CHILDREN'S PROBABILISTIC REASONING

WITH A COMPUTER MICROWORLD

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Hollylynne Stohl Drier

B.S., The Pennsylvania State University, 1991

M.A. Ed., The College of William and Mary, 1995

May, 2000

ABSTRACT

This dissertation investigated children's probabilistic reasoning during a two-month teaching experiment. As part of the research process, the researcher developed a computer microworld environment, *Probability Explorer*, for children's explorations with probability experiments. The design of the microworld is based on a constructivist theory of learning, design of mathematical computer microworlds, and research on students' understanding of probability and rational number concepts. Two major features in the microworld include a dynamic link between numerical, graphical and iconic representations of data that are updated simultaneously during a simulation, and the ability to design experiments and assign probabilities to the possible outcomes.

The teaching experiment was conducted with three nine-year-old children. The children participated in 10 hours of teaching sessions using the microworld. Each child also participated in pre- and post- task-based interviews to assess their reasoning in probabilistic situations. Each teaching session was videotaped, and computer interactions were recorded through internal mechanisms to create a video, including children's audio, of all actions in the microworld. These tapes provided the basis for analysis and interpretation of the children's development of probabilistic reasoning while using the microworld tools.

The individual case studies detail the children's probabilistic reasoning during the pre-interview, teaching experiment, and post-interview. After extensive coding, several themes were identified and discussed in each case study. Some of the major themes included: understanding and interpretation of theoretical probability in equiprobable and

unequiproable situations; theories-in-action about the law of large numbers; and development of part-whole reasoning as it relates to probability comparisons, *a priori* predictions, and analysis of relative frequencies.

The children's development of probabilistic reasoning and their interactions with the computer tools varied during the study. The children employed different strategies and utilized different combinations of representations (e.g., numerical, graphical, iconic) to make sense of the random data to enact their own theories-in-action. The results from this study imply that open-ended microworld tools have the potential to act as agents for children's development of intuitive–based probability conceptions. Dynamically linked multiple representations and flexibility in designing experiments can facilitate an exploratory approach to probability instruction and enhance children's meaning-making activity and probabilistic reasoning.

Department of Curriculum, Instruction, and Special Education

Curry School of Education

University of Virginia

Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, “Children’s Probabilistic Reasoning with a Computer Microworld”, has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Joe Garofalo, Advisor

____
Maria Timmerman

____
Glen Bull

____
Walter Heinecke

______________________________Date

This research is dedicated to the ones I love, my family. Without their unconditional love and support, the time and energy put into this research would not have been possible. My family, especially my mother, is truly my primary source of inspiration.

ACKNOWLEDGMENTS

I would first and foremost like to thank the children who participated in my pilot and actual research studies. Their young minds and enthusiastic approach to life and learning not only helped me fulfill my research goals, but reminded me of the joy of mathematics through the eyes of a child.

My sincerest gratitude is extended to Joe Garofalo, my major professor, dissertation chair, co-teacher/researcher, colleague, and friend. His guidance, reflective feedback, and encouragement are invaluable. Appreciation is extended to Glen Bull, who intuitively knew that great things would transpire when he paired me with Paula White for the Technology Infusion Project. My experiences with Paula and her elementary students sparked my interest in children's probabilistic reasoning and the incredible mathematical thinking that was possible when young children were engaged with meaningful computer activities.

Thank you to Paula and Maria Timmerman for inspiring my work with children, and many thanks to Glen for inspiring me to develop a computer microworld that would fulfill my teaching, learning, and research goals. I also appreciate the time, effort, and reflective feedback given by Maria Timmerman and Walter Heinecke.

A final acknowledgement is certainly due to John Olive. His collegial guidance and continual support throughout my dissertation research is greatly appreciated.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Probability is an essential part of daily living. What will the weather be like today? Will my favorite sports team win tonight? What are the chances of having a disease given that a blood test was positive? These questions are all probabilistic in nature. Knowledge of probability can help citizens in reading newspapers, evaluating information given, analyzing the validity and possibility of events, and making predictions or decisions based on that information. In this regard, there is probably "no other branch of the mathematical sciences that is important for *all* students, college bound or not, as probability and statistics" (Shaughnessy, 1992, p. 466).

