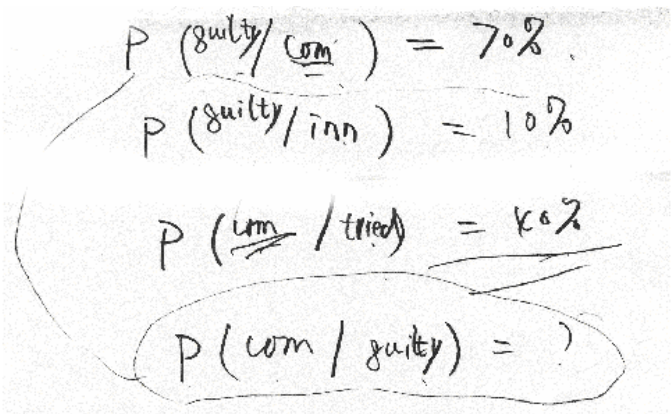

Figure 4. A participant's written work for the atypical version of the Conditional problem, illustrating use of spatial reorganization of given information

Computational method We also coded the *computational method* used by the problem solver, that is, the means by which the problem solver actually computes the answer required by the problem. We did this because in the course of coding the student protocols, we noticed that sometimes visual representations were used very early in the problem solving process, for example while the problem solver seemed to be still trying to understand the given problem information or to classify the problem, and sometimes later in the solution process, for example when the subject was trying to compute the actual numerical answer. In an effort to begin to understand what specific role or function external visual representations are serving in probability problem solving, we decided to separately code the method by which the problem solver actually computed the numerical answer required in each of these problems. We classified this later stage of each problem-solving protocol into three broad classes of computational method: *formula, procedural,* and *graphical*.

The computational method was characterized as *formula-based* problem solving if the participant wrote down an explicit (standard) formula, then substituted in quantities and solved the problem. An example of a formula-based computational method would be the use of the combinations formula followed by the necessary calculations: ${}_5C_2 = 5!/(2!)(3!) = (5)(4)/(2)(1) = 10$. This complete and rather formal method was distinguished from a *procedural* approach, which was used to code solutions carrying out a calculation involving only numbers without reference to any general formula or underlying principle. An example of a procedural approach would be if the participant calculated the probability of getting three heads in three flips of a coin by simply multiplying $(1/2)(1/2)(1/2) = 1/8$ without indicating any rationale for that procedure. A computational method was considered *graphical* if the subject used an external visual device to solve the problem, but only if the graphical device was judged to be instrumental to the method by which the student arrived at the actual numerical solution. An example of a graphical computational method would be if the subject multiplied two conditional probabilities that were taken from branches of a tree diagram.

No computational method was coded if the participant did not attempt to solve the problem. Occasionally, multiple computational methods were coded for a single problem. This occurred only when a participant attempted the problem, then abandoned that attempt, and attempted another computational method.

In order to assess reliability of the coding of the written protocols for external representations and computation method, a second rater coded all student solutions. Initial percent agreement between the two raters was over 90% for both external representation and computation method. Discrepancies were discussed by the two raters and the resulting consensus was used in all analyses reported.

Coding of audio protocols In order to better understand how the external visual representations are being used by problem solvers, we also transcribed and examined the audio portion of the videotapes capturing the participants' think-aloud protocols. Each utterance in a participant's audio transcript was coded to indicate if the participant was engaged in either of two broad phases or stages of problem solution: 1) a problem-representation phase that involves understanding the problem text and reformulating the problem in mathematical form; or 2) a solution phase, that involves selecting a solution strategy and implementing it. The video track of the tapes focused on participants' written inscriptions, including use of external representations. The video tapes were used to match uses of external visual devices with verbal statements by the participants about their thoughts and actions and the general phase of problem solving that they were engaged in: either problem representation or strategy selection and solution.

2.4. RESULTS

Preliminary analyses showed that individual problems and problem types varied considerably in difficulty. The rightmost column in the Appendix shows the proportion of subjects who correctly solved each problem. These proportions vary from a low of .08 for Problem P8 (*Conditional probability*, atypical) to .73 (for Problems P3 and P4, the typical and atypical *Sequential* problems). Regarding problem type, it was found that participants were most successful at solving the *Sequential* problems (.73 correct overall), followed by the *Permutations* problems (.48), then the *Combinations* problems (.29), and finally the *Conditional probability* problems (.25). These differences in solution rate among problem types were significant: in a log-linear analysis with factors Problem

Type, Typicality, and Correctness (the dependent variable), the Problem Type × Correctness association was significant ($\chi^2(3) = 31.56$, $p < .05$).

We were also interested in whether the atypical variant of each problem type was more difficult for problem solvers than the typical variant. Results indicated that typical and atypical variants did not differ in mean difficulty across all four problem types: mean proportion correct for the four typical problems encountered by each subject was .43, whereas for the four atypical problems it was .44. However, it is clear from the solution rate for individual problems (see Appendix) that the typical-atypical difference in solution rates varies across problem types. This interaction was tested by the three-way association of Problem Type × Typicality × Correctness in the loglinear analysis described above. This association was significant, ($\chi^2(3) = 17.00$, $p <.05$). Consequently, it is necessary to examine the effects of typicality separately for each problem type.

In particular, a difference in the expected direction was found for the *Conditional Probability* problems, with 42% of participants correctly solving the typical version of the problem versus only 8% for the atypical version. Unexpectedly, for the *Combinations* problem the solution rate for the atypical variant of the problem was much higher (at 73%) than for the typical variant (at 12%). In order to understand this unexpected result, we analyzed participants' written protocols to identify the specific solution methods used by participants for these problems. We found that most participants did not use the combinations formula at all to solve the atypical variant of the *Combinations* problem; rather they tended to solve this problem by treating it as a "sequential" problem involving sampling without replacement. For example, the problem can be solved using the formula: $P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1)$.

Inspection of the individual student protocols revealed that 92% (24 out of 26) of the participants selected this alternate method to solve the atypical *Combinations* problems. This probably occurred because the atypical problem is extremely difficult using the *Combinations* approach: the only two participants who tried this approach both failed to solve it. In contrast, exactly half of the 24 participants who adopted the sequential-events approach for the atypical variant succeeded in solving it. Note that many participants (73%) also tried to solve the typical *Combinations* problem using a sequential-events approach. However, all of these subjects failed to solve the problem, contributing to the overall low solution rate for the typical version. The difference in apparent difficulty of the sequential approach to these two problems probably involves that fact that in the typical *Combinations* problem, order is not important (but the use of the sequential solution method tends to elicit a solution attempt involving ordered pairs). Thus, many participants gave the answer 1/90 for this problem using the sequential approach, whereas the correct answer is 1/45. In the atypical variant, in contrast, there is a symmetry to the outcomes in the outcome space such that order is irrelevant.

