Journal of Mathematics Education at Teachers College

Spring – Summer 2010
Inaugural Issue

A Century of Leadership in Mathematics and its Teaching

TABLE OF CONTENTS

TABLE OF CONTENTS, continued

Notes from the Classroom

Other

The *Journal of Mathematics Education at Teachers College* is a publication of the Program in Mathematics and Education at Teachers College Columbia University in the City of New York.

This issue's cover and those of future issues will honor past and current contributors to the Teachers College Program in Mathematics and Education. Photographs are drawn from the Teachers College archives and personal collections.

This issue honors NCTM 2010 Lifetime Achievement Medalist, Dr. Henry O. Pollak, who has completed 22 years as a member of the Program in Mathematics and Education at Teachers College. Dr. Pollak has contributed so much to the mathematical preparation of the Program's graduates and to the communities of mathematics and mathematics education professionals in the United States and throughout the world.

David Eugene Smith, also pictured on the front cover, was the founding professor of the Teachers College Program in Mathematics and Education. Like Dr. Pollak, Professor Smith was widely respected by both mathematicians and educators.

Aims and Scope

The *JMETC* is a re-creation of an earlier publication by the Teachers College Columbia University Program in Mathematics and Education. As a peer reviewed, semi-annual journal, it is intended to provide dissemination opportunities for writers of practice-based or research contributions to the general field of Mathematics Education. Each issue of the *JMETC* will focus upon an educational theme. Themes planned for the 2010-2011 issues are: *Teacher Education, International Education, Curriculum, Technology, and Equity*—all centered upon mathematics and its teaching. The *JMETC* will have a distinctive niche in the world of education publishing. Our readers are educators from pre K-12 and college and university levels, and from many different disciplines and job positions—teachers, principals, superintendents, professors of education, and other leaders in education.

Manuscript Submission

We seek conversational manuscripts (2500-3000 words in length) that are insightful and helpful to mathematics educators. Articles should contain fresh information, possibly research-based, that gives practical guidance readers can use to improve practice. Examples from classroom experience are encouraged. Articles must not have been accepted for publication elsewhere. All manuscripts may be submitted electronically at www.tc.edu/jmetc. This system will help keep the submission and review process as efficient as possible.

Abstract and keywords. All manuscripts must include an abstract with keywords. Abstracts describing the essence of the manuscript should not exceed 150 words. All inquiries should be sent to Ms. Krystle Hecker, P.O. Box 210, Teachers College Columbia University, 525 W. 120th St., New York, NY 10027.

More Information available online
www.tc.edu/jmetc

Call for Papers

The "theme" of the fall issue of the *Journal of Mathematics Education at Teachers College* will be *International Mathematics Education*. This "call for papers" is an invitation to mathematics education professionals, especially Teachers College students, alumni and friends, to submit articles of approximately 2500-3000 words describing research, experiments, projects, innovations, or practices related to international or comparative mathematics education. Articles should be submitted to www.tc.edu/jmetc by September 1, 2010. The fall issue's guest editor, Dr. Juliana Connelly, will send contributed articles to editorial panels for "blind review." Reviews will be completed by October 1, 2010, and final drafts of selected papers are to be submitted by November 1, 2010. Publication is expected in late November, 2010.

Call for Volunteers

This *Call for Volunteers* is an invitation to mathematics educators with experience in reading/writing professional papers to join the editorial/review panels for the Fall 2010 and subsequent issues of *JMETC*. Reviewers are expected to complete assigned reviews no later than 3 weeks from receipt of the blind manuscripts in order to expedite the publication process. Reviewers are responsible for editorial suggestions, fact and citation checking, and identification of similar works that may be helpful to contributors whose submissions seem appropriate for publication. Neither authors' nor reviewers' names and affiliations will be shared; however, editors'/reviewers' comments may be sent to contributors of manuscripts to guide further submissions without identifying the editor/reviewer.

If you wish to be considered for review assignments, please request a *Reviewer Information Form* from Ms. Hecker. Return the completed form to Ms. Krystle Hecker at JMETC@tc.columbia.edu or Teachers College, Columbia University, 525 W 120th St., Box 210, New York, NY 10027.

Looking Ahead

Anticipated themes for future issues are:

Spring 2011	Curriculum
Fall 2011	Technology
Spring 2012	Equity
Fall 2012	Leadership
Spring 2013	Psychology

TO OBTAIN COPIES OF JMETC

Journal of Mathematics Education at Teachers College
Spring–Summer 2010, Volume 1

An Analysis of a Misconception of Probability among Future Mathematics Teachers*

Patricia Jendraszek
Mercy College

Probability is receiving increased coverage in all levels of education. Knowledge of the Law of Large Numbers may be an important factor in developing a true understanding of the subject. Study results evidence a lack of awareness of this fundamental principle among teachers of mathematics. Study results also suggest that additional probability preparation in teacher education may address this deficiency and allow probability to be taught more effectively.

According to a well-known evolutionary biologist, "misunderstanding of probability may be greatest of all general impediments to scientific literacy"; yet, misunderstandings and misconceptions in probability are common (Gould, 1996, p. 7). Part of the difficulty experienced in understanding probability may be attributed to its ambiguous nature.

The nature of probability has been the subject of many philosophical debates. Philosophers of probability distinguish between classical probability, frequentism, subjectivism, Baysianism, and many other types and subtypes of probability (Galavotti, 2005; Gillies, 2000; Hacking, 1975; Mellor, 2005). These distinctions are sometimes difficult to grasp and may be irrelevant to understanding probability in any practical sense; however, for the purpose of understanding the role of probability in education, it is important to distinguish between classical probability, frequentism, and subjectivism.

