PRESERVICE ELEMENTARY TEACHERS AND THE FUNDAMENTALS OF PROBABILITY

CLARK DOLLARD
Metropolitan State College of Denver
cdollard@mscd.edu

ABSTRACT

This study examined how preservice elementary teachers think about situations involving probability. Twenty-four preservice elementary teachers who had not yet studied probability as part of their preservice elementary mathematics coursework were interviewed using a task-based interview. The participants' responses showed a wide variety of misconceptions about the meaning of probability. In particular, when they were asked to think about the probability of an irregularly shaped object, many participants had misconceptions about the classical and frequentist interpretations of probability. These findings suggest that instruction for preservice elementary teachers should address the meaning of probability, including the subjective, classical, and frequentist interpretations of probability.

Keywords: Statistics education research; Preservice teacher; Teacher education

1.   INTRODUCTION

Probability is a relative newcomer to the K-12 curriculum. In recent years, however, probability has become a part of the mainstream curriculum in many countries, including the United States, Great Britain, and Australia (Jones, 2005; Jones & Tarr, 2007). The *Principles and Standards for School Mathematics* document published by the National Council of Teachers of Mathematics (2000) states that students in the elementary grades should understand probability as a way of quantifying the likelihood of an event. The *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* (Franklin et al., 2007) goes further. Level A (the most basic level) of the framework for K-12 statistics education recommends that students should also have experiences using empirical data to estimate probabilities and that they should develop a sense of the Law of Large Numbers (p. 34). However, most teachers have little or no experience with probability from either their own schooling or their preservice teacher training. They are not confident of their abilities to teach probability, they think that it is difficult to teach, and they are uncomfortable with the uncertainty involved in hands-on activities involving probability (Stohl, 2005). These perceptions and attitudes cannot help but impact the effectiveness of classroom instruction in probability.

One way to improve this situation is to ensure that new teachers graduating from teacher education programs have a good understanding of the fundamental concepts of probability. Although many mathematics texts designed for preservice teachers include at least some probability, this does not mean that preservice teacher education programs are adequately preparing new teachers to teach these concepts. Greer and Mukhopadhyay (2005) report that "there is continuing evidence of insufficient preparation of teachers to teach probability" (p. 306). If probability is going to be taught in a meaningful way in the

schools, then mathematics courses for preservice teachers need to prepare teachers to teach probability.

In order to develop effective instruction in probability, teacher educators need to know what conceptions and misconceptions preservice teachers bring with them when they enter the preservice mathematics classroom (Shaughnessy, 1992). There has, however, been little or no research specifically focused on preservice elementary teachers and probability other than Canada's (2004, 2006) study of preservice elementary teachers' understanding of variation in a probability context. Because Canada's study is focused on preservice teachers' understanding of variation and this study is focused on their understanding of the meaning of probability, they are quite different. Both studies do, however, describe preservice elementary teachers' misunderstandings of basic concepts related to probability. In many ways, they complement each other. Taken together, they form a more rounded picture of how preservice elementary teachers think about probability than either study does on its own.

Other related research has focused on K-12 students' understanding of probability (e.g., Fischbein, 1975; Jones, Langrall, Thornton, & Mogill, 1999), undergraduates' understanding of probability (e.g., Konold, 1989; Shaughnessy, 1977), preservice elementary teachers' understanding of statistics (Groth & Bergner, 2005), and preservice elementary teachers' understanding of mathematics in general (e.g., Ball, 1988). There has also been a significant amount of psychological research focused on understandings and misunderstandings of probability in the adult population at large (e.g., Kahneman, Slovic, & Tversky, 1982) and in specific professions such as medicine (e.g., Gigerenzer & Edwards, 2003). None of this research, however, focused on the question of how preservice elementary teachers think about probability. The current study was designed to fill this gap in the research literature and to provide teacher educators with information about how preservice elementary teachers think about probability before they have studied probability as part of their preservice elementary course work.

2. THEORETICAL PERSPECTIVE ON PROBABILITY

Although the concept of probability should be quite familiar to readers of this journal, it is important to remember that probability is difficult to define precisely. As Gigerenzer, Hoffrage, and Kleinbolting (1991) note

> Today, mathematicians, statisticians, and philosophers are still wrangling over the proper interpretation of probability: Does it mean a relative frequency, a propensity, a degree of belief, a degree of evidentiary confirmation, or yet something else? Prominent thinkers can still be found in every camp, and it would be bold unto foolhardy to claim that any interpretation had a monopoly on reasonableness. (p. 525)

Three different interpretations of probability are commonly discussed in the research literature: the classical interpretation (equal possibilities based on physical symmetry), the frequentist interpretation (observed frequencies of events), and the subjective interpretation (degrees of subjective certainty or belief). Each of these interpretations has its advantages and limitations. If students, especially future teachers, are to develop a meaningful understanding of probability, it is necessary to acknowledge these different interpretations and to explore the connections between them and the different contexts in which one or the other may be useful (Steinbring, 1991). Otherwise, the study of probability can easily become an academic counting exercise with little or no connection to the real world.

2.1. THE CLASSICAL INTERPRETATION OF PROBABILITY

Arguably the most common interpretation of probability used in the classroom is the classical interpretation, commonly referred to as "theoretical probability." The classical interpretation only applies to situations that can be described in terms of equally likely outcomes and it defines probability as the ratio of desired outcomes to total possible outcomes. It is commonly used in classroom instruction because it is easily applied to random chance devices such as dice and spinners where the definition of the sample space in terms of equally likely outcomes is relatively straightforward, and it allows instruction to avoid the uncertainty of real random events.

The classical interpretation does not, however, in and of itself, provide any means of determining whether the outcomes in a given situation actually are equally likely (Batanero, Henry, & Parzysz, 2005). This question of whether the outcomes in a given situation are equally likely is not trivial. Lecoutre (1992) describes an *equiprobability bias*, a misconception held by many adults that all possibilities in a random experiment are equally likely. Pratt (2000) and Metz (1998) noticed similar misconceptions in fifth-grade students and kindergartners respectively.

2.2. FREQUENTIST INTERPRETATION OF PROBABILITY

The frequentist interpretation of probability—commonly referred to as "experimental probability"—defines probability as the hypothetical number towards which the relative frequency of a given outcome tends to stabilize when a random experiment is repeated a large number of times under identical conditions (Batanero et al., 2005). This definition requires at least an intuitive understanding of the law of large numbers, the tendency of the relative frequency of a given outcome to stabilize over a large number of repetitions. The law of large numbers also connects the frequentist interpretation to the classical interpretation. If an experiment with equally likely outcomes is repeated a large number of times, the relative frequency of a given outcome (an estimate of the frequentist interpretation of probability) will be approximately equal to the theoretical probability (as defined by the classical interpretation).