What kinds of external visual representations are used? For each specific form of external visual representation, we calculated the percentage of participants who used that representation at least once. As shown in Table 2, we found that participants most often used *reorganization of the given information* (used at least once by 96.2% of the participants), followed by use of *picture*s (by 84.6% of the participants), *novel schematic representations* (65.4% of the participants), *trees* (53.8%), *outcome listings* (38.5%), *contingency tables* (7.7%) and finally *Venn diagrams* (3.8%).

Table 2. Frequency and percentage of participants using each type of external visual representation at least once, with frequency and percentage use of each representation across all problem solutions

	By participant (*N*=26)		By problem solution (*N*=208)	
Representation	Frequency	%	Frequency	%
Reorganize	25	96.2	72	34.6
Outcome Listings	10	38.5	20	9.6
Contingency Tables	2	7.7	6	2.9
Venn Diagrams	1	3.8	1	0.5
Trees	14	53.8	27	13.0
Novel Schematic	17	65.4	24	11.5
Pictures	22	84.6	64	30.8

Are different types of external visual representations used with different types of problems? We investigated the relationship between the type or topic of the probability problem (*Combinations, Sequential, Permutations,* and *Conditional*) and the type of representation that participants chose to use for it. In this analysis no distinction was made between the typical and atypical versions of each problem type. Table 3 summarizes how often each type of external representation was used for each type of problem. Because there were two problems of each type, each entry in this table is calculated across a total of 52 problem solutions.

Table 3. Frequency and percentage of problems of each type for which a given type of external representation was used (out of N=52 problem solutions), with χ^2 goodness-of fit tests evaluating differences in the frequencies of use of each representation across the four problem types

Representation	Combinations		Sequential		Permutations		Conditional		$\chi^2(3)$
	Freq	%	Freq	%	Freq	%	Freq	%	
Reorganize	26	50.0	4	7.7	4	7.7	38	73.1	47.6*
Outcome Listings	9	17.3	7	13.5	4	7.7	0	0.0	9.2*
Contingency Tables	0	0.0	0	0.0	4	7.7	2	3.8	---
Venn Diagrams	0	0.0	0	0.0	0	0.0	1	1.9	---
Trees	5	9.6	6	11.5	5	9.6	11	21.2	3.7
Novel Schematic	0	0.0	0	0.0	24	46.2	0	0.0	72.0*
Pictures	16	30.8	31	59.6	15	28.8	2	3.8	26.4*

*$p < .05$

For each row of the table, we used a chi-square goodness-of-fit test to test if each visual representation was used with unequal frequencies across problem types (i.e. columns). The chi-square goodness-of-fit test revealed that the frequency of use of *reorganization of given information* differed significantly across problem types ($\chi^2(3)$= 47.6, $p < .05$). This representation was used most often for the *Conditional Probability* problems (73.1% of the time) and the *Combinations* problems (50%). Usage was also distributed unequally across problem type for *outcome listings* ($\chi^2(3) = 9.2$, $p < .05$), with the most frequent use being for *Combinations* (17.3%) and *Sequential* (13.5%) problems. Use of *Novel schematic representations* was also distributed unequally across problem type ($\chi^2(3) = 72.0$, $p < .05$), because these representations were used only for the *Permutations* problems (46.2% of the time). *Novel schematic devices* may be tried especially often for the specific permutations problems used here because these problems are difficult for novices to recognize as permutation problems. That is because the cover stories for these particular permutation problems involve matching two sets of entities (e.g., tutors with students) rather than simply ordering one set of objects. This situation does not plug neatly into any formula or solution schema that students had been taught. This situation apparently spurred participants to try to understand these relatively unusual problems by inventing or adapting “novel” graphical representations.

Also, the use of *pictures* was distributed unequally across problem type ($\chi^2(3) = 26.4$, $p < .05$), due to very frequent use of pictures for the *Sequential* problems (in 59.6% of problem solutions), *Combination*s (30.8%), and *Permutations* (28.8%) and infrequent use (3.8%) for the *Conditional Probability* problems. The use of *trees* did not vary significantly across problem type ($\chi^2(3) = 3.7$, $p < .05$). Inspection of Table 3 reveals that trees were used about 10% of the time or more for all four problem types. This result seems to show that at least for these types of probability problems, trees were perceived by study participants as widely applicable. The use of *Contingency tables* and *Venn diagrams* was too infrequent to be tested in the manner.

Sometimes problem solvers used more than one form of external visual representation in a single problem solution. Figure 5 shows the percentage of use of single and multiple representations across all problem solutions, separately by problem and problem type. Across the eight problems, multiple external visual representations were used in 23.6% of the problem solutions. The most common combinations of multiple representations were *pictures* with *reorganization* (used in 13% of the problem solutions), and *pictures* with *trees* (used in 6%; see Figure 2 for an example). All other instances of multiple external representations occurred less than 2% of the time. Multiple representations were used most often for *Combinations* and *Sequential* problems. This may simply reflect the fact that pictures were used quite often for these problem types, as shown in Table 3.