Both the classical and frequentist approaches assign probabilities based on mathematical calculations. Put simply, classical probability is theoretical, or *a priori*, whereas frequentism is experimental, or *a posteriori* (Jones, Langrall, & Mooney, 2007). For example, a classicist would evaluate the probability of the outcome of a "6" on one toss of a single die as one favorable outcome out of six possible equi-probable outcomes, or 1/6. A frequentist would toss the die many times, record the results, and evaluate the probability of a "6" as the ratio of the number of times a "6" was tossed over the total number of tosses. If enough tosses were observed, the frequentist and classicist determinations would likely be nearly the same.[1] Many educators have argued that both facets of probability should be taught in order to cultivate both a theoretical and experiential understanding of the topic (Kvatinsky & Even, 2002; Steinbring, 1991).[2]

Subjectivism or subjectivist probability, by contrast, measures probability as a degree of personal belief. Confusion could result from a failure to recognize the dichotomy inherent in considering probability as an objective mathematical computation versus viewing probability as a subjective assessment of personal belief. Cultivating an awareness of this dichotomy may allow a deeper understanding of probability as well as a more reasoned approach to the subject (Greer & Mukhapadhyay, 2005; Hacking, 1975; Kvatinksy & Even, 2002).

While some misconceptions may be due to the ambiguous nature of probability, many others result from the misapplication of common mathematical or logical algorithms. Such misconceptions in probabilistic reasoning are common even among the mathematically sophisticated. Even a mathematician and educational researcher who specializes in probability has reported sometimes "falling prey" to these misconceptions (Shaughnessy, 1992).

The Research Study

Regardless of the conceptual difficulties, the study of probability has received significantly greater emphasis at the elementary, secondary, and college levels (Lutzer, Rodi, Kirkman, & Maxwell, 2007; NCTM, 2000). Whether this increased emphasis will result in greater student comprehension depends greatly upon the probability knowledge of teachers and their ability to address student misconceptions (Stohl, 2005). In order to evaluate the mathematics teachers' understanding of the concept of probability, the investigator tested future elementary, secondary, and college mathematics teachers

*This article presents certain findings from a dissertation study that tested many different types of misconceptions of probability and considered other factors that may affect understanding of probability concepts (Jendraszek, 2008).

[1]See discussion of the Law of Large Numbers below.

[2]Note that not all probabilities can be calculated using the pure classical approach; for example, a weather forecaster cannot rely solely on this approach to calculate the likelihood of rain on a particular day.

for evidence of misconceptions of probability and correlated the incidence of these misconceptions with aspects of their background, attitude, and education (Jendraszek, 2008). The method and overall results of the research study are described below; however, this article focuses on the portion of the study related to knowledge of the Law of Large Numbers.

The Law of Large Numbers and Applicable Mathematics

One common misconception considered in the study concerned ignorance of the effect of increasing sample size, called the Law of Large Numbers. Consider, for example, the probability of getting 2 heads in 3 coin tosses compared with the probability of getting 200 heads in 300 tosses. Many people see that the ratio of heads to tails is equal and assume that the probabilities are, therefore, the same.

It is simple to calculate the probability of getting 2 heads in 3 tosses because all eight possible results could be easily listed (HHH, HHT, HTH, HTT, THH, THT, TTH, and TTT). Three of these results show 2 heads; therefore, the probability of getting exactly 2 heads in 3 tosses is 3/8 or 0.375. It is considerably more difficult to enumerate each possible outcome for 300 tosses. For 300 tosses, the number of possible sequences of heads and tails is 2^{300} (2.037×10^{90}). In this case, the binomial distribution formula can be used to calculate the probability of getting 200 heads in 300 tosses.

If an event occurs N times (for example, a coin is tossed N times), then the binomial formula can be used to determine the probability of obtaining exactly r successes in the N outcomes, where π is the probability of success on any one trial:

P(r successes in N trials)

$$= \frac{N!}{r!(N-r)!} * \pi^r (1-\pi)^{N-r}$$

In the example above, N would equal 300, and r would equal 200. In any toss of a fair coin, the probability of heads, π, is ½ (50%), thus:

P(200 heads in 300 tosses)

$$= \frac{300!}{200!(300-200)!} * (0.5)^{200}(1-0.5)^{100}$$

$$= 2.041 * 10^{-9} = 0.00000000241$$

(rounded to four significant figures)

Thus, the probability of 200 heads in 300 tosses is 0.00000000241, substantially less than 0.375, the probability of 2 heads in 3 tosses.

This problem could be answered without any calculations if one were familiar with the Law of Large Numbers. The Law of Large Numbers states that as the sample size increases, the probability distribution in the sample will represent the theoretical probability distribution in the parent population more and more closely. It was first stated without proof by the Indian mathematician Brahmagupta in the 7th century and later, in the 16th century by the Italian mathematician Cardano. A rigorous proof was not published until 1713, by Jacob Bernoulli in his work Ars Conjectandi. The Law of Large Numbers was not known by its present name until Poisson first referred to it as “La loi des grands nombres,” in 1835 (Sheynin, 1968).3

In the case of the coin toss example, since the probability of heads is ½, it is clear from the Law of Large Numbers that the greater the number of tosses, the more likely it is that the number of heads will be closer to ½ of the tosses. For 300 tosses, there will likely be about 150 heads. The graphs on the following page illustrate the probability distributions of heads in coin tosses, as the sample size increases from 3 to 30 to 300 to 3000. It is evident from these graphs that the probability of deviation from the mean decreases greatly as sample size increases.