The frequentist interpretation of probability is important because it allows the idea of probability to be applied to situations that cannot be described in terms of equally likely outcomes. It is, however, limited in that it cannot provide a probability for an event that has not already been repeated a large number of times under the same conditions. Furthermore, the frequentist approach can, in the real world, only provide estimates for a hypothetical true probability, which is assumed to exist. This is pedagogically problematic because it tends to create confusion between the abstract mathematical object, the probability, and the empirical observed frequencies that estimate the probability (Batanero et al., 2005). In spite of these conceptual difficulties, research such as Horvath and Lehrer's (1998) work with second grade students and Pratt's (2000) work with fifth graders has shown that elementary students can use the frequentist interpretation of probability in a meaningful way.

2.3. SUBJECTIVE INTERPRETATION OF PROBABILITY

The third interpretation of probability is the subjective interpretation. Although both the classical and the frequentist interpretations view probability as an objective property of the event in question, the subjective interpretation takes the point of view that the probability of an event can be different for different observers depending on the amount

of knowledge each observer has. Examples of this are common in card games where different players have different knowledge about the distribution of the cards. Subjective probability can be defined formally (Lindley, 1994) and in formal academic contexts it is often referred to as Bayesian probability. However, rigorous Bayesian definitions of probability are too abstract for the elementary classroom.

Informally, however, the subjective interpretation of probability is important for the elementary classroom because it is the most consistent with the common everyday notion of the "likelihood" of an event. Konold (1991) points out that students have developed understandings of words such as *likely* and *unlikely* that allow them to use them in common everyday discourse and they bring these understandings with them to the classroom. One of the challenges of making probability meaningful for students is to help them integrate their informal understanding of the likelihood of an event with the formal definitions they encounter in the classroom.

3. THE DESIGN OF THE STUDY

3.1. THE PARTICIPANTS

The participants in this study were undergraduate students at a state supported university in the Rocky Mountain region of the United States. They were enrolled in the first semester of a three-semester series of mathematics courses required for preservice elementary teachers at this university. Because probability is a part of the second semester course, the participants had not yet studied any probability as a part of their preservice teacher education at this institution. One hundred and fifty-five students were given the opportunity to volunteer to be interviewed and 52 of these volunteered. Twenty-six of these volunteers were interviewed. The volunteers who were not interviewed could not be contacted, had scheduling conflicts, or failed to show up for a scheduled interview. Two interviews were not included in the analysis due to technical difficulties with the recording equipment or irregularities in the interview. Thus, 24 participants were included in the final analysis.

3.2. THE RESEARCH QUESTION

The results reported here were some of the findings of a larger study focused on the question "How do preservice elementary teachers think about situations involving fundamental concepts of probability?" In this paper, I will focus on the results related to simple probability (which includes the meaning of probability), and the law of large numbers. The descriptions of the participants' thinking that are presented in this paper focus on identifying misunderstandings or gaps in understanding that need to be addressed by teacher educators in mathematics classes for preservice elementary teachers in order to prepare them to effectively teach basic concepts of probability.

3.3. THE INTERVIEW

Each participant was interviewed once for approximately 45 minutes using a task-based interview protocol. The results that are the focus of this paper came from two of the interview tasks, Task 4 and Task 5. The two questions in Task 4 (Questions 4.a and 4.b; see the results section for the specific questions and tasks) focused on the participants' understanding of the meaning of the term "probability." Task 5 had two parts focused on two different manipulatives. The first part of Task 5 (Questions 5.a.i and 5.a.ii) focused

on an ordinary six-sided die. The second part of Task 5 (Questions 5.b.i, 5.b.ii, 5.b.iii, and 5.b.iv) focused on a toy house from a popular board game.

3.4. INITIAL CODING AND ANALYSIS

The analysis in this study was based on Miles and Huberman (1994), who describe qualitative analysis as "consisting of three concurrent flows of activity: data reduction, data display, and conclusion drawing/verification" (p. 10). During the initial phases of data analysis, a rubric was developed for each question—or in some cases, each aspect of a complex question—that assigned a value from zero to five representing the degree to which a participant's response to that question was consistent with the understanding one would expect from someone who was well prepared to teach probability to elementary students. A more detailed description of this understanding is given below. Before coding began, rough working definitions for each rubric were written at levels zero, three, and five. As coding progressed, detailed definitions of the various levels of each rubric were written and modified to accurately describe the various responses given by the participants. All coding was then reviewed to ensure that it was consistent with the final definitions of the levels of the rubric.

These rubrics were then condensed down to four "response levels" for each question: strong, adequate, limited, and inadequate, which provided a framework for grouping responses with similar characteristics. Descriptions of the common characteristics of the responses at each level were then written. Generally, condensing the rubric from six levels to four was fairly straightforward. In almost all cases, either one or more levels of the rubric had not been used or some of the differences between different levels of the rubric that had seemed important in the early stages of coding seemed trivial after all of the participants responses were taken into account.

3.5. ASSUMPTIONS ABOUT WHAT TEACHERS SHOULD KNOW

The analysis focused on describing participants' thinking in ways that would be useful to mathematics teacher educators who are designing instruction in probability for preservice elementary teachers. In order to identify the participants' misunderstandings or gaps in understanding that need to be addressed in their mathematics classes, it was necessary to have a model for the type of understanding one would expect from an elementary teacher who is well prepared to teach probability to his or her students. I assumed that a well-prepared elementary teacher would understand the following concepts.

- The probability of an event is a measure of how likely the event is to occur in a given context.
- If the possible outcomes for a given situation are equally likely, the probability of an event can be defined as the ratio of the desired outcomes to the possible outcomes. However, this definition cannot be used if the possible outcomes are not equally likely.
- The probability of any repeatable event can be determined—or at least estimated—by conducting a reasonably large number of trials. In this case, the probability of the event is the ratio of the number of times the event occurs to the total number of trials. The larger the number of trials, the greater the accuracy of the estimate.

These expectations are generally consistent with the understanding of probability recommended for students in Level A (the most basic level) of the framework for statistics instruction described in the GAISE report (Franklin et al., 2007). The report

recommends that students at Level A "should understand that *probability is a measure of the chance that something will happen*" (p. 33). It also recommends that students have experiences estimating probabilities through experimentation or simulation and that they should develop a sense of the law of large numbers. The authors of the report suggest that students' experiences with probability should primarily be with "simple models based on equally likely outcomes or, at most, something based on this, such as the sum of two number cubes" (p. 34), which suggests that teachers should be familiar with the classical interpretation of probability and the concept of equally likely outcomes.