Probability has not been a traditional component of the K-12 mathematics curriculum in the United States (US). In contrast, according to Shaughnessy (1992), probability (and statistics) has been an integral part of the mathematics curriculum in Europe for some time. The study of probability began to slowly appear in US school curricula in the 1970s and has been given increased attention in the past decade – mostly stemming from recommendations in the National Council of Teachers of Mathematics' *Curriculum and Evaluation Standards* (1989) and updated state curriculum guides (e.g., *The Virginia Standards of Learning,* 1995). In the 1989 document, the NCTM stated that "collecting, organizing, describing, displaying, and interpreting data, as well as making decisions and predictions on the basis of that information, are skills that are increasingly important in a society based on technology and communication" (p. 54). The NCTM further advocates that a strand of probability and statistics be taught throughout the K-12 curriculum. The current draft of the new *Principles and Standards of School Mathematics* (NCTM, 1998) continues to support a strong K-12 strand of probability and statistics.

Current Status of Probability in K-12 Schools

Although no large-scale survey has been done on how and when probability is taught in K-12 classrooms, there are several indicators that probability is at least beginning to make its way into curriculum and instruction. For example, all of the newly developed NSF-funded middle school and high school curricula (e.g., *Mathematics in Context*, *Integrated Mathematics Project*) include substantial attention to probability and statistics. Other recently developed elementary curriculums (e.g., *Everyday Mathematics*) include many lessons on probabilistic concepts. There are also a variety of supplementary workbooks and packets sold through organizations such as Creative Publications and Dale Seymour that provide a wealth of probability activities for K-12 classrooms.

A Glimpse at Achievement in Probability

Since there are no surveys of the current status of probability instruction in K-12 schools, perhaps national achievement data can provide a glimpse of students' current achievement with probability concepts and skills. The most recent National Assessment of Educational Progress (NAEP, 1996) included several questions that assessed students' ability to list a sample space (all possible outcomes from a random event), determine a probability, and to use probabilistic reasoning. Students' understanding of sample space is a critical component in their ability to analyze any probabilistic situation. On the 1996 NAEP assessment, only 24% of 4th grade students were able to list all possible outcomes from picking two colored marbles from a bag (Zawojewski and Heckman, 1997). The scores improved to 59% for 8th graders on this item. Yet for listing all possible combinations for a set of objects with replacement (this item included a much larger sample space), only 13% of 8th graders and 24% of 12th graders correctly listed the entire sample space. Overall, these results indicate that students have a difficult time in constructing all possible outcomes for a given probability situation. Since this concept is a fundamental component in determining probability and using probabilistic reasoning, the other results are not surprising.

In determining probability, a little more than 50% of 4th graders could find a simple probability (1 out of N) when the total number of outcomes was given. When considered in isolation, this result seems promising. However, when asked the question in Figure 1.1, only 23% of 4th graders and 55% of 8th graders could correctly identify the probability of an event occurring. When asked to find the probability *and* explain the result, only 13% of 8th graders could correctly answer the question. It appears that although 4th grade students can determine simple probability when the total number of outcomes is given, both 4th and 8th graders demonstrate limited understanding of determining relatively simple theoretical probabilities.

There are 3 fifth graders and two sixth graders on the swim team.

Everyone's name is put in a hat and the captain is chosen by picking one name. What are the chances that the captain will be a fifth grader?

A. l out of 5

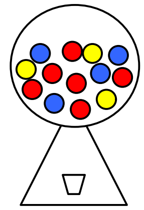
B. l out of 3

C. 3 out of 5

D.2 out of 3

Figure 1.1. Sample NAEP question to assess "determining a probability" (correct answer in bold). National Center for Educational Statistics, 1996 Assessment Mathematics–Public Release: Grade 4 available at http://nces.ed.gov/nationsreportcard/sampleq/index.shtml.