There are three problems in the study that could be answered using the Law of Large Numbers. The provenance of these problems is discussed in the next few pages under “Prior Studies.” The first asks a question similar to the example above:

> The likelihood of getting heads at least twice when tossing a coin three times is:
>
> a. smaller than
> b. equal to
> c. greater than *(Correct)*
> the likelihood of getting heads at least 200 times when tossing a coin 300 times.
> d. None of the above answers is correct.

Note that this problem compares the probability of getting 2 *or more* heads in 3 tosses with the probability of 200 *or more* heads in 300 tosses. In this case, there are four successful outcomes (HHT, HTH, THH, and HHH). The total number of possible outcomes is 8; therefore, the probability is 4/8 = 1/2. The probability of 200 or more heads in 300 tosses is more difficult to calculate. It requires 101 iterations of the binomial formula, summing of the probabilities of 200 heads, 201 heads, 202 heads, etc., up to 300 heads. This process is most easily accomplished using a computer or calculator. The formula for the calculation is:

P(r or more successes in N trials)

$$= \sum_{r}^{N} \frac{N!}{r!(N-r)!} * \pi^r (1-\pi)^{N-r}$$

P(200 or more successes in 300 trials)

$$= \sum_{r=200}^{300} \frac{300!}{r!(300-r)!} * 0.5^r (1-0.5)^{300-r}$$

[3] Afterwards, other mathematicians, notably Chebyshev and Kolmogorov, contributed to the refinement of the law, resulting in a Strong Law of Large Numbers and a Weak Law of Large Numbers. They are not really different laws, and the distinction is not really necessary to understand the general concept presented herein.

Figure 1: Probability Distribution: 3 Coin Tosses

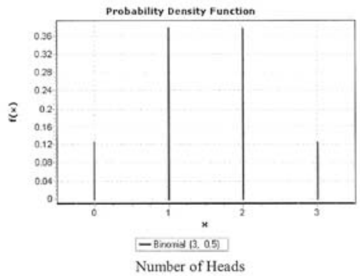


Figure 2: Probability Distribution: 30 Coin Tosses

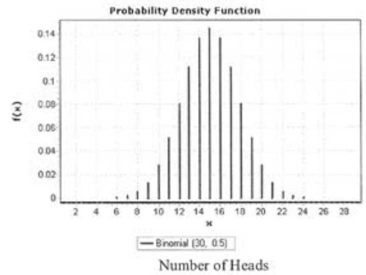


Figure 3: Binomial Distribution: 300 Coin Tosses

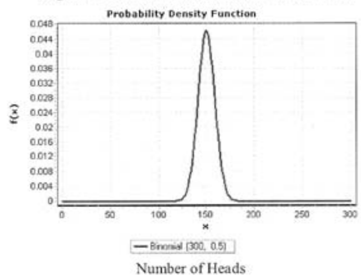
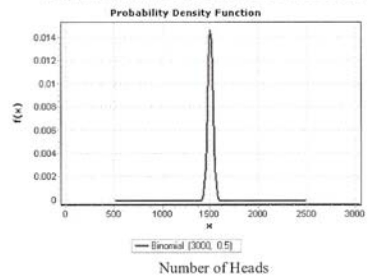


Figure 4: Binomial Distribution: 3000 Coin Tosses



The resultant probability is 4.00 * 10^{-9} (four out of a billion), significantly lower than 0.5, the probability of 2 or more heads in 3 tosses.

The second problem in the study dealing with this issue concerned hospital births:

> In a certain town there are two hospitals, a small one in which there are an average of about 20 births a day and a big one in which there are an average of about 60 births a day. The likelihood of giving birth to a boy is about 50%, the same as that of giving birth to a girl. However, there are days on which more than 50% of the babies born were boys, and there are days on which more than 50% of the babies born were girls. Both hospitals like to keep track of the days when the rate significantly deviates from 50%, favoring either male or female births (in other words, when 60% or more of the births are of either sex). Consider, for example, the number of days in which the number of **boys** born exceeded 60% in the past year. In which of the two hospitals are there likely to be more such days?
>
> a. In the big hospital there were likely more days recorded where more than 60% boys were born.
> b. In the small hospital there were likely more days recorded where more boys were born. *(Correct)*
> c. The number of days for which more than 60% boys were born is likely to be equal in the two hospitals.
> d. You cannot tell.

The Law of Large Numbers can also be applied to this question. Since the Law of Large Numbers indicates that the deviation from the expected 50% boy births would lessen as the sample size increases, it is clear that the number of days on which the percentage of boy births would exceed 60 is likely to be larger at the smaller hospital. This can be checked by evaluating the possibility of a single day of 60% or more male births at each hospital.

At the smaller hospital it would be:

$$= \sum_{r=12}^{20} \frac{20!}{20!(20-r)!} * (0.5)^r (1-0.5)^{20-r}$$

$$= 0.2517$$

(rounded to four significant figures)

At the larger hospital it would be:

$$= \sum_{r=36}^{60} \frac{60!}{60!(60-r)!} * (0.5)^r (1-0.5)^{20-r}$$

$$= 0.07750$$

(rounded to four significant figures)

The third question dealing with the effect of sample size was a sports related question:

> Assume that the Yankees have a history of winning about 60% of their games. Which is more likely?
>
> a. The Yankees win 80 out of 100 games
> b. The Yankees win 8 out of 10 games *(Correct)*
> c. (a) and (b) are equally likely.
> d. None of the above answers apply.