3.6. RESPONSE LEVELS

As discussed above, the initial six-level rubric was condensed down to four response levels: strong, adequate, limited, and inadequate. These response levels provided a framework for describing individual participants' responses to individual interview questions and their understanding of the particular concepts addressed by each question. Due to the variation in the types of misunderstanding seen in the participants' responses, the data do not support the use of the response levels to describe a participant's understanding of probability in general. Although the specific definitions of the various response levels unavoidably varied from question to question, every effort was made to keep response levels consistent with the descriptions given below.

Strong A "strong" response indicates an understanding that is consistent with the understanding of a well-prepared teacher described above. A "strong" response indicates that the participant understands the concepts addressed by a particular question and is comfortable using them to think about situations involving probability. Specifically, the participant can recognize situations to which these ideas apply and voluntarily (without prompting) use them to think about the situation.

Adequate An "adequate" response indicates an understanding that is generally but not entirely consistent with the understanding of a well-prepared teacher. The response may include statements that indicate minor misunderstanding or lack of understanding, the participant may have trouble explaining his or her thinking, or he or she may need prompting before applying appropriate thinking to a given situation. However, the participant's understanding seems to be close enough to a strong understanding to be viable. In other words, his or her understanding of a particular concept appears to be strong enough to be useful for solving problems or building understanding of concepts for which this concept is a foundation.

Limited A "limited" response includes some thinking consistent with a strong understanding of the concept addressed by the question. However, the response indicates a misunderstanding or lack of understanding that is great enough that it does not appear that the participant has a viable understanding of this concept. In other words, his or her understanding does not appear to be strong enough to be useful for solving problems or building understanding of concepts for which this concept is a foundation.

Inadequate An "inadequate" response indicates a major misunderstanding or lack of understanding of the concept addressed by the question. The response contains little or no thinking consistent with the understanding of a well-prepared teacher described above. The response shows little or no understanding that would be useful for solving problems or building understanding of concepts for which this concept is a foundation.

3.7. REVIEW OF CODING FOR CONSISTENCY

A stratified random sample of six of the interviews was selected for review by a second coder. The reviewer worked from transcriptions of the participant's responses and the descriptions of the various levels on the original rubric. The reviewer's code was considered to agree with the original code if it placed the response in the same response level (strong, adequate, limited, or inadequate) as the original code. Working from partial transcripts of the six interviews, the reviewer's codes matched the original codes for 104 of the 125 codes reviewed. This represents an initial inter-rater reliability of 83%, which was considered to be strong considering that the reviewer was using the transcripts rather than listening to the sound files and was basing her coding solely on the written definitions of the codes.

The reviewer then listened to actual sound files of the responses for which there was disagreement, received clearer definitions of some of the rubric levels, and discussed some of the participants' responses with the original coder. Based on this information, the reviewer agreed with the original coding for 19 of the 21 remaining responses. One of the original codes had clearly been an error. The reviewer and the original coder agreed that the remaining response could be interpreted in two ways, and they agreed to disagree on the coding.

4. RESULTS

4.1. INTRODUCTION

Because this was a qualitative study, the reporting of results below focuses on giving detailed descriptions of the participants' responses. Categories of responses are described as clearly as possible and numerous examples of participants' responses are given, using their words and phrasing as much as possible. Names used to identify the participants are pseudonyms, which were selected by the participants at the start of the interview. In transcribing the interviews, I used a convention of two dots (..) to indicate a pause of about a second, three dots (…) to indicate a pause of about two to three seconds, and four dots (….) to indicate a pause of about four to five seconds. When a participant interrupted themselves, in other words, they cut one thought off to begin another, this was indicated with a semi-colon followed by two dots. For example, this is an excerpt from one of the interviews, "Uh huh, I do, because, it doesn't;.. I just think, it's not as specific."

The number of participants giving a particular type of response is provided to give the reader a sense of how common a particular type of response was among the participants in this study. However, because this was, in quantitative terms, a fairly small non-random sample, these numbers are at best rough indicators of how common a particular type of thinking may be in the population of preservice elementary teachers as a whole.

4.2. THE MEANING OF PROBABILITY

The definition of probability Question 4.a asked the participants directly, "What do we mean when we talk about the probability of an event?" Most of the participants (17 of 24) answered this question as the "likelihood" of an event or the "chance" or "chances" of an event happening. One participant defined probability as the "possibility" of an event happening and provided a reasonable example. Two other participants gave definitions that were somewhat loosely based on the classical interpretation of probability, for

example, Amanda's response was "the amount of times something will happen out of the total options available." The responses of these 20 participants were classified as adequate. As none of the participants mentioned equally likely outcomes or a large number of repetitions, none of the responses to this question were considered to be strong responses.

The remaining four participants gave responses that were considered to be "inadequate" because they were clearly different from the mathematically accepted meaning of the word "probability." Three of these participants said that probability was what could happen. For example, Amber said that "I think it means the different outcomes that can happen .. from an event or a specific happening. All the different answers you can get." The remaining participant, Jessica, said that probability was "whether or not it's going to happen." Thus, the responses to this question indicate that, although most preservice elementary teachers have an adequate intuitive sense of the meaning of the word "probability," some of them have significant misconceptions.

Probability as a number Question 4.b further explored the participants' conceptions about the meaning of probability by asking the following: "When we talk about probability in terms of a number, for example, when we say that the probability of an event is three-fourths or we say that the probability of an event is two-thirds, what does that mean?" Most of the participants (14 of 24) gave "adequate responses" to this question in that they were able to talk about the numbers in some meaningful way that was reasonably consistent with either the classical or frequentist interpretations of probability. An example of this is Lisa's response:

Lisa: It's saying like, three out of four times it might happen, or like, two out of three. How many times it could happen if you did it that many times.
Interviewer: OK, I think I, so you said three out of four times it could happen.
Lisa: Right, so if you do it four times, it'll probably happen three of those times.

Because none of the participants said anything about a large number of repetitions or equally likely outcomes, none of the responses to this question were considered to be strong responses.

Five of the participants gave responses that were considered to be "limited" because they only talked about the meaning of the numbers in terms of whether they were greater than or less than one-half. For these participants, a probability greater than one-half implied that an event is likely to happen, whereas a probability less than one-half implied that an event is unlikely to happen. For example, Bob said "That means that, say, you have three-fourths chance of something happening. You're saying that it has a greater chance of it happening than not."