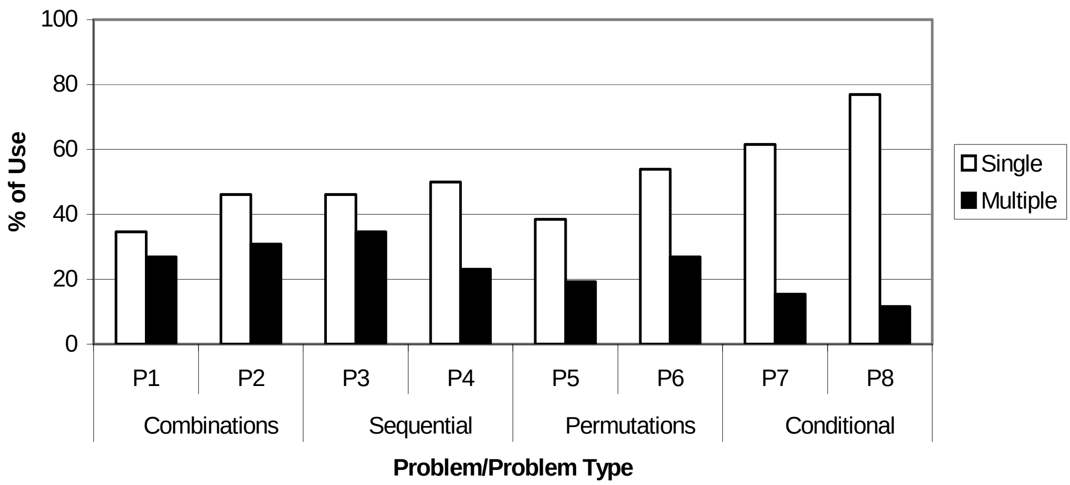

Figure 5. Percentage of use of single and multiple external representations in the problem solutions (N=208), by problem and problem type

The above results showing differences in frequency of use of specific representations across the four problem types demonstrate that participants are selecting representations based on the type of problem they are trying to solve, presumably reacting to differences in the problem schema for the four problem types. This suggests that participants’ solution methods (at least, their use of external visual representations) vary depending on the problem's underlying schema or "deep structure." We return to this issue in the Discussion section.

Is solution success associated with use of external representations? If external visual devices do indeed serve some purpose for problem solvers, then we might expect an association between solution success and the specific external representation used (if any). We explored this idea by estimating the conditional probability of solution success given use of each type of external visual representation. The results, shown in Table 4, show that use of particular external visual representations was associated with higher rates of solution correctness for some problem types (compared to baseline performance for that problem type), and with lower rates of success for others. For example, the use of *reorganization* is associated with a higher rate of solution success for *Combinations* problems (.32 versus .29), but with a lower rate of success for *Sequential* (.50 versus .73), *Permutations* (.00 versus .48), and *Conditional* (.15 versus .25) problems.

Table 4. Proportion of correct solutions, given the use of a particular representation, separately by problem type. Dashed lines indicate a cell with fewer than four uses of that representation (i.e., $n \leq 3$).

Representation	Combinations	Sequential	Permutations	Conditional
Reorganize	.32	.50	.00	.15
Outcome Listings	.43	.42	--	--
Contingency Tables	--	--	.00	--
Venn Diagrams	--	--	--	--
Trees	.60	.67	.25	.25
Novel Schematic	--	--	.44	--
Pictures	.23	.63	.59	--
Mean P(correct)	.29	.73	.48	.25

Table 4 shows that for the *Combinations* problems, use of *reorganization*, *outcome listings*, or *trees* were all associated with higher rates of solution success, whereas use of *pictures* was associated with a lower rate of success. Presumably, the first three types of representations are useful here because the essence of such combinatorics problems is to identify the number of outcomes in the outcome space. However, trees are not usually useful for problems involving simultaneous sampling of multiple objects (where order is not important). We therefore reexamined participants' solutions to try to understand this association. We found that trees were used in only five solutions for the *Combinations* problems, and all of these were cases where the problem solver was treating the problem as a sequential problem rather than using the combinations formula.

For the *Sequential* problems, use of any external visual representation was associated with a lower rate of solution success. *Sequential* problems were the easiest type of problem overall, with *P*(correct) = .73, so it may be that participants did not feel any need to call upon visual representations unless they were among the few who experienced difficulty with these problems.

For the *Permutations* problems, use of *reorganization*, *contingency tables,* and *trees* was associated with lower rates of solution success. Contingency tables in particular do not seem appropriate for permutation problems, which involve ordering a single set of objects. Trees are rarely used to represent sequential sampling without replacement, though in principle they could be applied. However, use of *pictures* was associated with a higher rate of solution success for these *Permutations* problems. *Pictures* may be especially useful for these particular (unusual) permutations problems, which are unusual in that they describe matching *two* sets of objects rather than ordering a single set. It may be that pictures facilitated the realization that the ordering of one of these sets is arbitrary.

Finally, for the *Conditional probability* problem, external visual representations were not often used. Use of *reorganization* was associated with slightly lower rates of solution success. It is surprising that trees were not often used for these problems because their use for such problems was explicitly described in the course.

Thus, this table seems to offer mixed evidence concerning the usefulness of external visual representations in probability problem solving. The positive associations found seem easily explainable. We argue that the few observed negative associations between external visual representations and solution success do not prove that use of external representations is harmful in probability problem solving. Rather, the negative associations may arise because external visual representations are more often called upon when the problem is especially difficult for the problem solver (cf. Hegarty & Kozhevnikov, 1999; Lowrie & Kay, 2001).

In addition to the analysis of solution correctness given the use of a particular external visual representation, we correlated the number of times a participant used *any* external visual representation (which ranged from 2 to 18 for participants) and the participant's overall solution success (defined as number of problems correct out of 8 for a participant). Results from this analysis show that there is a significant *negative* overall correlation between the use of an external visual representation and solution success (r = -.40, p < .05). We also found marginally significant negative correlations between use of certain specific representations and solution success. Specifically, the use of *reorganization of the given information* was negatively correlated with solution success (r = -.37, p = .06), as was the use of *outcome listings* (r = -.37, p = .06).

What computation methods are used in PPS? Are external visual devices used in computing problem solutions? We calculated how often each of the three computational methods (formula, procedural, graphical) was used for each problem. Results showed that students used the procedural computational method most often (on average in 5.5 out of 8 problems) followed by formula-based computational methods (1.19 out of 8 problems) and finally, graphical solutions (0.42 problems out of 8.) Thus, external visual representations were rarely used to compute solution. For 13.5% of problems overall, subjects did not complete the problem to the point of computing a solution. Multiple computation methods (coded only when the participant made multiple solution attempts) were observed only 2.4% of the time.