The Law of Large Numbers can also be used to answer this question. For this problem, it was presumed the team had an overall winning percentage of 60%. Applying this rule, it is more likely to achieve the higher winning percentage in the smaller sample of games. Thus, it is more likely for the team to win 8 out of 10 games than 80 out of 100. This can be checked using the binomial formula, assuming the probability of winning any one game, π, is 0.6:

P(winning 8 games in 10)

$$= \frac{10!}{10!(10-8)!} * (0.6)^8(1-0.6)^2$$

$$= 0.1209$$

(rounded to four significant figures)

P(winning 80 games in 100)

$$= \frac{100!}{80!(100-80)!} * (0.6)^{80}(1-0.6)^{20}$$

$$= 1.053 * 10^{-5} = 0.00001053$$

(rounded to four significant figures)

Prior Studies

Misconceptions of probability have been the subject of numerous studies by both psychologists and mathematics education researchers. Most of the pioneering work in identifying and analyzing misconceptions in probability, often referred to as "judgmental heuristics," was undertaken by cognitive psychologists.[4] Some of the earliest studies were published by Tversky and Kahneman, beginning in the early 1970s. Indeed, Tversky and Kahneman (1974) included the hospital question described above. Many of these studies addressed the incidence of probabilistic misconceptions, often in young adults and children, but few have tested misconceptions in teachers.

As early as 1992, Shaughnessy called for more study of misconceptions of probability among teachers. A recent review of research indicated that there is a deficiency of research addressing teachers' probability content knowledge (Jones, Langrall, & Mooney, 2007). Several recent studies have considered issues regarding the probability knowledge of teachers and prospective teachers (Begg & Edwards, 1999; Carnell, 1997; Carter & Capraro, 2005; Fischbein[5] & Schnarch, 1997; Koirala, 1998, 2002; Quinn, 2004; Watson, 2001; Zaslavksy, Zaslavsky, & Moore, 2001). Although illuminating, they are generally limited in scope. Further, no studies were found that considered the effect of possible variations in content knowledge among different levels of mathematics teachers or the effect of other factors such as the amount of the teachers' probability preparation.

Four of the above studies posed questions regarding sample size to college students or teachers. Tversky and Kahneman (1974) included the hospital question, Fischbein and Schnarch (1997) included the hospital question and a version of the coin toss question, Carter and Capraro (2005) included another version of the coin toss problem, and Koirala (1998) addressed the sample size issue in the context of determining experimental probability.[6] The Yankees problem was originally used by Rubel (2002).

In Tversky and Kahneman (1974), of the 95 undergraduate students who answered this question, only 21, or 22%, correctly answered that the smaller hospital was likely to have more aberrant days. Of the incorrect answers, 53, or 56%, answered that the number of such days was about the same, and 21, or 22%, answered the larger hospital. Both of the incorrect answers demonstrate ignorance of the Law of Large Numbers.

Fischbein and Schnarch (1997), whose study included a group of 18 prospective teachers along with groups of students in grades 5, 7, 9, and 11, used two problems testing the effect of sample size. The first involved a version of the hospital problem, originally used by Tversky and Kahneman (1974), above. Not one prospective teacher answered this problem correctly. All of the prospective teachers that chose to answer this question, 89% of the sample, indicated that the number of days should be about the same in both hospitals. Only one participant in the study, a ninth grader, was correct. In contrast, when asked to compare the probability of getting 2 heads in 3 coin tosses and 200 heads in 300 tosses, 50% of the prospective teachers answered correctly. Forty-four percent thought the likelihoods equal, and 6% thought it would be more likely to get 200 heads in 300 tosses.

Carter and Capraro (2005), in their online study of 108 pre-service elementary teachers, also included a question on sample size and the Law of Large Numbers. They asked for a comparison of the probability of getting 3 tails in 5 coin tosses with the probability of getting 3000 tails in 5000 coin tosses. Only 13.9% of the prospective teachers answered correctly. Of the participants, 81.5% thought both events were equally likely.

Koirala (1998) did not specifically address the sample size question but did address the Law of Large Numbers. He found, in the conduct of a study of 16 secondary teachers with backgrounds in probability, that 8 of the 16 teachers expressed unease with generating a probability based on the frequentist approach because they were not sure how many trials were necessary. Koirala used an

---

[4](Kahneman & Tversky, 1972, 1973, 1982; Tversky & Kahneman, 1971, 1973, 1974, 1982). There are numerous examples of psychologists who continued these cognitive studies (Falk & Konold, 1992; Fischbein, 1975, 1999; Konold, 1991; Konold, Pollatsek, Well, Lohmeier, & Lipson, 1993).

[5]The work of Efraim Fischbein was the inspiration for this study. There have been other studies following Fischbein investigating the development of misconceptions across age (Rubel, 2002), and across age and gender (Kennis, 2006), but these studies exclusively tested middle and high school students.