The results for probabilistic reasoning paint an even dimmer picture. For tasks similar to that in Figure 1.2, approximately 20% of 4th grade students could reasonably predict and justify their reasoning. The results for similar tasks only reached as high as 40% for 12th graders. Knowing all possible outcomes and the distribution of outcomes in the sample space does not appear to aid students in using probabilistic reasoning. I conjecture that the poor performance on these types of items is possibly due to the lack of probability instruction spiraled through K-12 mathematics, as well as students' documented difficulties in using rational numbers and proportional reasoning (Behr, Harel, Post, and Lesh, 1992; Bezuk and Bieck, 1993; Cramer, Post, and Currier, 1993). Furthermore, when probability is included in the curriculum, the type of instruction traditionally used tends to lack an experimental approach where students are asked to predict outcomes based on given information in a probabilistic situation (Shaughnessy, 1992).



The gum ball machine has 100 gum balls; 20 are yellow, 30 are blue, and 50 are red. The gum balls are well mixed inside the machine. Jenny gets 10 gum balls from this machine. What is your best prediction of the number that will be red?

Answer: __________ gum balls

Explain why you chose this number.

Figure 1.2. Sample NAEP question to assess "using probabilistic reasoning." National Center for Educational Statistics, 1996 Assessment Mathematics–Public Release: Grade 4 available at http://nces.ed.gov/nationsreportcard/sampleq/index.shtml.

Efforts to Enhance Probability Instruction

During the 1990s, there have been several initiatives to enhance K-12 probability instruction. The largest initiative occurred in the early 1990s in Australia. The Chance and Data project produced several curriculum guides and a professional development CD-ROM (Watson and Moritz, 1997) full of many activities, data sets, software, and video clips of children solving probability tasks.

In the US, Konold, Sutherland, and Lockhead (1993) developed ChancePlus, a computer-based curriculum in probability and statistics for high school and introductory college level courses. In fact, many of the activities and software programs from ChancePlus were included in the Data and Chance CD-ROM in Australia. Most recently, the ChancePlus curriculum was used with high school students at the SummerMath program at Mt. Holyoke. About 20% of the students were able to reason probabilistically before the two-week workshop, while 73% reasoned appropriately afterwards (see http://www.umass. edu/srri/chanplus.html).

There were other smaller initiatives throughout the past decade. For example, in Virginia, the State Department of Education (1996) developed a 15-hour institute for K-5 teachers to enhance their understanding of probability and statistics. This workshop has been conducted numerous times across the state over the past few years. I personally have conducted numerous 1-day workshops for teachers in elementary and middle school on how to develop probabilistic concepts. I believe that many teachers and mathematics supervisors are recognizing the need for improvement in the teaching and learning of probability. This recognition will hopefully continue to evolve into more systemic efforts to help students develop probabilistic reasoning.

Technological Influences

All of the aforementioned initiatives, at some level, utilize technological tools to help teachers and students learn probability concepts. This is not surprising since the power and availability of technology in K-12 schools has increased dramatically during the past decade. One of the most predominant uses of the technology is to quickly generate random data similar to that resulting from experiments done with physical objects such as coins and dice. Technology is then often used to create a graphical display of the data for analytical purposes.

There is almost universal agreement that technology should play a predominant role in probability and statistics education (Shaughnessy, 1992). Technology allows students to generate a large amount of data, and manipulate and represent the data in various ways that would be nearly impossible to do within the time constraints of school curriculum and instruction. Furthermore, generating large sets of data allows students to quickly experience phenomena like the law of large numbers in a meaningful way. The *law of large numbers* states that as an experiment is repeated a large number of times, the experimental distribution will approximate the theoretical distribution. Without technology, this law is often only stated as fact and must be blindly accepted by students. If probability instruction is to be more experimental, exploratory and meaningful (NCTM 1989), then there is a clear need to utilize technology tools that facilitate students' experimentation.