Are different computation methods used with different problem types? We calculated frequency and percentage of use of each of the three types of computation method across problem types. Note that more than one type of solution method could be coded for a given solution, and that if the problem solver did not attempt to compute a numeric solution no computation method was coded. Results showed that formulas were used most often for the *Conditional* problems (21.2% of the time) and *Combinations* (21.2%). A procedural computation method was used most often for *Sequential* problems (84.6%), *Combinations* (75.0%), and *Permutations* (75.0%). Finally, a graphical computational method was used most often with the *Conditional Probability* problems (9.6%). Note that for the *Conditional Probability* problems the observed student solutions were not purely graphical; rather the tree graphs were typically used in conjunction with procedural calculations. We also performed a chi-square goodness-of-fit test to determine if each computational method was used equally often across all four problem types. The results (Table 5) show differential use of the procedural computation method across problem types ($\chi^2(3) = 8.57$, $p < .05$). Frequency of use did not differ significantly across problem types for the procedural computation method ($\chi^2(3) = 6.05$, $p > .05$), nor for the graphical computation method ($\chi^2(3) = 5.21$, $p > .05$). The pattern of results in Table 5 suggests that a procedural solution method is used relatively less often for the *Conditional Probability* problems.

Table 5. Frequency and percentage of problem solutions (n = 52) of each type for which a given computational method was used, with χ^2 goodness-of-fit tests of differences in frequency of use of each method across the four problem types

Computational method	Combinations		Sequential		Permutations		Conditional		$\chi^2(3)$
	Freq	%	Freq	%	Freq	%	Freq	%	
Formulas	11	21.2	3	5.8	6	11.5	11	21.2	6.05
Procedural	39	75.0	44	84.6	39	75.0	21	40.4	8.57*
Graphical	3	5.8	3	3.8	0	0.0	5	9.6	5.21

*$p < .05$

Is solution success associated with computational method? We also checked for associations between solution success and computational method, separately by problem type (Table 6). Use of a formula-based computation method was associated with a higher rate of solution success only for *Combinations* problems. This makes sense, because the combinations problems are arguably best solved via formulas. Use of a procedural method was associated with the highest rates of solution success for the other three problem types. This is probably because if it is intuitively clear to a student how to solve a problem, only the computations need be written down (and the solution would be coded as a procedural one). Using a graphical method to aid in computing the solution was observed infrequently, except for the *Conditional probability* problems. For these problems, the tree can be used to organize the procedural calculations.

Table 6. Proportions of correct solutions given the use of a particular computation method, separately by problem type. Dashed lines indicate a cell with fewer than four uses of that representation (i.e., n ≤ 3).

Computational method	Combinations	Sequential	Permutations	Conditional
Formula	.38	--	--	.11
Procedural	.30	.77	.57	.46
Graphical	--	--	--	.30
Mean	.29	.73	.48	.25

Are there differences between typical and atypical problems in the use of external visual representations or computational method? We investigated whether there is a difference in the rates of use of visual representations for the typical and atypical problems. To test this, for each type of representation we compared the summed frequency of its use for the four problems presented in their typical versions to the summed frequency of its use for the four problems presented in their atypical versions. The results show that the only significant difference in use of an external representation between typical and atypical problems was for pictures (paired-sample $t(25) = -3.86$, $p < .05$). Specifically, *pictures* were used more often for atypical problems (for 38.5% of problems) than for typical problems (23.0%), as shown in Table 7. This result is not at all surprising – pictures may be used especially often to try to better understand or structure a difficult problem, especially one that does not plug in easily to a familiar solution schema. On the other hand, if the solution method is obvious, nothing is gained (and time and effort are expended) in drawing a picture. The only other representation used more often for atypical problems (though the difference is not significant) is *reorganization.* As for pictures, it can be argued that this is a general-purpose method that is often useful when the problem text itself is difficult to understand.

Table 7. Frequency and percentage use of each type of external representation, separately for the four typical and four atypical problems experienced by each participant. (= significant difference between the total number of uses for typical and atypical using a dependent samples t test, df = 25)*

Representation	Typical		Atypical	
	Freq	%	Freq	%
Reorganize	34	32.8	38	36.5
Outcome Listings	12	11.5	8	7.8
Contingency Tables	4	3.8	2	2.0
Venn Diagrams	1	1.0	0	0.0
Trees	15	14.5	12	11.5
Novel Schematic	13	12.5	11	10.5
Pictures*	24	23.0	40	38.5

*$p < .05$

Table 8 reports the rates of use of different computation methods for typical versus atypical problems. There were no significant differences in the use of different computation methods for typical and atypical problems.

Table 8. Percentage use of each type of computational method, summed across problems, separately for the four typical and four atypical problems experienced by each participant

Computational Method	Typical		Atypical	
	Freq	%	Freq	%
Formula	15	14.5	16	15.5
Procedural	73	70.3	70	67.3
Graphical	6	5.8	10	10.0

2.5. ANALYSIS OF AUDIO PROTOCOLS

As described in the Methods section, each utterance in the audio track of the session videotapes was coded as relevant to either the problem solvers' problem-representation phase or as part of the solution execution phase (cf. Novick & Hmelo, 1994). We also matched any use of an external visual representation in a solution (as captured in the video track) to any utterances made simultaneously. This enabled us to classify uses of external visual representations as being associated with either or both of these broad temporal stages of problem solving.

Of the 2,756 utterances in the audio transcripts, approximately 63% of them were coded as part of the Problem Representation phase (1,734 utterances) and approximately 32% of the utterances were coded as Solution Execution (881 utterances). The remaining utterances were not coded as part of either phase. This occurred if the utterance was a meta-comment or irrelevant, for example concerning how difficult the problem was or the temperature of the room in which the study was being conducted.

Of the 1,734 utterances that were coded as part of the Problem Representation phase, 417 (24.08%) of them were matched with the use of an external visual representation. Of the 881 utterances that were coded as involving the Solution Execution phase, 31 utterances (3.68%) were matched with the use of an external visual representation. Although participants did use external visual representations during both phases of problem solving, they tended to use them more often during the Problem Representation phase than during the Solution Execution phase.