[6]In a study of 22 in-service and 12 pre-service New Zealand teachers, Begg and Edwards (1999) included a hospital question similar to the one above. It asked whether a large or small hospital would be more likely to have days with more than 80% female births. However, no results for this question were reported, although it is clear that few, if any, correct answers were given.

experiment concerning whether tacks would land on their side or on their backs when thrown. Two thought that even 1000 trials would be insufficient to generate a reliable estimate of probability, indicating that they were relatively unfamiliar with the Law of Large Numbers. Note that this knowledge is crucial to the understanding of how to use experimentation in the probability curriculum, which many would say is essential for fostering a true understanding of the concept of probability in students (Kvatinksy & Even, 2002; Steinbring, 1991; Stohl, 2005).

Gender Studies

Since this study found statistically significant gender differences in performance, it is relevant to review the research on this issue. Studies examining gender differences in probability and statistics performance are rare, although Liu and Garfield (2002) investigated gender differences in probability reasoning and misconception levels in student samples in Taiwan and in the United States, using the Statistics Aptitude Assessment test. The most striking finding was the higher level of performance by Taiwanese students of both genders in probability reasoning and freedom from misconceptions. In the US sample, women had a higher incidence of misconceptions than men, but Taiwanese men had significantly higher performance scores and fewer misconceptions than Taiwanese women, who outperformed US men on both criteria.

With respect to the related topic of statistics, in 1996, Schram published a meta-analysis of male and female achievement in college-level applied statistics classes. Schram found that undergraduate men had an advantage in statistics achievement as measured by exam scores, but women outperformed men when achievement was measured by course grade (see also, Brooks, 1987; Buck, 1987).

More research has been done on gender differences in mathematics performance in general, which has some relevance to this study. In general, males tend to out perform females on higher-level problem solving tasks which may include probability reasoning (Fennema, 1974). Gender differences on mathematical performance tests tend to increase with age (Hyde, Fennema, & Lamon, 1990) and among the gifted (Benbow & Stanley, l980, 1983; De Lisi & McGillicuddy, 2002).

Study Purpose and Method

The purpose of the study was to evaluate teacher knowledge of certain pivotal probability topics, given that such understanding has been identified as a key factor in educational results (Shulman, 1986). Recent reviews of research in probability education have specifically mentioned the shortage of studies that evaluate the probability knowledge of teachers (Jones, Langrall, & Mooney, 2007; Stohl, 2005).

The study examined misconceptions of probability among students (n=66) at a graduate school of education. All participants intended to teach mathematics at the elementary, secondary, or college levels. Most participants had teaching experience. Participants filled out a questionnaire concerning their educational background, their views of the importance of the study of probability, and their understanding of the concept of probability (including the effects of their beliefs on their view of probability). All reported some formal instruction in probability. Participants also completed a supervised 19-question multiple-choice test of probabilistic concepts and misconceptions and provided explanations for their answers.

It was necessary to assign a value to probability preparation that would permit comparison among participants. A numerical score was assigned to the level of probability coverage in high school and college based on the answers provided in the background questionnaire. These questions asked whether the participant had taken any full-semester or year-long courses in probability and for descriptions of the level of probability coverage in other courses. Numerical scores from 0 to 4 were assigned separately to high school probability preparation and college probability preparation based upon the criteria in the following table:

Criteria for Comparing Levels of Probability Preparation

Level of Coverage	**Numerical Score**
No coverage	0
Minimal or unmemorable coverage	1
Some but not substantial coverage in one or more courses.	2
One course covering probability in a substantial manner.	3
More than one course covering probability in a substantial manner.	4*

*There were no scores of 4 at the high school level in the sample.

Overall Study Results

Overall, the average rate of correct responses on all of the test problems was 56%. Average success rates on the individual test items ranged from 98% to 3%. All participants demonstrated a basic understanding of the concept of probability and could carry out simple probability calculations. Participants at all levels showed evidence of the equiprobability bias (miscounting of outcomes in a question concerning two dice), exhibited ignorance of the effect of sample size, and were seldom successful on counter-intuitive conditional probability problems.

Probability preparation, especially at the college level, was found to be strongly related to performance on the probability test. Participants who intended to teach at the secondary and college levels generally performed better than those who intended to teach at the elementary level, which may be related to the caliber and timing of their probability training.

Gender differences were observed. The overall correct response rate of males (64%) was significantly greater than that of females (52%). Males and females also tended to answer differently, based on the type of question; many of these differences were statistically significant, including those with respect to sample size.

Specific Results for Sample Size Problems

Three problems were included to test for knowledge of the effect of sample size. The first involved comparing the probabilities of outcomes from coin tosses.

> The likelihood of getting heads at least twice when tossing a coin three times is:
>
> a. smaller than
> b. equal to
> c. greater than *(Correct)*
> the likelihood of getting heads at least 200 times when tossing a coin 300 times.
> d. None of the above answers is correct.

Analysis of Correct Responses to Coin Toss Problem

Participant Level	Number Correct	Number Incorrect	n=66 Total	% Correct	% Incorrect
Doctoral	5	9	14	36%	64%
Masters	6	15	21	29%	71%
Elementary	4	27	31	13%	87%
Total	15	51	66	23%	77%

Overall, 23% of participants answered this question correctly. There were no significant differences among the levels of participants (χ^2=3.457, df=2, p=.178), but there was a tendency for higher scores at higher levels of mathematics education. The 23% accuracy rate on this problem is less than the 50% rate reported by Fischbein and Schnarch (1997) and somewhat greater than the 13.9% reported by Carter and Capraro (2005).

Most of the errors resulted from a mistaken attempt to apply proportionality. The most commonly given incorrect answer was (b). Forty-one participants, 62% of the sample, answered that the likelihood of getting heads at least twice in three tosses is the same as getting heads at least 200 times in 300 tosses. Proportionality was specifically given as the reasoning employed by 24 participants, although it may be implied in even more. Answers to this question show a fundamental lack of knowledge of sample size. Only five participants specifically mentioned it.