Purpose of this Study

Although probability has not been a traditional part of the US mathematics curriculum, recent efforts in the past decade have given probability a place in school mathematics. However, the achievement data from the 1996 NAEP (Zawojewski and Heckman, 1997) still indicates that students have difficulty in using probabilistic reasoning and understanding very basic probability concepts such as sample space. The recent efforts to improve probability instruction have made progress in localized settings. These projects often used technology tools as a critical component of the instructional process and advocated student exploration and experimentation with data.

Research on the teaching and learning of probability, to be discussed in Chapter 2, has shown that students have difficulty understanding chance happenings. Research has also shown that students possess many informal and intuitive ideas about probability concepts before instruction, and that these ideas can often lead to inappropriate probabilistic reasoning. The majority of studies conducted thus far have focused on secondary and college-age students. In the past decade, a few researchers have utilized technological tools to help students, in grades 7-12 and college, develop an understanding of probability concepts.

Recent research on the use of computer microworlds to enhance mathematics teaching and learning at the elementary level has shown promising uses to enhance children's conceptions in a variety of domains; however, to date, research has not been conducted in the domain of probability. Hence, I have developed a computer microworld intended for elementary students to explore probability concepts. By tapping into children's informal understandings and intuitive notions of chance and randomness, I conjecture that this microworld can help children develop more normative understandings of probability concepts. This dissertation research will serve two purposes: (1) to further understand children's conceptions of probability in a technological environment; and (2) to further develop the software based on children's use and their development of probabilistic reasoning.

In Chapter 2, I provide an in-depth review of the literature on students' probabilistic reasoning and instructional interventions, both with and without technology, designed to help students improve their probabilistic reasoning. This review of the literature provides a framework for my research on children's understandings of probability and their interactions while problem solving with a technological tool. In Chapter 3, I discuss the design of the software environment based on research of probabilistic misconceptions and instructional design for mathematical learning. Chapter 4 includes a description of the methodology employed for this research study. Chapters 5-7 include detailed description of the children's development of probabilistic reasoning. Chapter 8 includes a cross-case analysis and summary of results, while Chapter 9 includes a discussion of the implications from this study.

CHAPTER 2

REVIEW OF LITERATURE

Much of the early research on teaching and learning probability came from European mathematics educators, mathematicians, and psychologists. In the past two decades, US researchers have investigated students' conceptions of probability and their probabilistic reasoning abilities. Since probability concepts are increasingly making their way into US curricula, it is critical that we analyze prior research to better inform the teaching and learning process. It is my intent to review the relevant research on how concepts in probability are developed and identify misconceptions students often use in their probabilistic reasoning. I will also review prior research on instructional intervention and its effect on students' understanding of probability concepts.

Two Critical Perspectives on the Development of Probability Concepts

Two critical perspectives – the probabilistic and cognitive perspectives – appear in the literature on teaching and learning probability. Authors from each of these perspectives have contributed research and position papers on the teaching and learning of probability. A brief overview of each perspective will provide a framework for analyzing the literature and prior research on probabilistic reasoning. A third perspective, an educational one, will be discussed as part of a summary of research on instructional interventions designed to enhance students' probabilistic understandings.

The Probabilistic Perspective

Borovcnik, Bentz, and Kapadia (1991) provide a detailed discussion of the nature and study of probability from a probabilistic perspective. Their discussion includes an historical development of the study of probabilities and various probability theories. They also discuss four main approaches to the nature of probability which are relevant in the teaching and learning of school mathematics: 1) the classical view; 2) the frequentist view; 3) the subjectivist view; and 4) the structural view.

The classical view. The classical view of probability allows for the calculation of probabilities before any trial is made. In this regard, the classical view is said to be *a priori* since the ratio of favorable outcomes to all possible outcomes can be determined by assuming equal likelihood of outcomes in the sample space (Borovcnik, Bentz, & Kapadia, 1991). In this view, the entire sample space is known (assuming equiprobability of outcomes), the favorable outcomes are counted, and the probability is expressed as a ratio $\frac{favourable\ outcomes}{all\ \ possible\ outcomes}$.