The responses of the remaining five participants indicated an inadequate understanding of the meaning of probability as a number. Three of these participants were unable to give a meaningful answer. The remaining two participants merely rephrased the question in terms of chance:

Crystal: That means that there's a three-fourths chance that an event will happen or not happen, or .. same thing with two-thirds, a two-thirds chance of that (inaudible) happen.
Interviewer: When you say a three-fourths chance, can you say any more about what that number means, three-fourths chance?
Crystal: Well, if you did it like, in percentage, it would be a 75% chance. So, yeah.

As might be expected, participants who gave inadequate responses when asked about the meaning of probability did poorly when asked about the meaning of probability as a number. Of the four participants who gave inadequate responses when asked about the meaning of probability in Question 4.a, three of them gave inadequate responses to questions 4.b, whereas the fourth gave a limited response to 4.b. An additional observation from these responses was that many of the participants seemed to be more comfortable talking about probability as a percentage rather than as a fraction. Of the 24 participants who responded to this question, 11 used the term "percent" at some point in their answer.

Summary of Task 4 Although most of the participants seemed to have some understanding of the meaning of probability and the meaning of probability as a number, six of the 24 participants gave an inadequate response to at least one of these questions. An additional four could only describe the numeric meaning of probability in terms of it being greater or less than one-half. Thus, only 14 of the 24 participants gave responses that indicated an adequate understanding of the meaning of the word "probability" and the meaning of numbers expressing probability.

4.3. THE DIE AND THE HOUSE

Introduction Task 5 focused on two manipulatives, an ordinary six-sided die and a toy house. The participants were first given an ordinary six-sided die and asked the probability of rolling a two (Question 5.a.i) and the probability of rolling a number divisible by three (Question 5.a.ii). They were then given a small wooden house that is used as a playing piece in a popular board game. This house was similar in size to the die, but the house had seven sides instead of six and the sides were not all the same size and shape. The participants were asked Question 5.b.i: "What I want you to think about is, if I roll this house, what is the probability that it will land on its roof?" If the participant's answer did not at least strongly imply that the irregular shape of the house affected the probability of landing on the roof, the participant was asked Question 5.b.ii: "Does it make a difference that the sides are different sizes and shapes?"

The participants were then asked Question 5.b.iii: "Is there any way to answer this question by rolling the house?" It should be noted that in the original design of the interview, this question was intended to be a follow-up question that would only be asked if the participant did not suggest rolling the house in response to Question 5.b.i. However, because none of the participants suggested rolling the house in response to Question 5.b.i, this question (5.b.iii) became a standard interview question. If a participant's response to Question 5.b.iii indicated that they had a reasonably good sense of how to find the probability, and if their answer to Question 5.b.iii did not indicate a need to roll the house a large number of times, they were asked Question 5.b.iv: "How many times would we have to roll it?"

Because the sides of the house were not uniform and there was no reason to expect that the house would be equally likely to land on each of its sides, this task presented a situation that did not meet the necessary conditions for applying the classical interpretation of probability. The only reasonable way to determine the probability of the house landing on its roof was to use the frequentist interpretation of probability and conduct a number of trials. One of the points of interest in the analysis of this question was whether the participants recognized that, due to the irregular shape of the house, one could not apply the classical interpretation of probability and count up the number of sides. Another point of interest was whether the participant recognized the frequentist

interpretation of probability as a possible way to approach this problem and if they did, whether they recognized the need to conduct a large number of trials. Results from studies using similar tasks (Konold, 1989; Watson & Moritz, 2003) had indicated that this may be an area of confusion for many participants.

An ordinary six-sided die As expected, almost all of the participants (22 of 24) gave strong responses when asked about the probability of rolling a two on an ordinary six-sided die (Question 5.a.i). They were able to say that the probability of rolling a two was one-sixth (or an equivalent expression such as "one out of six") and they were able to give a reasonable explanation for their answer. However, two participants gave answers that were classified as inadequate. One of these was Samantha.

Samantha: I have no idea. .. I know there are, there are six sides. Six sides so, I'm not sure how to make that a probability. 'Cause if there are six sides, you have six chances that there;.. but they don't all roll a two. … That's all I can think of.

Eighteen of the participants, all of whom had given strong responses to Question 5.a.i, were also asked Question 5.a.ii: "What is the probability of rolling a number divisible by three?" Four participants who gave strong responses to questions 5.a.i were not asked question 5.a.ii simply because I neglected to ask it. Although many participants hesitated a minute before they gave their answer, sometimes obviously considering the meaning of "a number divisible by three" in this context, all of the participants who were asked this question were able to give strong responses. This question did not appear to be appreciably more difficult than Question 5.a.i, the question about the probability of rolling a two.

The toy house: Irregularity and the classical interpretation The participants were then presented with the toy house and asked the questions described above (Questions 5.b.i-5.b.iv). Samantha and Jessica, the two participants who gave inadequate responses to Question 5.a.i, the question about rolling a two on an ordinary die, were not included in the analysis of this question. Because neither of these participants seemed to have a workable understanding of the meaning of probability in the context of a common chance object with equally likely outcomes, it did not make sense to expect them to understand the meaning of probability in the context of an uncommon chance object that did not have equally likely outcomes. Thus, the responses of 22 participants were considered in the analysis of the series of questions focused on the house.

When asked to think about the probability of the house landing on its roof, 11 of these 22 participants noted the irregularity of the sides before being asked about it directly. Most of these participants were somewhat baffled by the question of how to determine the probability of an outcome when rolling an irregular object. An example of this would be Mildred's response:

Mildred: I don't know because they're all, like, the area is not the same on all sides of 'em. So maybe you'd have to figure out how, how big, like, each side is in comparison to the other sides because the bottom is the biggest part. … um … I don't really know.
Interviewer: So you're not sure, but what I'm hearing you saying is that the fact that the sides are all different areas has something to do with it.
Mildred: (Very emphatically) Yes!

If they were a little vague in their initial answer and were asked directly about the irregular shape, all of these participants clearly said that they felt that the irregular shape affected the probability. These responses were considered to be strong responses to these two questions (5.b.i and 5.b.ii).

However, the other 11 participants who were asked this question did not mention the irregularity of the sides in their initial response. All of these participants counted the sides on the house and used this as the basis for their answer. Ten of these participants said that the probability of the house landing on the roof was two-sevenths because there were seven sides to the house and two of them made up the roof. The other participant could not decide whether the probability should be two-sevenths or one-third because she perceived three categories of possible outcomes: base, side, or roof.