The second problem, the hospital question, was first used by Kahneman and Tversky (1974).

> In a certain town there are two hospitals, a small one in which there are an average of about 20 births a day and a big one in which there are an average of about 60 births a day. The likelihood of giving birth to a boy is about 50%, the same as that of giving birth to a girl. However, there are days on which more than 50% of the babies born were boys, and there are days on which more than 50% of the babies born were girls. Both hospitals like to keep track of the days when the rate significantly deviates from 50%, favoring either male or female births (in other words, when 60% or more of the births are of either sex). Consider, for example, the number of days in which the number of boys born exceeded 60% in the past year. In which of the two hospitals are there likely to be more such days?
>
> a. In the big hospital there were likely more days recorded where more than 60% boys were born.
> b. In the small hospital there were likely more days recorded where more boys were born. *(Correct)*
> c. The number of days for which more than 60% boys were born is likely to be equal in the two hospitals.
> d. You cannot tell.

Analysis of Correct Responses to Hospital Question

Participant Level	Number Correct	Number Incorrect	n=66 Total	% Correct	% Incorrect
Doctoral	3	11	14	21%	79%
Masters	7	14	21	33%	67%
Elementary	6	25	31	19%	81%
Total	16	50	66	24%	76%

Twenty-four percent of participants correctly answered this question. Mathematics masters students did slightly better than the other two groups. Doctoral and elementary students performed similarly on this problem. There were no significant differences between the levels of participants (χ^2=1.409, df=2, p=.494). The correct response rate of 24% is very close to the 22% rate reported by Tversky and Kahneman (1974). No prospective teacher in the Fischbein and Schnarch (1997) study answered this question correctly.

With respect to the 50 incorrect answers, 22 participants indicated that sample size does not matter. The greatest number of participants chose option (c) indicating that the number of days should be the same in both hospitals, often mentioning that 60% was the same proportion in both hospitals. Option (d), the second most popular answer, indicated that you cannot tell which hospital would be more likely to have days with over 60% male births. Those who answered in this manner often mentioned in their explanations that sample size does not

Results by Gender for Sample Size Questions

Question Description	Question Number	n=46 Females Correct	% F	n=20 Males Correct	% M	n=66 Total Correct	% Total	M or F
Hospital	6	6	13%	10	50%	16	24%	M**
Coin Flips	10	8	17%	7	35%	15	23%	M-
Yankees	14	5	11%	12	60%	17	26%	M**

The final column indicates when either gender outperforms the other, using the designations F or M, and whether the result is of statistical significance.

(* = Statistical Significance at .05, ** = Statistical Significance at .01, - = not statistically significant)

matter. They also often expressed that other unspecified factors were in play. For example, one participant indicated, "Size of the hospital does not matter since you are looking at percentages not numbers of births."

The third and final problem dealing with the effect of sample size involved numbers of winning games for the Yankees.

> Assume that the Yankees have a history of winning about 60% of their games. Which is more likely?
> a. The Yankees win 80 out of 100 games
> b. The Yankees win 8 out of 10 games *(Correct)*
> c. (a) and (b) are equally likely.
> d. None of the above answers apply.

In this problem, 26% of all participants answered correctly. This compares to 29% in Rubel (2002). The results evidence the greatest difference between the elementary education participants and masters and doctoral students compared to the other two sample size problems. This difference was significant (χ^2=8.004, df=2, p=.018).

Sixteen participants provided correct explanations, including references to sample size or the Law of Large Numbers, many more than the five participants who mentioned this with respect to the coin toss problem. Most of the participants who erred in this problem chose option (c), indicating that the errors resulted from a mistaken attempt to apply proportionality. All but two of the 24 participants who provided explanations for option (c) specifically mentioned that 8/10 = 80/100. The two other participants indicated that the Yankees always had a 50% chance of winning, somewhat surprising because the underlying winning percentage was given as 60%.[7]

Compared to the other sample size questions, this question had the greatest number of answers that reflected an inability to choose which event was more likely, with 20 participants choosing option (d). In 18 of the 20 answers, this was most often due to confusion between the stated 60% historical winning record and the two possibilities stated. For example, Participant 4 stated "6/10 and 60/100 are more likely," and Participant 27 indicated "8/10 = 80/100 = 80% but 60% = 6/10 = 60/100." Nearly all of the participants who incorrectly answered option (d) were also concerned with applying proportions.

Analysis of Correct Responses to Yankee Problem

Participant Level	Number Correct	Number Incorrect	n=66 Total	% Correct	% Incorrect
Doctoral	6	8	14	43%	57%
Masters	8	13	21	38%	62%
Elementary	3	28	31	10%	90%
Total	17	49	66	26%	74%

Analysis of Probability Questions and Problems by Gender

With respect to the sample size questions, males consistently outperformed females.[8] For the coin toss question, 17% of females answered correctly, while 35% of males answered correctly, but the difference was not significant (χ^2=2.461, df=1, p=.117). For the hospital question, 13% of females answered correctly, while 50% of males answered correctly, a significant difference (χ^2=10.366, df=1, p=.001). For the Yankees question, 11% of females answered correctly, while 60% of males answered correctly, also a significant difference (χ^2=17.595, df=1, p=.001).