These associations were broken down by type of representation used (Table 9). Results indicated that participants more often tended to use the external visual representations to help understand and organize the problem text (i.e., in problem understanding) than to select or execute solutions. This trend was especially strong for *pictures*, *reorganization of the given information*, and *novel schematic representations*. For example, *pictures* were used significantly more often during the problem representation phase because participants claimed that they helped them visualize the problem more clearly. One subject explained, “I drew the ten cookies because I needed literally to visualize it, and then based on what the information is in this problem, there’s obviously….there’s ten different types of cookies.” (Subject #16)

Table 9. Total frequency of use of particular representations, by problem solving phase

Representation	Problem Representation	Strategy & Execution
Reorganize	169	9
Outcome Listings	17	3
Contingency Tables	11	3
Venn Diagrams	4	0
Trees	50	13
Novel Schematic	45	1
Pictures*	120	4
Total	417	32

*$p < .05$

3. DISCUSSION

Our results show that students sometimes choose to use self-generated external visual representations while solving probability word problems. Presumably, this is because problem solvers believe that these representations are useful in solving the problems, because in this study they were requested merely to “show their work,” and not explicitly requested to produce any diagrams or other visual devices. A skeptical observer might worry that the verbal prompts occasionally issued by the experimenter here could have served as a general prompt to try alternative representations. However, Russell (2000) found similar levels and patterns of use of these types of representations in students’ actual answers to course assignments, lending confidence to the conclusions that problem solvers choose to use such representations because they are thought to be useful.

The results also document what types of spatial and graphical devices are used in probability problem solving. Using a very broad definition of external visual representations, the types we identified included (in decreasing order of frequency of use): *reorganization of the given information, pictures, novel schematic representations, trees, outcome listings, contingency tables,* and *Venn diagrams* (cf. Russell, 2000). Of course, our reported relative frequencies of use for these representations may not generalize to other curricula, other specific sets of problems, and other problem solvers.

We found evidence that the frequency of use of each of these different external visual representations depends on the *type* of probability problem being solved. Type of problem refers to the basic problem schema, and not to surface characteristics of the problem. The types of probability problems studied here were *Conditional Probability, Combinations, Permutations,* and *Sequential* problems. Our results showed that *pictures* were used most often for *Sequential, Combinations*, and *Permutations* problems; *outcome listings* were used more often for *Combinations* and *Sequential* problems; *trees* were used most often for *Conditional* problems; *novel schematic representations* were used mainly with *Permutation* problems; and *reorganization of given information* was used more often with *Conditional* and *Combinations* problems. One way to interpret these findings is that problem solvers attempt to select representations appropriate to the problem's structure, and that the appropriateness of the representation is determined by the problem's solution schema, not by surface characteristics.

However, any conclusions as to *specific* associations between the type of visual device and the problem type must be tempered by consideration of the particular set of problems used here to represent these general types. First, the *Permutation* problems studied here were unusual in that they described situations in which two sets of entities (e.g., tutors and students) were to be matched in a one-to-one fashion, but the ordering of one set was arbitrary, making the problem isomorphic to an ordering problem. This "schema mismatch" may have made these problems particularly difficult for our problem solvers, spurring more attempts to use novel schematic representations and pictures. Second, problems were experienced with our manipulation of typical and atypical problems for the *Combinations* problems. The atypical *Combinations* problems were most often solved by an alternative method, using a sequential-events approach, that resulted in a higher rate of success than for the typical variant of this problem type. Thus, although it was our intention to manipulate problem typicality in such a way that the atypical problems were at least as difficult as the typical ones, this did not happen for this problem type. In future studies, we hope to more fully refine and explore the notions of problem typicality and difficulty, and to try to disentangle their effects experimentally by careful development and piloting of materials.

In future research, we also hope to more fully investigate aspects of the schematic devices that play a part in determining the appropriateness of a representation for a given problem. We believe that the seven types of external visual representations studied here differ in some important ways. Three of the visual representations (reorganization, outcome listings, and contingency tables) can be considered forms of *tabulation*. Another three (Venn diagrams, trees, and novel schematic representations) could be classified as *schematic devices*, and the final type (pictures) refers to iconic representations of concrete aspects of the problems. We term the second group of representations (Venn diagrams, trees, and novel schematic representations) *schematic* because structural aspects of the graphs symbolically represent meaningful aspects of the problem.

Novick and Hurley (2001) propose that different types of schematic devices (or "diagrams") have structural aspects or properties that determine their range of applicability. The associations we have found between use of the different types of representations and specific problem types suggest that properties of the diagrams and properties of the problem schema are being matched (though not always successfully) by participants. For example, trees seem naturally appropriate for sequential problems such as the results of multiple coin flips or successive spins of a spinner, whereas contingency tables and Venn diagrams are particularly appropriate for representing joint or compound events.

These seven identified types of external representations also differ in terms of their degree of structure. This has implications for how broadly or how narrowly the visual representation may apply. Specifically, we argue that *reorganization of the given information, pictures,* and *outcome listings* are relatively general representations that can be applied to a wide variety of problems, whereas *trees, contingency tables,* and *Venn diagrams* have more inherent structure, thus may be applicable to a more limited set of problems. Finally, the category of *novel schematic representations* is by definition not limited to any specific type of structure, thus this category of representation is also widely applicable (although any specific novel graph may have limited applicability). However, novel schematic representations seem to be used only when the problem solver encounters a very atypical or unusual problem that does not seem to plug into any familiar schema.

If we are right that *reorganization of the given information, pictures,* and *outcome listings* are very general tools, whereas the schematic devices (*trees* and *Venn diagrams*) and contingency tables are more limited in scope of application because they have more constrained structures, then those variations in scope of application ought to show up in our data. Calculating the average percentage of problems for which each type of representation was used (Table 2), provides some supporting results for this idea. The three types of representation argued here to be general ones (*reorganization, pictures,* and *outcome listings*) were used in 25% of problem solutions on average, whereas the three specific types were used in only 6% of problem solutions on average. However, the picture given by Table 3 is a bit less clear. Here it can be seen that uses of *reorganization of the given information* and *pictures* are spread across all four problem types, and *outcome listings* are used for three out of four types, whereas the more constrained types of representation *Venn diagrams* and *contingency tables* are used for only one or two types of problem. However, *trees* are used across all four problem types. Thus, except for *trees*, the predicted pattern does hold.

3.1. DO SPONTANEOUSLY SELF-GENERATED VISUAL REPRESENTATIONS HELP PROBABILITY PROBLEM SOLVERS?

The present study provides mixed evidence for the idea that external visual devices are used by probability problem solvers because they are helpful (i.e., they aid in solving the problem). For example, we found higher rates of solution success given use of *reorganization*, *outcome listings*, and *trees* for the *Combinations* problems. The finding regarding outcome listings makes sense intuitively because the essence of the combinations problems involves determining the number of outcomes in the outcomes space. Furthermore, use of trees is associated with a higher success rate (60%) for those solving the *Combinations* problems via a sequential approach, and use of trees is a relatively successful strategy (67% success rate) for the "true" *Sequential* problems. This finding seems easily interpretable, because these sequential events problems have structures that map directly onto tree diagrams. Specifically, the *Sequential* problems used here described a sequence of trials or events, each of which had several possible outcomes. Thus the process determining the outcome space can be described by a branching set of possibilities. In the corresponding tree, each node of the tree graph corresponds to one of the sequential events (e.g., one spin of the spinner), and the branches that ensue from that node represent the several possible outcomes of that uncertain event.