When these 11 participants were asked directly about the irregular shape, two of them clearly felt that the irregular shape affected the probability. These two responses were classified as adequate because they showed some awareness of the need for equally likely outcomes. Four participants gave responses that indicated a limited understanding; they acknowledged that the irregular shape might affect the probability, but their answers were somewhat vague and/or they seemed unsure about what they were saying. For example, Bertha, when asked "Does it make a difference that the sides are different sizes and shapes?" said, "Yeah, to me, it kind of does. It makes me think about it a little more." Another example of a vague response is Kelli, who said, "Probably, because, it's probably more likely to land on one of the bigger sides than like maybe these sides of the house, so that gets a little complicated, but that probably has some effect on it."

Five of the 22 participants gave inadequate responses to this set of questions; they did not feel that the irregular shape had an effect on the probability. Some of these participants were not sure, but they all at least leaned toward the shape not affecting the probability. For example, Bob said,

Bob: It could, but I'm not too sure. That's kind of like a physics kind of thing. (laughs) But I think overall if you're just talking about the probability of it happening, there is two sides on the top, and just two more places it can land, so I don't think it'd be a factor.

Another example is Amber's response to this question:

Interviewer: Does it make a difference that the sides are different sizes and shapes?
Amber: A difference in what?
Interviewer: In, how you would think about the probability, as compared to where everything was the same size and shape.
Amber: No, I don't think so, I think they're just all counting the sides.

It should be noted that none of the nine participants who gave inadequate or limited responses to this question had any trouble determining that the probability of rolling a two on a standard six-sided die was one-sixth. They could use the classical interpretation of probability to determine a probability by counting options, but they did not have an understanding of the necessary conditions for applying this interpretation. This led them to apply the classical interpretation to the house in a context for which it was not appropriate.

To summarize the findings from these two questions, 13 of the 22 participants who were asked this pair of questions clearly recognized that the irregular shape of the house would affect the probability, although two of the 13 did not mention the irregularity until they were asked about it directly. These participants seemed to have a strong or at least

adequate understanding of the need for equally likely outcomes when applying the classical interpretation of probability. Four of the 22 participants gave limited responses; they acknowledged that the irregular shape may have an effect on the probability, but their answers were vague or unsure and it did not appear that they had a clear sense of the effect of the irregular shape on the probability. The remaining five participants gave inadequate responses to this pair of questions. When asked directly about the effect of the irregular shape on the probability, either they did not know, they stated that the shape would have no effect on the probability, or they were unsure but they strongly leaned toward the shape having no effect on the probability.

The toy house and the frequentist interpretation After they responded to the pair of questions discussed above, the participants were asked whether there was any way to approach this problem—the probability of the house landing on the roof—by rolling the house (Question 5.b.iii). Based on results from studies conducted by Konold (1989) and by Watson and Moritz (2003), some participants were expected to suggest rolling the house when asked to think about the probability of the house landing on the roof (Question 5.b.i). However, none of the participants mentioned the possibility of rolling the house until they were specifically asked, so none of the responses to Question 5.b.iii were classified as strong.

Ten of the 22 participants gave adequate responses to this question. Although they did not suggest rolling the house until asked directly and some of their answers were a bit vague or unclear, all of these participants seemed to have a reasonably good idea of how to find the probability of the house experimentally. For example, Deborah was a little bit vague but she seemed to have the general idea:

Deborah: Um, you could probably roll it a bunch of times and see how many times it rolled on that, and that ..
Interviewer: That would be one way we could approach the problem?
Deborah: Yeah.

April had to think for a bit, but then she seemed to understand the basic idea of experimental probability:

April: Is there any way we could approach it by rolling it? .. Um .. (out loud) You mean, say, if we were to roll this house 100 times, how many times would it land on the roof? Sure! You could do that and just see, by your end result, if you had more roof landings than base or side landings, then you could probably gather that there's something about this house that makes it land upside down.

Ten other participants gave responses that indicated only a limited understanding. They acknowledged the possibility of finding the probability by rolling the house, but either a lack of clarity or misconceptions in their response indicated that they did not have a good sense of how to find a probability by experimentation. Due to the wide variety of misunderstandings that can be seen in these responses, it is worthwhile to examine a number of them.

Molly seemed to be having difficulty separating the classical interpretation and the question of shape from the idea of finding the probability of the house by rolling it:

Molly: Yeah, 'cause if you rolled it and it landed on the roof, then that, you could still have that as the one, …I don't know. 'Cause there are seven sides to it now, seven surfaces. What was the …
Interviewer: Like if you rolled it, say a number of times, and kept track of what happened,

would that help you figure out the probability?

Molly: It might, it you counted the roofs as one, the two sides of the roofs as one. Because they would, they would, if they were flat, then it would be the same size as the bottom, so maybe if you did that, yes.

Mildred said that she thought the house would land on the bottom every time. She did not seem to realize that this would be information about the probability of landing on the roof:

Mildred: Yeah, but it probably wouldn't be very productive because I think every time it would land on the bottom, because the bottom is biggest. So, you could do it that way, but I don't think it would be effective.

Bonnie freely admitted that she was not sure whether one could determine the probability of landing on the roof by rolling the house:

Bonnie: Um, … maybe. I don't really know (laughs). Like maybe doing testing, like rolling it a couple times and see how many times, ..
Interviewer: Would that give you any information about the probability?
Bonnie: .. It probably would.
Interviewer: But you're not sure?
Bonnie: Yeah.

George's response indicated that he had misconceptions about how one calculates the probability of an event based on a number of trials:

George: You could but if it didn't land on the roof at all, within the 7 times, because you have 7 sides, that would change the probability of it landing on the roof greatly, to say, one in a hundred .. possibly.

Lisa's response also indicated misconceptions, but her misconceptions had to do with how one would conduct a random experiment to determine the probability of an event:

Interviewer: You said you could do a test. What would that test look like?
Lisa: Maybe just see how many times it'll land on a certain side, and then if it's; .. like if it never lands on one side then maybe you could .. keep trying until it does.

These are only five of the 10 limited responses. As can be seen in the above examples, the responses given by the different participants contained a wide variety of misconceptions and degrees of uncertainty. No commonly occurring misconceptions were noted among the participants with this level of understanding. The commonality observed in these answers is a vague sense that it is possible to find the probability of an event through experimentation, but a lack of clarity about how to do so.

The responses of the remaining two participants did not indicate any real sense that it might be possible to find the probability of the house landing on the roof by rolling it. For example, Chloe said,

Chloe: Like, you .. roll it? (rolls house once) (laughs) Well, that's what happens. So, uh, no. Well, you could roll it, like physically, yes, but it probably wouldn't be .. the best way.

The findings from this question suggest that many preservice elementary teachers are not familiar with the frequentist interpretation of probability or with the idea that one can

estimate the probability of an event through experimentation. None of the participants suggested the possibility of rolling the house until they were specifically asked about it. Even when asked directly, only 10 of the 22 participants had a reasonably clear sense of the possibility of determining the probability experimentally.