The males in the sample did better in all questions concerning sample size. It is possible that men may be more aware of the Law of Large Numbers. This may be especially true with respect to the Yankees question, because men may be more familiar with sports statistics.

[7]This may be an attempt to use the 50-50 approach (Rubel, 2007).

[8]Males do better in some categories of misconceptions and females in others. Males do seem to outperform females in more categories and by greater amounts, but this may be in part due to the much greater concentration of males who will teach at the secondary and college levels.

Probability Preparation

The following table reflects the participants' average levels of probability coverage in high school and college and their total coverage (obtained by adding the high school and college figures) by level and gender. Not unexpectedly, the math doctoral and masters students show a much higher level of coverage in college. Elementary participants show slightly higher levels of high school preparation, which may be due to their age, since probability instruction has recently become more prevalent in grade school curricula.

Preparation by Level and Gender

Participant Category		Coverage in High School	Coverage in College	Total Coverage
Doctoral	Females	1.0	3.3	4.3
	Males	0.6	3.6	4.2
	Total	0.9	3.4	4.3
Masters	Females	0.9	3.2	4.1
	Males	1.3	3.0	4.3
	Total	1.1	3.1	4.2
Elementary	Females	1.3	1.5	2.9
	Males	1.8	2.0	3.8
	Total	1.4	1.6	3.0
Total	Females	1.2	2.2	3.4
	Males	1.2	3.0	4.2
	Total	1.2	2.5	3.7

It was found that total score was significantly related to college preparation ($r=.35$, $p=.004$). As college probability preparation increased, total score increased. It was also found that total score was significantly related to total probability preparation ($r=.29$, $p=.017$). As with college preparation, as total probability preparation increased, total score increased; however, the relationship is not as strong as that of college-level probability preparation. Total score was not significantly related to high school probability preparation alone ($r=.03$, $p=.80$).

That college-level preparation appears to have the strongest relationship with total score seems reasonable in that it is the most recent level of instruction for the participants, and is less likely to have been forgotten. It is also possible that higher-level coverage is more likely to address the areas of probability included in the probability test used in the study. The level of high school probability preparation may have less relevance to the misconceptions of probability tested, which are generally regarded as conceptually difficult.

Females and males had about the same average high school preparation scores (1.2 units); male participants had better college preparation (men had 3.0 units and women had 2.2) and also had better overall preparation (men had 4.2 units and women 3.4 units). Since college-level preparation was significantly related to performance, it is possible that the lower level of probability preparation of female participants accounts for much of the reported gender differences in these questions.

Recommendations

The primary limitations of the study concern the composition of the sample. All participants came from a single select school of education and were not randomly selected from schools of education or from the teaching population. A larger sample of randomly selected participants would generate results that could be considered more generally applicable to the entire teaching community. Nonetheless, these participants may fairly be considered representative of this institution and others like it, and, thus, the results may be considered applicable to this population.

Additional participants, more evenly matched as to gender and mathematics education level, would be needed to form reliable conclusions concerning gender-related issues. This may be difficult, because the elementary education community is predominantly female. Further studies concerning gender differences in probability concepts and performance may contribute to an understanding of how to structure probability education for teachers as well as general education in probability.

Implications for Teacher Education

Recognition of the counterintuitive nature of probability and how to avoid falling victim to probabilistic misconceptions may be an essential and as yet unaddressed component of teacher education. Specifically, knowledge of the Law of Large Numbers and the effect of sample size is an important tool in understanding probability and statistics and may be key to teaching it properly.

From the results of the study, it appears that teachers may need more instruction in order to properly recognize the effect of sample size.[9] This can be accomplished with more instruction concerning the frequentist approach, which establishes probabilities through experimentation (Stohl, 2005). This instruction could be included in either a mathematical knowledge course or a pedagogy course. Steinbring (1991) advocated a dual classical/frequentist approach to probability to develop a deeper understanding in students, but it must first be addressed in teacher

[9] The study also showed some instruction may be needed in compound probability, basic combinatorics, and conditional probability.

education in order to allow it to be taught properly (Kvatinsky & Even, 2002).

One of the unique aspects of this study was the calculation and reporting of overall results and their correlation with probability preparation. A strong correlation was found between probability preparation and performance, especially probability preparation at the college level.[10]

As suggested by Shulman (1986), Ma (1999), and others, teachers' profound understanding of their subject is crucial to the quality of their teaching. The strong relationship found between probability preparation and performance in this study suggests that additional preparation in this area may allow future teachers to gain the fundamental understanding necessary to teach this subject more effectively.

References

Begg, A., & Edwards, R. (1999). *Teachers' ideas about teaching statistics.* Presented at the combined meeting of Australian Association for Research in Education and New Zealand Association for Research in Education in Melbourne. (ERIC Document Reproduction Service No. ED455195)

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science, 210*, 1262.

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222, 1029.*

Bernoulli, J. (2006). *The art of conjecturing: Together with a letter to a friend on sets in court tennis* (Translated with an introduction and notes by E. D, Sylla). Baltimore, MD: Johns Hopkins University Press. (Originally published, *Ars conjectandi,* Basel, Switzerland, 1713)

Brooks, C. I. (1987). Superiority of women in statistics achievement. *Teaching of Psychology, 14*, 45–46.

Buck, J. L. (1987). More on superiority of women in statistics achievement: A reply to Brooks. *Teaching of Psychology, 14, 45-46.*

Carnell, L. (1997). *Characteristics of reasoning about conditional probability (Pre-service teachers).* Unpublished doctoral dissertation, University of North Carolina, Greensboro.