For example, if a regular six-sided die were tossed, the geometrical symmetry of the object allows the probability of 1/6 to be assigned for any of the faces to land facing up. Although Borovcnik, Bentz, and Kapadia focused on outcomes with equal probability, I also include objects such as the spinner in Figure 2.1 where the possible outcomes are not equally likely but the probability of a specific color occurring can still be calculated *a priori* as 1/6 blue, 1/3 red, and 1/2 yellow.

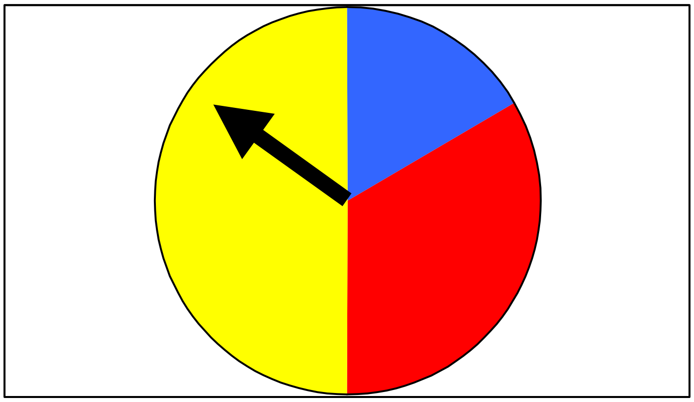


Figure 2.1. Spinner with unequiprobable areas.

The frequentist view. The frequentist view of probability uses the observed relative frequency of an event in repeated trials to determine the probability of that event. Frequentists use an *a posteriori* experimental approach to estimate probabilities after many actual trials have been done. In this view, one must repeat the trials enough times to observe all possible outcomes and to obtain enough data to establish patterns in the results so that the relative frequencies will reflect theoretical probabilities (calculated from a classical view).

For example, given the same six-sided die as mentioned above, a frequentist would roll the die many times, tally the frequencies for each of the six possible results, and use the frequencies to determine the probability for each face being rolled. Most likely, if enough trials are completed, the frequentist will obtain a probability similar to the classical one. Thus, theorems like the law of large of numbers connect the classical and frequentist view of probability. Both views result in objective measures of probability, and are thus considered part of an objectivist position.

The subjectivist view. The subjectivist view of probability includes evaluating probabilities based on personal beliefs, often based on implicit preference patterns between decisions and primary intuitions, based on naive knowledge or prior experiences (Borovcnik, Bentz, & Kapadia, 1991). To a subjectivist, probabilities provide a degree of confidence in uncertain events. Subjectivists also consider symmetry (from a classical view) and frequency (from a frequentist view) when evaluating probabilities. However, subjectivists often update their probabilities based on a "learning from experience" model (Borovcnik, Bentz, & Kapadia, 1991, p. 42).

For example, given the same six-sided die, a subjectivist might recall instances (perhaps from game-playing experiences) where it appeared difficult to roll a "six." Thus they might believe that all outcomes are not equally likely, even though the die appears to be symmetric. Upon conducting many trials, if the frequencies of all outcomes were about equal, they would adjust their theory to account for the new data. If the outcome of "six" does not appear as often as the other possible outcomes, this evidence might further ground their belief that a six is less likely to appear than the other numbers.

<u>The structural view.</u> In the fourth view of probability, the structural view, one uses the formal system of mathematical axioms, definitions, and theorems to determine probabilities – without justification of numerical values in applications. The structural view provides an underlying framework for developing concepts of probability through both an *objectivist* and *subjectivist* position. For example, Kolmogorov's axioms must be obeyed in order to work rationally with probabilities. Kolmogorov's axioms state:

1) $P(E) \geq 0$ for any event E.

    The probability of an event occurring is non-negative.

2) $P(S) = 1$ for the whole sample space S.

An event within the sample space is certain to occur in any given trial and this measure of certainty is defined as 1.

3) If E_1 , E_2 , ... is a sequence of mutually exclusive events, then

$$P(\bigcup_i E_i) = \sum_i P(E_i)$$

Probabilities of mutually exclusive events are additive.