The toy house and the law of large numbers Question 5.b.iv asked the participants how many times they would have to roll the toy house in order to determine the probability of landing on the roof. Because this question does not make sense if one does not have a sense that it is possible to find the probability by rolling the house, it was not asked of all participants. A total of 14 participants were asked this question, including all 10 participants who gave an adequate response to 5.b.iii, the question about whether there was any way to find the probability of landing on the roof by rolling the house. Because the decision of whether to ask this question or not had to be made by the interviewer during the course of the interview, four participants were asked this question even though later analysis determined that their responses to 5.b.iii showed only a limited understanding. The responses of these four participants will be considered separately as it is questionable whether they should have been asked the question.

Five of the 10 participants who gave adequate responses to Question 5.b.iii either implied or clearly stated a sense of the need for a large number of repetitions in their response to 5.b.iii. Four of these five were then asked the question directly and they confirmed this sense of the need for a large number of repetitions. All of these responses were considered to be strong responses to this question. A good example of this is April's response:

Interviewer: You said roll it 100 times, could you roll it 5 or 6 times, would that work, or would you need to roll it a lot of times?
April: You could roll it as many times as you want, but you're going to have the best information if you were to roll it a hundred plus times, because obviously the more times you roll it, the better representation you're going to have of how often it's going to land on the roof.

Three of the 10 participants who recognized the possibility of rolling the house to determine the probability of it landing on the roof did not mention a large number of trials until they were specifically asked, but they did show an awareness of this need when they were asked directly. These responses were considered to be adequate because they did indicate that the participant recognized the need for a large number of repetitions. An example of this can be seen in Amber's response to 5.b.iii and 5.b.iv. This response is also interesting because it implies that she has had some experience with similar situations.

Amber: Well you can see if you roll it (rolled house), well, it lands on the roof that time. Then you can keep doing that like ten times and see how many times out of that, it lands on the roof.
Interviewer: Would ten times be; .. give us a good answer, or ..
Amber: Probably not, you'd probably have to do it more.
Interviewer: You think you should do it more. OK, how many times do you think you'd have to do it to get a pretty good answer for the probability of landing on the roof?
Amber: Well, I usually do, or I've usually done, 100.

The remaining two participants who responded that it would be possible to find the probability of the house landing on the roof by rolling it gave "inadequate" responses to this question. They did not show an awareness of the need for a large number of repetitions even when asked directly. They seemed to have similar misconceptions; both

of them indicated that the number of repetitions that would be needed would depend on the results. Crystal just said, "It depends on what the results were. Like if there wasn't a solid number, then you'd want to keep going until you got a pretty solid number." This response caught me completely off guard so I did not pursue it further to find out how she would determine the appropriate number of rolls. However, when Jill showed a similar confusion I pressed a little harder. Her answer indicated confusion between finding the probability of an event and answering a yes/no question of whether an event was biased one way or the other.

Interviewer: Would ten times be enough?
Jill: I would have to roll it ten times to find out. (laughs) I don't know because, I really don't know where it would land.
Interviewer: OK, so if you did that, if you rolled it ten times to find out, what would tell you that that was a pretty good answer, and when would you think that you would have to roll it more?
Jill: If they; .. if the numbers were pretty close, so let's say it landed on the roof five times and on the bottom five times, then you'd obviously have to keep going because it, they were exactly the same. Um or even if it was like four and six, that's really close. So if it landed on the roof one time and on the bottom nine times, then we would know that that's; .. chances are that it's probably going to land on the bottom, that the chances are .. larger that it'll land on the bottom than on the top.

As mentioned above, four participants were asked about the need for a large number of repetitions even though later analysis determined that they only had a limited understanding of the possibility of finding the probability of the house landing on the roof experimentally. Not surprisingly, three of them did not indicate a sense of the need for large numbers. Their responses generally reflected their confusion about finding the probability of the house experimentally. An example of this is Lynn's response:

Lynn: If you rolled it, say, 'cause there's seven sides, seven times, um … maybe, it would land on its roof one fourteenth of the time? Does that make sense? (laughs) Like a fourteenth of the time, every fourteen times you roll it would, I guess, since there's two sides here, so it'd be one, .. one in every seven rolls. That's my guess.
Interviewer: So if I rolled it seven times and that didn't happen, would that mean that you're wrong?
Lynn: No, (mumbling) I don't think so.
Interviewer: If I rolled it a whole bunch of times, would that give us a better answer? Like if instead of rolling is 7 times, we rolled it 7000 times. ... Would that help you understand the probability?
Lynn: I think that the average would be more, so it'd probably help you out more.

Amanda, however, was an exception. Her response to 5.iii was very tenuous and did not indicate that she had a good sense of how to determine the probability experimentally but when I pushed her, she did indicate a sense that a larger number of repetitions would be desirable.

Amanda: Maybe if we rolled it like 25, 50 times and just recorded it, we would see some sort of a trend. But I don't think it would be scientific enough to make a scientific probability .. of it. But maybe that would be some sort of indication of whether or not it was fair.
Interviewer: So it would be some kind of indication of; .. if we rolled it, you said like 25, 50 times.
Amanda: Somewhat, maybe more.

Interviewer: Maybe more, would more give us a better answer?
Amanda: Probably.

To summarize these findings, a total of nine participants recognized the need for a large number of repetitions and gave adequate or strong responses to Question 5.b.iv. One of these, however, gave a limited response to Question 5.b.iii that indicated she was very uncertain about finding the probability of an event by experimentation. Of the 10 participants who had a reasonably good sense of how to find a probability experimentally in Question 5.b.iii, eight had some sense of the need for a large number of repetitions, whereas two did not. Combining the findings from Questions 5.b.iii and 5.b.iv, only eight of 22 participants who were asked these questions had both a sense of how to find the probability of an event experimentally and a sense of the law of large numbers. Because the need for a large number of repetitions is an essential part of the frequentist interpretation of probability, 14 of the 22 participants who were asked this series of questions did not appear to have an adequate understanding of the frequentist interpretation of probability.

Furthermore, of the eight participants whose responses indicated an adequate understanding of the frequentist interpretation of probability, three did not recognize the need for equally likely outcomes in question 5.b.ii. Two of these participants gave inadequate responses to 5.b.ii whereas the other gave a limited response. Thus, only five of the 24 participants in the study gave responses that indicated an adequate understanding of both the classical and frequentist interpretations of probability.