For the *Permutations* problems, the use of *pictures* led to a higher success rate. Our explanation for this finding is based on the point we have already made, that these particular permutation problems were atypical in that their semantic content (i.e., the real-world situation they described) describes a matching process between two sets of objects (e.g., tutors and students). Typical permutation problems the students had seen in their course consisted of problems in which a *single* set of objects is randomly ordered. Thus, it is only through a relatively sophisticated symmetry argument (requiring what is perhaps a rare or difficult insight) that the student was likely to see that one set of objects could be arbitrarily ordered, hence ignored, reducing the problem to one about ordering a single set of objects. For this reason we suspect that the increased solution success associated with use of pictures for this problem type may indicate a facilitative effect of pictures for *problem restructuring*. Such restructuring seems a necessary insight to deal with this relatively novel type of problem.

Although problem solvers showed relatively frequent use of *novel schematic representations* for the *Permutations* problems, these novel or invented types of external graphical representations were apparently not always useful, because their use did not lead to increased solution success.

Why are positive correlations between specific types of visual representations and specific problem types relatively rare in our data? As Novick and Hmelo (1994) observe, having an appropriate problem representation does not guarantee that the problem can be solved, because computational or other issues may intrude, lowering correlations between initial problem representations and solution success. Furthermore, even if graphics could be helpful, prior research shows that students are not always successful in finding correct representations for problems (Novick, 1990). Our data provide additional evidence that this is true. Additionally, some evidence from our study suggests that choosing an *inappropriate* representation might be harmful to a student's chance of successfully solving a problem. For example, for the *Permutations* problems solution success was negatively associated with use of *contingency tables* and with use of *reorganization of the given information.* The former finding can be explained because contingency tables are not appropriate for representing problems involving the ordered selection of a single set of objects. The latter finding can be explained by viewing the *reorganization* strategy as a response commonly chosen when the problem solver is confused. Thus, the negative association may indicate that when a student is stymied by a problem, rewriting the given information might be seen as a general-purpose strategy, to be tried if the student is merely casting about for any approach that might help.

We also found lower rates of success associated with use of *outcome listings* for the *Sequential* problems. Here, we suspect that the choice of representation could be based on a wrong understanding of the problem situation, or might just be an unfortunate (being potentially misleading) choice. For the *Sequential* problems used here, the listings seem to be appropriate, but they may cue (incorrect) approaches based on treating the outcome space as consisting of equally-likely outcomes.

In addition to lower rates of solution success associated with particular external visual representations, we also found a significant negative correlation ($r = -.4$, $p < .05$) between solution success and the overall use of external visual representation, suggesting that our participants were often using the external visual representations in futile solution attempts. Looking more closely at the correlations, we found that *reorganization of the given information* and *outcome listings* were marginally significantly correlated with solution failure. These two types of representations are very general tools for problem solving and participants may use these types of representations mainly when they are having trouble solving the problems. This tendency could produce such correlation with solution failure. A related possibility is that participants who are adept at solving probability problems may not need to use or report the use of a visual representation; whereas weaker problem solvers may choose very general external visual representations in the absence of insights that might allow them to select a more specific representation, leading to a lower rate of success overall for these lower-ability participants who needed to resort to external representations.

Prior research by Hegarty and Kozhevnikov (1999) suggests that there are two main types of external visual representations, schematic and pictorial. They found that the use of schematic representations was positively correlated with solution success and the use of pictorial representations was negatively correlated with success. In our data, solution success was negatively (but non-significantly) correlated with use of all types of external visuals, but the correlations were more negative with the use of *pictures*, *reorganization*, and *outcome listings* than with *contingency tables* and with the schematic representations (*Venn diagrams*, *trees*, and *novel schematic representations*), lending tentative support to the importance of this distinction.

In summary, appropriate use of a correct external visual representation may generally be helpful in problem solving, but this effect is difficult to measure in the present type of study, in which the student only generates an external representation if he or she so chooses. First, there is evidence (Lowrie & Kay, 2001; Table 7) that self-generated external visuals may be tried more often for difficult or novel problems, which have a lower solution rate in general. It should be easier to demonstrate facilitative effects of external visuals in less naturalistic studies in which the visual representations are provided to the student, or the student is explicitly asked to generate an appropriate representation before attempting to compute the answer. That type of study has been common in the literature on uses of visual representations in (non-mathematical) problem solving. However, the present data showing which types of graphical representations are spontaneously used for which types of problems may aid in designing such experimental studies and educational interventions.

Finally, facilitative effects of using visual representations may not be easy to detect in the present type of experiment because choosing the correct representation is a non-trivial task, and may require a certain level of problem understanding to accomplish (Novick, 1990; 2001; Novick & Hmelo, 1994; Novick & Hurley, 2001). With novice problem solvers, knowledge of why one representation is more appropriate than another may still be incomplete, because they have not yet mastered the appropriate schemas. Riley, Greeno and Heller (1983) found that failure to solve word problems might be caused more often by a lack of appropriate schemas than by poor arithmetic skills. They observed that problem solvers often carried out correct arithmetic procedures on incorrect representations of the problems. The negative associations we found between certain types of (presumably inappropriate) representations and solution success seem consistent with their conclusions. Interestingly, De Bock, Verschaffel, Janssens, Van Dooren, and Claes (2003) also found negative effects on solution success of asking students to generate *specified* diagrams for geometry problems, showing that not all experimenter-selected representations are useful as well (cf. Tversky, 2001; Mayer & Gallini, 1990; Scaife & Rogers, 1996), or perhaps merely indicating that not all student-generated visual representations are produced correctly, even when appropriately cued.