In the die examples given for each of the other views above (classical, frequentist, and subjectivist), Kolomogorov's axioms were implicitly applied. In the examples above, every possible outcome on the die was given a probability of occurring (satisfying axiom 1), it was assumed that on any given roll one of the sides would land facing up (satisfying axiom 2), and each of the outcomes were mutually exclusive allowing the probability of a "one" and a "three" to be added to determine the probability of either occurring (satisfying axiom 3). All four views on the teaching and learning of probability do not necessarily have to be used independently. In fact, many students tend to rely on a combination of several of the views when making probabilistic decisions. A closer look at the cognitive processes underlying decisions under uncertainty will illustrate the interconnectedness of the four views.

The Cognitive Perspective

Several psychologists and mathematics educators have investigated the cognitive processes used in probabilistic reasoning. In an attempt to describe how students develop probabilistic reasoning, Piaget and Inhelder (1975/1951) proposed a three-stage cognitive development model. In the first stage (under 7 years of age), they believed a child could not distinguish between necessary (*will* happen) and possible (*may* happen) events. Here, a child’s weak concept of randomness would hinder him or her from understanding and analyzing probability events. In the second stage (up to 14 years of age), a child could recognize the distinction between necessary and possible, but could not systematically generate a list of all possible outcomes for an event. Thus, students may not be able to make a model of a probability experiment and express the likelihood of an outcome as a ratio. In the third stage (over 14 years of age), Piaget and Inhelder suggested that a person would have developed the combinatoric reasoning skills to analyze a probability situation and be able to list possible outcomes.

Piaget and Inhelder were two of the first researchers to develop a model for how students develop probability concepts. However, their work focused on *a priori* (the classical view) probability and the “spontaneous” development of concepts, and did not account for instructional intervention or a child’s social experiences. In contrast, Fischbein (1975) postulated that children, even in the Piagetian pre-operative stage, possess a pre-conceptual understanding of both relative frequencies and probabilities based on intuitive foundations, and that the transformation of these intuitions into operative concepts of probability can be mediated through *instructional intervention*.

“Fischbein’s perspective allows an exploration of intuitive foundations and precursors to probabilistic knowledge ... Fischbein is looking for the existence of partially-formed probability concepts whereas Piaget is observing the lack of completed concepts” (Hawkins & Kapadia, 1984, p.352). The intuitive foundations that Fischbein refers to stem from early social experiences with chance happenings, and are embedded in children’s thinking prior to instruction in probability. Thus, he terms these *primary* *intuitions*. The intuitive notions formed after instruction and extended social experiences are considered *secondary intuitions* (Fischbein, 1975).

Hawkins and Rapadia briefly discuss several other theories related to probability cognition: 1) information theory; 2) communication theory; and 3) Estes' (1964) stimulus sampling theory. *Information theory* presupposes that individuals continually receive and process information under the guise of uncertainty. The nature of received information is only predictable with a degree of certainty and that information is processed in a probabilistic manner. *Communication theory* suggests that children use language in a probabilistic manner since their developing language reflects the pattern, but not exact nature, of the language heard around them. Again, this theory indicates that individuals are naturally capable of processing probabilistic data from the environment. Estes' *stimulus sampling theory* models how individuals develop concepts by repeatedly sampling information from their environment (e.g., a child touching a stove-top burner several times) and making statistical and probabilistic inferences from that information (the burner is usually hot when mother is preparing food, otherwise it is usually cool). Humans employ probabilistic judgments when making sense of these sampled stimuli.

These three theories, along with Piaget's and Fischbein's theories, imply that probabilistic reasoning is a cognitive activity inherent in our students. Recognizing this, educational development of probabilistic reasoning must be closely connected with cognitive development. Since it is apparent that children do use probabilistic inferences to make daily decisions, the teaching and learning of probability should build upon on those experiences.

The probabilistic and cognitive perspectives each contribute to understanding how probabilistic reasoning is developed. Together, these perspectives provide a focused lens for which to analyze and interpret the research on students' understanding of probability concepts and the misconceptions that arise in their probabilistic reasoning. By reviewing the literature and prior research through this lens, an educational perspective on how probability concepts are best taught will emerge.