5.   IMPLICATIONS FOR INSTRUCTION

The results given above must be interpreted with caution. The participants were a relatively small non-random sample of preservice teachers from one university. Nevertheless, these results suggest that mathematics teacher educators cannot assume that preservice teachers enter their classrooms with an adequate understanding of probability. Specifically, teacher educators cannot assume that their students understand that probability is a measure of the likelihood of events. Although most (but not all) of the participants had no problem calculating simple probabilities involving a six-sided die, their responses to the questions about the toy house indicated many of them did not recognize the need for equally likely outcomes when applying the classical interpretation of probability to a situation such as this. Participants also did not seem to have a good understanding or familiarity with the process of determining a probability experimentally. None of them mentioned experimentation in the context of the toy house until they were specifically asked about it. Even when asked directly, many of them did not have a good understanding of how to determine the probability experimentally or the need for a large number of repetitions. These results suggest that mathematics classes for preservice teachers should address these issues. In particular, it should help them develop a rich understanding of the classical and frequentist interpretations of probability, the connections between them, and how they are related to the subjective interpretation of probability as the likelihood of an event.

Five mathematics textbooks that are specifically intended for preservice elementary teachers (Bassarear, 2005; Beckmann, 2005; Bennett & Nelson, 2004; Billstein, Libeskind, & Lott, 2001; O'Daffer, Charles, Cooney, Dossey, & Schielack, 2002) were examined to determine whether or not they address issues such as these. Four of these textbooks (Bassarear; Bennett & Nelson; Billstein et al.; O'Daffer et al.) have fairly similar probability chapters that are almost entirely focused on counting problems— sometimes fairly complex counting problems—based on the classical interpretation of

probability. Stohl (2005) points out that an approach such as this "relies on counting techniques, leads to a single theoretical answer to the probability of an event, and avoids realistic interpretation of that value" (p. 347). In each of these textbooks, experimental probability and the law of large numbers are mentioned only briefly, if at all, and no attempt is made to connect experimental probability, theoretical probability, and the law of large numbers. Any practice problems having to do with experimental probability are trivial. For example, students are asked to find the experimental probability of events such as "tossing a paper cup and having it land with its bottom down if it landed in this position 18 times in 150 tosses" (Bennet & Nelson, p. 523). Exercises such as this are purely computational and do not require any understanding of the meaning of probability to complete successfully. These textbooks do not address the issue of helping preservice elementary teachers create meaning for the concept of probability.

Beckmann's (2005) chapter, although it is relatively brief, does focus more on the meaning of probability and less on complex counting problems. It does not address all of the basic concepts that need to be addressed, but it does seem to be focused on helping the student construct meaning for the concept of probability. It is a step in the right direction.

These observations are not intended to point a finger at textbook writers or publishers. I assume that these books were written to meet the expectations of the teacher educators who use them. These observations do, however, indicate that there is work to be done to develop meaningful instruction in the basic concepts of probability for preservice elementary teachers. In particular, preservice elementary teachers need classroom experiences that require them to think in a meaningful way about the probability of an event that cannot be analyzed analytically. They need experiences that help them construct a meaningful understanding of the connections between theoretical probability, experimental probability, and the law of large numbers.

The question of how to do this is by no means trivial and I consider the tasks I use in my own classroom to be works in progress. However, I have been fairly pleased with the results from beginning my unit on probability by giving groups of students the toy house and asking them, "If I roll this house one thousand times, how many times do you think it would land on the roof?" If groups use the classical interpretation and say the probability is one-seventh, I offer to play a game. "We'll roll the house a bunch of times and every time it lands on the roof, you pay me a dollar. Every time it lands on one of the sides (which are much smaller than the two halves of the roof), I'll pay you a dollar." Nobody wants to play that game and the ensuing discussion brings out the need for equally likely outcomes. Usually, one or two groups start rolling the house and other groups see them doing this and get the idea that they should try gathering some data. The sample sizes the groups use are always pretty small, but the resulting variation in their answers leads to a good discussion of random variation and the law of large numbers. I am still working on improving this lesson, but starting the unit on probability by engaging with a situation that cannot be analyzed using the classical interpretation helps to interrupt students' assumption that probability is, as one interview participant put it, "just all counting the sides."

One resource that my colleagues and I have found to be a useful source of ideas for developing probability tasks for preservice elementary teachers is *Navigating through Probability in Grades 6-8* from NCTM (Bright, Frierson, Tarr, & Thomas, 2003). In particular, we have adapted one activity from this book (p. 26) for use with our students. Students working in groups are asked to perform a probabilistic experiment twenty times and keep a record of a particular outcome. For example, they may be asked to roll a toy house twenty times and keep track of the number of times the house lands on the roof or they may be asked to roll a die and keep track of the number of times they get a number

divisible by three. We have found it useful to have the students gather data for two events, one that can be analyzed theoretically (such as the die) and one that cannot (such as the house). Students are then asked to calculate—for each of the twenty rolls—the ratio of the number of times this particular outcome occurred to the total number of rolls and to make a graph showing the ratio as a function of the number of rolls. The groups then trace their graph onto an overhead transparency and these are stacked on top of each other on the overhead projector.

The image of convergence created by the stacked graphs is quite powerful and forms a good starting point for a class discussion of the meaning of probability and the law of large numbers. We have found that these are not trivial concepts for students. Assessments that ask students to describe their understanding in writing have indicated that many of them still do not understand these ideas as well as we would like. We are continuing to try to improve these activities and develop new ones.

I will add a word of caution. I have seen tacks (which can land point up or point down) and bottle caps suggested as objects for which the probability cannot be analyzed theoretically. I have used both in activities asking the students to perform a large number of trials and have found the probability to be so close to 50% that the activity did nothing to dispel the equiprobability bias (Lecoutre, 1992), the idea that all outcomes of any random event have equal probabilities. Any number of irregularly shaped but consistently manufactured objects will work—I use the house, a colleague uses hair clips—but be sure to test your object out before using it with your students.

Teaching probability in a meaningful way is not an easy task. Real probability (not contrived textbook probability) unavoidably involves uncertainty, which makes many adults, including teachers, uncomfortable (Stohl, 2005). Probability activities and demonstrations do not always go as planned and often the results are not very accurate given the number of repetitions that can reasonably be accomplished in a classroom. Computer simulations can possibly help but today's students have seen computer simulations in which all kinds of things can happen (people can fly, dragons exist, etc.) so they do not quite trust them. In my experience, students need real experiences with tangible objects in order to get a sense of probability as a real phenomenon. However, after students have had tangible experience with a given situation, computer simulations can be useful to simulate large numbers of trials to help students better understand the distribution of the outcomes.