3.2. SOME FINAL ISSUES

One potential limitation of the present study is the question of how well the results will generalize to other populations of students. Participants in the present study were graduate students in social sciences and education, who were finishing or had recently finished an introductory course in probability and statistics. However, the participants were actually quite diverse in terms of mathematics background, including people who had not taken any mathematics in college and those who had taken a number of undergraduate or graduate mathematics and statistics courses. Thus, we believe that our results would generalize to other populations, such as high school students who had completed a similar probability course. But this question, and the question of how well the results would generalize to a wider set of probability problems, should be addressed by future research.

Another factor possibly affecting the generalizability of these results is that participants were taught probability problem solving using external visual representations, and associations between specific problem types and specific types of external graphical devices may have been implicitly or explicitly taught. Thus, the results of our study may just be a reflection of the instruction. However, the use of "novel" graphical representations by some participants indicates that although the participants were taught to use visual devices when solving probability problems, they do not always use the specific representations they were taught to use in class. This finding suggests that students believe that visual representations are useful, and try to use even representations that they have not been explicitly taught.

Studying how students solve probability problems (or any type of mathematics word problem) is a complex endeavor. One reason is that students can use any of several solution methods or strategies for many problems. Even worse, an individual student may switch approaches across similar problems, or even during the solution of a single problem. As an example of how multiple solution strategies can complicate the research process, we designed each of our probability problems with a particular formula or problem-solving schema in mind. However, in producing atypical problems for a given method, in at least one case (the atypical *Combinations* problem) we produced a problem that could easily be solved by another method entirely (treating the problem as involving sequential events), with a different appropriate external representation.

Thus, another limitation to the present study is that our manipulation of problem "typicality" was not fully successful, due to the use of alternative solution strategies by many participants. In a well-controlled experimental study with novice probability problem solvers, this problem could be avoided by introducing only one solution method or probability principle at a time. However, in a naturalistic study like the present one, where participants have been taught an array of probability problem-solving techniques, the problem of alternate strategy choice is difficult to avoid. Certainly such effects could be minimized by more careful piloting of materials in future studies.

Another issue deserving of future study is to more closely investigate the temporal process of probability problem solving. In the present study we have used a coding scheme that separates uses of external visuals for problem understanding and representation from the type of method used to compute the problem solution, but we still do not have a clear picture of the temporal stages of probability problem solving. We plan future studies that will use think-aloud protocols and structured interviews to try to distinguish sequential stages of probability problem solving, and that will examine specifically when and how external visual representations are used in the temporal process of PPS. We also plan to investigate the coordination of external visual representations with internal visualizations (cf. Scaife & Rogers, 1996). Results of these studies may bring us to a more complete understanding of the role played by visual representations and visualization skills in probability problem solving.

Results from the present study might be useful in improving instruction in the domain of probability problem solving. The very general types of external representations considered here (*pictures*, *reorganization*, and *outcome listing)* might be taught to students as general methods that can help them restructure particularly difficult problems. In contrast, the schematic representations studied here (*Venn diagrams* and *outcome trees*) and *contingency tables* could be taught as applicable to particular problem types. In line with the work of Novick and colleagues (e.g., Novick, 1990; Novick & Hmelo, 1994), abstract aspects or "features" of problems and of specific graphical representations could be taught to students, and it could be emphasized that a given representation will most likely be useful when these structural aspects of the problem and the visual device match. To some extent such principles may already be employed by instructors of probability courses, but future research should explore and better document the success of such practices.

REFERENCES

Batanero, C., Godino, J., & Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education, 12*(1).
[Online: www.amstat.org/publications/jse/v12n1/batanero.html]

Chance, B., & Garfield, J. (2002). New approaches to gathering data on student learning for research in statistics education. *Statistics Education Research Journal, 1*(1)*,* 38-41.
[Online: www.stat.auckland.ac.nz/~iase/serj/SERJ1(2).pdf]

De Bock, D., Verschaffel, L., Janssens, D., Van Dooren, W., & Claes, K. (2003). Do realistic contexts and graphical representations always have a beneficial impact on students' performance? Negative evidence from a study on modeling non-linear geometry problems. *Learning and Instruction, 13*(4), 441-463.

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

Gelman, A., & Nolan, D. (2002). *Teaching statistics: A bag of tricks.* Oxford, England: Oxford University Press.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective Probability* (pp. 129-161). Somerset, NJ: John Wiley & Sons.

Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice.* New York: Cambridge University Press.

Hall, V., Bailey, J., & Tillman, C. (1997). Can student-generated illustrations be worth ten thousand words? *Journal of Educational Psychology*, *89*(4), 677-681.

Hampson, P., & Morris, P. (1990). Imagery and working memory. In P. J. Hampson, D. E. Marks, & J. T. E. Richardson (Eds.), *Imagery: Current developments* (pp. 78-102). New York: Routledge.

Hannafin, R. D., Burruss, J. D., & Little, C. (2001). Learning with dynamic geometry programs: Perspectives of teachers and learners. *The Journal of Educational Research, 94*(3), 132-144.

Hannafin, R. D., & Scott, B. (1998). Identifying critical learner traits in a dynamic computer-based geometry program. *The Journal of Educational Research, 92*(1), 3-12.

Hegarty, M., & Just, M. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*(6), 717-742.

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

Hollebrands, K. (2003). High school students' understandings of geometric transformations in the context of a technological environment. *Journal of Mathematical Behavior, 22*(1), 55-72.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge; New York: Cambridge University Press.

Kaufmann, G. (1990). Imagery effects on problem solving. In P. J. Hampson, D. E. Marks, & J. T. E. Richardson (Eds.), *Imagery: Current developments* (pp. 169-197). New York: Routledge.

Keeler, C. M., & Steinhorst, R. K. (2001). A new approach to learning probability in the first statistics course. *Journal of Statistics Education*, *9*(3). [Online: www.amstat.org/publications/jse/v9n3/keeler.html]

Kintsch, W., & Greeno, J.G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*(1), 109-29.

Koedinger, K. R., & Anderson, R. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, *8*, 30-43.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, *6*(1), 59-98.

Konold, C. (1995). Confessions of a coin flipper and would-be instructor. *The American Statistician*, *49*(2), 203-209.

Konold, C. (1996). Representing probabilities with pipe diagrams. *The Mathematics Teacher*, *89*(5), 378-384.

Konold, C., Pollatsek, A., Well, A. D., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education*, *24*, 392-414.

Kozhevnikov, M., Hegarty, M., & Mayer, R. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, *20*(1), 47-77.