Research on Students' Probabilistic Reasoning

In the past 30-40 years, researchers from the fields of mathematics education, mathematics, and psychology have all conducted inquiries in how children use probabilistic reasoning to solve tasks. Collectively, these inquiries have investigated the thinking of children as young as pre-school age through college students. The primary purpose of these inquiries was to identify: 1) whether students reason appropriately on probability tasks; 2) what strategies students use to solve probability tasks; and 3) what causes students to employ inappropriate probabilistic reasoning. In these studies, the researchers typically administered probability tasks via paper-and-pencil tests (both multiple choice and free response) and/or individual interviews. Researchers used both quantitative and qualitative methods of inquiry and data analysis in these investigations.

The studies relating to students' probabilistic reasoning have identified common strategies that students use to solve probability tasks. These strategies often lead to inappropriate reasoning and, when compounded with primary intuitions, can result in a misunderstanding of probabilistic situations. The four most common strategies used are based on: 1) the *representativeness* of the information given; 2) the *availability* of information; 3) the assumption that events are *equiprobable;* and 4) an *outcome based* orientation for determining future outcomes (in contrast to a frequentist approach). By examining students' use of these strategies and their orientations towards solving a probability task, the researcher gained a better understanding of students' conceptions of probability. This understanding will in turn inform my own efforts in designing appropriate software, instructional methods and interventions for developing children's conceptions of probability.

Representativeness

The heuristic of representativeness was initially studied and described by Kahneman and Tversky in the early 1970s (Kahneman & Tversky, 1972; Kahneman & Tversky, 1973; Tversky & Kahneman, 1973). When someone is presented with a probabilistic situation,

> They often estimate the likelihood of an event based on how well an outcome represents some aspect of its parent population. People believe that even small samples, perhaps a single outcome, should either reflect the distribution of the parent population or mirror the process by which random events are generated. (Shaughnessy & Bergman, 1993, p. 181-182)

The notion of representativeness is at the heart of much of statistics. Researchers try to draw "random" samples from a population in such a way that the sample is representative of the population and results can be used to infer characteristics of the whole population. Although the representativeness heuristic can often result in appropriate probability judgments, applying this heuristic can cause some predictable errors in certain situations.

For example, many people might think that when flipping a coin six times, the sequence HTTHTH is more likely to occur than either HHHHTH or HHHTTT. The sequence HTTHTH might appear more likely than HHHHTH since it is more representative of the expected 50-50 distribution of heads and tails. Likewise, people might consider the sequence HTTHTH more likely than HHHTTT since the second sequence does not appear to be representative of their intuitive understanding of the random process of flipping coins. In this regard, the use of the representative heuristic is an example of subjective probability.

There are many other instances of inappropriate applications of representativeness. Many of these applications are so common and predictable they have been termed as the gambler's fallacy, base rate fallacy, law of small numbers, and conjunction fallacy. A brief discussion of each will illustrate the reasoning processes students use in probabilistic situations.

Gambler's fallacy. People fall prey to the gambler's fallacy when they believe the probability of an event occurring is related to the outcomes of previous trials. For example, a person might predict that if the last six flips of a coin resulted in HTTTTT, then a head is more likely to occur next. This belief is also called the "negative recency effect" since a person believes recent repeated results (five tails) will cause a different result (heads) to have an increased probability of appearing. A person is applying the representativeness heuristic since he or she thinks the six known tosses are not representative of the expected 50:50 distribution (connected to the law of large numbers). A person also may not fully understand that results of individual tosses are independent of one another.

Another fallacious approach often used in gambling situations is termed the positive recency effect. A person might increase their bets on a certain number (e.g., from a roll of two dice) because that number is “hot” tonight and has won several times recently. Again, the gambler is ignoring the independence of individual rolls of the dice. Past outcomes of random events can not be used as indicators of future outcomes. Of course, in this situation, the person could be utilizing a subjectivist approach and might believe that the dice are not “fair” and that, learning from experience, a particular outcome might actually be more likely to occur.