Students' understanding of the basic concepts of probability is also difficult to assess. It is much easier to assess a student's response to a question such as "If two marbles are selected without replacement from a bag containing three red marbles and five blue marbles, what is the probability that they are both blue?" than it is to assess a student's written response to a question asking them to explain a situation involving the law of large numbers. Although the first question is much easier for an instructor to grade, I would argue that the second question shows much more about a future elementary teacher's ability to teach probability in a meaningful way.

The results of this study suggest that teacher educators cannot assume that preservice elementary teachers enter the preservice mathematics classroom with an adequate understanding of the basic concepts of probability. Although developing instruction that will enable preservice elementary teachers to construct a meaningful understanding of these concepts will not be easy, it is vital if future elementary students are to have a meaningful introduction to probability at an age when they are open to learning it. Although some may argue that such instruction is time consuming, I would respond to that argument with the following questions. What are we accomplishing by teaching preservice elementary teachers to solve problems involving contrived compound events

such as drawing marbles out of a bag if they do not understand what probability means? What are we accomplishing by teaching the subtleties of conditional probability, independence, dependence, and complex counting situations if they do not understand what probability means? What are we communicating to them about the meaning of probability when we do this? What are we communicating about the meaning of mathematics?

6. LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The participants in this study were a fairly small non-random sample from one university. Thus, the findings—in particular the proportions of participants who thought about probability in a particular way—may or may not generalize to the population of preservice elementary teachers as a whole. There is a need for more research with larger, more random samples to determine how common these ways of thinking are in the general population.

Furthermore, as mentioned above, the results reported here were part of a larger study focusing on a number of areas of probability. Because there had been virtually no research focused on preservice elementary teachers and probability, it was not clear when this study was designed that the questions discussed in this article would be as rich and interesting as they were. Now that the misunderstandings described above have been identified as possible areas of concern, there is a need for more qualitative research focused on a deeper understanding of preservice elementary teachers' thinking about these concepts.

Finally, this study focused on preservice elementary teachers' thinking about probability before they had studied probability as a part of their university coursework. There is a need for research that examines preservice elementary teachers' thinking about these concepts after they have completed their university mathematics coursework. In particular, there is a need to examine the effects of different types of instruction on preservice elementary teachers' thinking in these areas.

REFERENCES

Ball, D. L. (1988). *Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education.* Unpublished doctoral dissertation, Michigan State University, East Lansing, MI.

Bassarear, T. (2005). *Mathematics for elementary school teachers* (3rd ed.). Boston: Houghton Mifflin.

Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 15–37). New York: Springer.

Beckmann, S. (2005). *Mathematics for elementary teachers.* Boston: Pearson Education.

Bennett, A. B., Jr., & Nelson, L. T. (2004). *Mathematics for elementary teachers: A conceptual approach* (6th ed.). Boston: McGraw-Hill.

Billstein, R., Libeskind, S., & Lott, J. W. (2001). *A problem solving approach to mathematics for elementary teachers* (7th ed.). Boston: Addison Wesley.

Bright, G. W., Frierson, D., Jr., Tarr, J. E., & Thomas, C. (2003). *Navigating through probability in grades 6-8.* Reston, VA: National Council of Teachers of Mathematics.

Canada, D. L. (2004). *Elementary preservice teachers' conceptions of variation.* Unpublished doctoral dissertation. Portland State University.
 [Online: www.stat.auckland.ac.nz/~iase/publications/dissertations/04.Canada.Dissertation.pdf]

Canada, D. L. (2006). Elementary pre-service teachers' conceptions of variation in a probability context. *Statistics Education Research Journal, 5*(1), 36–63.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ_5%281%29_Canada.pdf]

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children.* Dordrecht-Holland: D. Reidel.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Schaeffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) Report: A pre-K-12 curriculum framework.* Alexandria, VA: American Statistical Association.
[Online: http://www.amstat.org/education/gaise/GAISEPreK-12_Full.pdf]

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal*, *327*(7417), 741–744.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*(4), 506–528.

Greer, G., & Mukhopadhyay, S. (2005). Teaching and learning the mathematization of uncertainty: Historical, cultural, social and political contexts. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 297–324). New York: Springer.

Groth, R. E., & Bergner, J. A. (2005). Pre-service elementary school teachers' metaphors for the concept of statistical sample. *Statistics Education Research Journal, 4*(2), 27–42.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ4%282%29_groth_bergner.pdf]

Horvath, J. K., & Lehrer, R. (1998). A model-based perspective on the development of children's understanding of chance and uncertainty. In S. P. Lajoie (Ed.), *Reflections in statistics: Learning, teaching, and assessment in grades K-12* (pp. 121–148), Mahwah, NJ: Lawrence Erlbaum.

Jones, G. A. (Ed.). (2005). *Exploring probability in school: Challenges for teaching and learning*. New York: Springer.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education*, *30*, 487–519.

Jones, D. L., & Tarr, J. E. (2007). An examination of the levels of cognitive demand required by probability tasks in middle grades mathematics textbooks. *Statistics Education Research Journal, 6*(2), 4–27.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ6%282%29_Jones_Tarr.pdf]

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

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

Konold, C. (1991). Understanding students' beliefs about probability. In E. von Glasersfeld (Ed.), *Constructivism in mathematics education* (pp. 139–156). Dordrecht, The Netherlands: Kluwer.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, *23*(6), 557–568.

Lindley, D. (1994). Foundations. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 3–15). Chichester, UK: John Wiley & Sons.

Metz, K. E. (1998). Emergent understanding and attribution of randomness: Comparative analysis of reasoning of primary grade children and undergraduates. *Cognition and Instruction*, *16*(3), 285–365.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage Publications.

O'Daffer, P., Charles, R., Cooney, T., Dossey, J., & Schielack, J. (2002). *Mathematics for elementary school teachers* (2nd ed.). Boston: Addison Wesley.

Pratt, D. (2000). Making sense of the total of two dice. *Journal for Research in Mathematics Education*, *31*(5), 602–625.

Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small-group, activity-based, model building approach to introductory probability at the college level. *Educational Studies in Mathematics*, *8*(3), 295–316.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.) *Handbook of research on mathematics teaching and learning* (pp. 465–494). New York: Macmillan.

Steinbring, H. (1991). The theoretical nature of probability in the classroom. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education*. Dordrecht, The Netherlands: Kluwer.

Stohl, H. (2005). Probability in teacher education and development. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 297–324). New York: Springer.

Watson, J. D., & Moritz, J. B. (2003). Fairness of dice: A longitudinal study of students' beliefs and strategies for making judgments. *Journal for Research in Mathematics Education*, *34*(4), 270–304.

CLARK DOLLARD
Metropolitan State College of Denver
Department of Mathematical and Computer Sciences
Campus Box 38, P.O. Box 173362
Denver, CO 80217
USA