Carter, T. A., & Capraro, R. M. (2005). Stochastic misconceptions of preservice teachers. *Academic Exchange Quarterly*, *9*, 105-111.

De Lisi, R., & McGillicuddy-De Lisi, A. (2002). Sex differences in mathematical abilities and achievement. In R. De Lisi & A. McGillicuddy-De Lisi (Eds.), *Biology, society, and behavior: The development of sex differences in cognition* (pp. 155–182). Westport, CT: Ablex.

Falk, R., & Konold, C. (1992). The psychology of learning probability. In F. S. Gordon and S. P. Gordon (Eds.), *Statistics for the twenty-first century* (MAA Notes #26), pp. 151-164). Washington, DC: Mathematical Association of America.

Fennema, E. (1974). Mathematics learning and the sexes: A review. *Journal for Research in Mathematics Education, 5*, 126-139.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. New York: Springer.

Fischbein, E. (1999). Intuitions and schemata in mathematical reasoning. *Educational Studies in Mathematics, 38*, 11-50.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*, 96-105.

Galavotti, M. C. (2005). *Philosophical introduction to probability*. Stanford, CA: CSLI Publications.

Gillies, D. (2000). *The philosophical theories of probability*. New York: Routledge.

Gould, S. J. (1996). *Dinosaur in a haystack: Reflections on natural history*. New York: Three Rivers Press.

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

Hacking, I. (1975). *The emergence of probability*. New York: Cambridge University Press.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Jendraszek, P. (2008). *Misconceptions of probability among future teachers of mathematics* (doctoral dissertation, Columbia University), published under the title *Misconceptions of probability among future mathematics teachers*. Saarbrucken, Germany: VDM Verlag Dr. Muller.

[10]Another new concept considered in this study involved the potential effect of holding beliefs that may conflict with probability concepts. Two participants in the present study did not believe in randomness, one believed all events like tossing dice and drawing balls were controlled by fate, and another believed the events could be controlled by a deity or supreme being. While only four participants in this sample had potentially conflicting beliefs, it is possible that more of these conflicts would be present in a more diverse sample of teachers or students. This study tested for the incidence of the outcome approach, which was observed in six participants. This naïve approach to probability in which probabilities are judged right or wrong based on the outcome that occurs could well interfere with probability reasoning and may be even more common in a more diverse sample of teachers.

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. In F. Lester (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 909–955). Charlotte, NC: Information Age.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237-251. Paper presented at the Sixth International Conference on the Teaching of Statistics, Hawthorn, VIC: International Statistical Institute.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition, 11*, 123-141.

Kennis, J. (2006). *Probabilistic misconceptions across age and gender.* Unpublished doctoral dissertation, Columbia University.

Koirala, H. P. (1998). Pre-service teachers' conceptions of probability in relation to its history. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd annual conference for the International Group for the Psychology of Mathematics Education,* Stellenbosch, South Africa.

Koirala, H. P. (2002). Secondary school preservice teachers' understanding of probability: An experiment with three dice. *Focus on Learning Problems in Mathematics, 24*, 20-34.

Kolmogorov, A. (1956). *Foundations of the theory of probability* (2nd English ed.). New York: Chelsea.

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

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

Kvatinsky, T., & Even, R. (2002). *Framework for teacher knowledge and understanding about probability.* Retrieved January 7, 2010, from http://www.stat.auckland.ac.nz/~iase/publications/1/6a4_kvat.pdf.

Liu, H., & Garfield, J. B. (2002). Sex differences in statistical reasoning. *Bulletin of Educational Psychology (National Taiwan Normal University), 34*, 123-137.

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). *CBMS 2005 survey: Statistical abstract of undergraduate programs in the mathematical sciences in the United States.* Providence, RI: American Mathematical Society.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. New York: Erlbaum.

Mellor, D. H. (2005). *Probability: A philosophical introduction*. New York: Routledge.

National Council of Teachers of Mathematics (NCTM). (2000). *Standards for school mathematics.* Reston, VA: Author.

Quinn, R. J. (2004). Investigating probabilistic intuitions. *Teaching Statistics, 26*, 86-88.

Rubel, L. (2002). *Probabilistic misconceptions: Middle and high school students' mechanisms for judgments under uncertainty*. Unpublished doctoral dissertation, Columbia University.

Rubel, L. H. (2007). Middle school and high school students' probabilistic reasoning on coin tasks. *Journal for Research in Mathematics Education, 38*(5), 531-557.

Schram, C. M. (1996). A meta-analysis of gender differences in applied statistics achievement. *Journal of Educational and Behavioral Statistics, 21*(1), 55-70.

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.

Sheynin, O. B. (1968). Studies in the history of probability and statistics. XXI.: On the early history of the Law of Large Numbers. *Biometrika, 55*, 459-467.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Steinbring, H. (1991). The theoretical nature of probability in the classroom. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 135-167). Dordrecht, the Netherlands: Kluwer.

Stohl, H. (2005). Probability in teacher education and development. In G. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 345-366). New York: Springer.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*, 105-110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124.

Tversky, A., & Kahneman, D. (1982). Availability: A heuristic for judging frequency and probability. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 163-178). New York: Cambridge University Press

Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education, 4*, 305-337.

Zaslavsky, T., Zaslavsky, O., & Moore, M. (2001). Language influences on prospective mathematics teachers' understanding of probabilistic concepts. *Focus on Learning Problems in Mathematics, 23*, 23-40.