Lehrer, R., & Schauble, L. (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching*, *35*(1), 3-25.

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

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.

Mayer, R. (1989). Systemic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology*, *81*(2), 240-246.

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

Mayer, R. (2001). *Multimedia learning*. Cambridge: The Cambridge University Press.

Mayer, R., & Anderson, R. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*(4), 484-490.

Mayer, R., & Anderson, R. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*(4), 444-452.

Mayer, R., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715-726.

Mayer, R., Mautone, P., & Prothero, W. (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology*, *91*(4), 171-185.

Molitor, S., Ballstaedt, S. P., & Mandl, H. (1989). Problems in knowledge acquisition from text and pictures. In H. Mandl & J. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 3-35). North-Holland: Elsevier Science Publishers.

Nemirovsky, R. (1994). On ways of symbolizing: The case of Laura and the velocity sign. *Journal of Mathematical Behavior*, *13*(4), 389-422.

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

Novick, L. (2001). Spatial diagrams: Key instruments in the toolbox for thought. In D. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory*, Volume 40 (pp. 279-325). San Diego, CA: Academic Press.

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

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

O'Connell, A. A. (1993). *A classification of student errors in probability problem-solving*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York.

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

O'Connell, A., & Corter, J. E. (1993, April). Student misconceptions in probability problem-solving. Paper presented at annual meeting of the American Educational Research Association, Atlanta, GA.

Penner, D. E., Giles, N. D., Lehrer, R., & Schauble, L. (1996). Building functional models: Designing an elbow. *Journal of Research in Science Teaching*, *34*(2), 125-143.

Pollatsek, A., Well, A., Konold, C., Hardiman, P., and Cobb, G. (1987). Understanding conditional probabilities. *Organizational Behavior and Human Decision Processes*, *40*(2), 255-269.

Presmeg, N. C. (1986). Visualization in high school mathematics. *For the Learning of Mathematics – An International Journal of Mathematics Education*, *6*(3), 42-46.

Qin, Y., & Simon, H. (1995). Imagery and mental models in problem solving. In B. Chandrasekaran, J. Glasgow, & N. H. Narayanan (Eds.), *Diagrammatic reasoning* (pp. 403-434). Menlo Park, CA: The AAAI Press.

Reusser, K. (1996). From cognitive modeling to the design of pedagogical tools. In S. Vosniadou, E. De Corte, R. Glaser & H. Mandl (Eds.), *International perspectives on the design of technology–supported learning environments* (pp. 81-103). Mahwah, NJ: Erlbaum.

Riley, M., Greeno, J., & Heller, J. (1983). Development of children's problem-solving ability. In H. Ginsburg (Ed.), *The development of mathematical thinking*. New York: Academic Press.

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

Santos-Trigo, M. (1996). An exploration of strategies used by students to solve problems with multiple ways of solution. *Journal of Mathematical Behavior*, *15*(3), 263-284.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, *45*(2), 185-213.
Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, *20*(4), 457-497.
Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129-184.
Sedlmeier, P. (2000). How to improve statistical thinking: Choose the task representation wisely and learn by doing. *Instructional Science*, *28*(3), 227-262.
Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*(3), 380-400.
Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 79-112). Cambridge, MA: MIT Press.
Zahner, D. C., & Corter, J. E. (2002, April). Clinical interviewing to uncover the cognitive processes of probability problem solvers. Paper presented at the 2002 Annual Meeting of the American Educational Research Association, New Orleans, LA.

James E. Corter
Doris C. Zahner
Department of Human Development
Teachers College, Columbia University
525 West 120th Street
New York, NY 10027

APPENDIX: PROBABILITY PROBLEMS

The eight probability problems (test form A) classified by topic (problem type) and level of typicality, with observed proportion correct for each item. Test form B counterbalanced surface content of the text and level of typicality for each problem.

Topic	Typicality	Problem Text	Proportion of correct responses
Combinations	Typical	There are 10 books on Mary's bookcase. She randomly grabs 2 books to read on the bus. What is the probability that the 2 books are *Little Women* and *War & Peace*? (Both these books are on her bookshelf.)	0.115
	Atypical	There are 10 cookies in a cookie jar. Three of the cookies are chocolate chip, seven are sugar. A child blindly picks 2 cookies from the cookie jar. What is the probability that both cookies are chocolate chip?	0.731
Sequential	Typical	There are three balls in an urn. One is red, one is white, and one is blue. Jane randomly draws a ball from the urn, then replaces it, three times in all. What is the probability that she draws a red ball on all three turns?	0.462
	Atypical	Three spinners are constructed. The first spinner has 2 equal areas (colored red and blue), the second has three equal areas (red, blue, and white), and the third again has two equal areas (red and white). All three spinners are spun and the result of each spin is recorded. What is the probability of getting 'red' on all three spins?	0.423
Permutations	Typical	An apartment building has four parking spaces in front (call them A, B, C, and D). There are four apartments in the building (#1, #2, #3, and #4), and each apartment has a single occupant with a single car. Every evening, all four occupants come home and park in a randomly chosen space. What is the probability that this evening they park so that the occupant of Apt #1 is in space A, the occupant of #2 is in space B, the occupant of #3 in space C, and the occupant of #4 in space D?	0.462
	Atypical	There are four math students (Ed, Fred, Mary, Pia) waiting to be randomly matched with four math tutors (#1, #2, #3, and #4). Each tutor works one-on-one with a student. What is the probability that Ed will be matched with tutor #1, and Fred will be matched with tutor #2?	0.731
Conditional Probability	Typical	Joe applies for a state-subsidized mixed-income housing project being built in his neighborhood. If he is classified as a low-income applicant, he has a 70% chance of getting an apartment. Applicants not classified as low-income have only a 10% chance of getting an apartment. Joe believes that on the basis of the records he is submitting that he has a 40% chance of being classified as low income. What is the probability that he gets an apartment?	0.500
	Atypical	Assume that in the city of Metropolis, if a criminal defendant in fact committed the crime, he has a 70% chance of being found guilty by the jury. A defendant who is in fact innocent has a 10% chance of being found guilty by the jury. Assume that 40% of defendants who are tried in Metropolis in fact committed the crime. We meet a Metropolis defendant in prison. What is the probability that he is fact committed the crime, given that we know he was found guilty by the jury?	0.077