Base-rate fallacy. Another application of the representativeness heuristic occurs when students are asked to make a probability judgment based on given information. For example, Kahneman and Tversky (1973) originally posed a task where participants were given a description of a person as being male, 45, conservative, ambitious, and with no interest in political issues. Then they were asked which is more likely: a) the person is a lawyer or b) the person is an engineer. The majority of participants overwhelmingly thought the person was more likely an engineer since the description more closely represents the stereotype of an engineer. Giving participants information that this person was randomly chosen from a population of 70% lawyers and 30% engineers had little effect on their choice and most still picked the engineer as more likely. The use of representativeness in these types of situations has come to be known as the base rate fallacy.

Law of small numbers. Many students have a belief that every sample of a population must be representative of the true proportion of the population. This belief has been termed the “law” of small numbers and can affect students’ reasoning similar to the gambler's fallacy. Using the "law" of small numbers, many students will misunderstand the effect of sample size on probability situations. For example, students will often believe that when flipping coins simultaneously the result of two heads and one tail is equally as likely as the result of 200 heads and 100 tails. Tversky and Kahneman (1982) reported that the belief in small numbers is even present in the thinking of research psychologists who put too much faith in statistically significant results in samples of small sizes. Consequently, these researchers "grossly overestimated the replicability of such results" (Tversky & Kahneman, p. 8).

Conjunction fallacy. Another common intuitive-based conception of probability is that the probability of events A & B occurring simultaneously (mathematically considered the intersection of the two events) is greater than the probability of only A occurring. In fact, just the opposite is true.

$$P(A) \cap P(B) \leq P(A)$$

Students have demonstrated use of this conjunction fallacy when completing problems such as:

Which is more likely:

A) a person is 55 years old and has had a heart attack

B) a person (regardless of age) has had a heart attack.

Shaughnessy (1992) reported that statistically naive (i.e., no formal statistical education) college students predicted that choice A was more likely. Since age is typically a characteristic most people associate with heart attacks, many people consider choice A as representative of people who have had heart attacks. Other researchers (Kahneman & Tversky, 1973; Slovic, Fischhoff, & Lichtenstein, 1976) have also confirmed the existence of the conjunction fallacy in secondary and college-age students. It is important to note that the wording of the above task could mislead some students to interpret the situation as a conditional (e.g., if a person is 55, they are more likely to have a heart attack) rather than a conjunction. On a multiple-choice test (which is often the way probability tasks are posed), it is difficult to know how students interpret the questions, and, in turn, difficult to assess their reasoning.

As shown above, representativeness can lead to different approaches, some appropriate and some fallacious, to probabilistic reasoning. Students in probability problem solving also use the other heuristics of availability, equiprobability, and the outcome-based approach. A discussion of these other heuristics and the outcome-based approach will contribute to understanding the other intuitive-based misconceptions of probability.

Availability

A subjective view of probability relies on personal experiences and knowledge. Developing a subjective judgment of probability naturally relies heavily on the availability of information to an individual. For example, if you have recently been in an accident at a particular intersection, you are more likely to think that accidents occur more often at that particular intersection than someone who has driven through that intersection accident-free for 10 years. The availability of your personal experience may have biased your opinion when, in fact, your accident might have been a rare occasion rather than just another tally in a large frequency. The other person’s accident-free experience could also bias his or her opinion that the intersection is safe if in fact many accidents have occurred there. Either way, both drivers would be relying on the availability of experiences to make a judgment about the likelihood of having an accident in that intersection.

Another example of availability occurs when students need to draw upon personal knowledge as well as mathematical facility with combinatorics. When asked whether it is possible from a group of 10 people to make up more committees of eight people or of two people, Kahneman and Tversky (1973) and Shaughnessy (1977) found that most students thought there were more ways to make a two-person committee. It is actually possible to make the same number of each size committee. The researchers attribute the choice to students being able to imagine more two-person committees and lacking combinatoric reasoning for forming unique committees of eight people, thus, having a greater availability to instances of the smaller committee.

Einhorn (1982) gave an example of how availability of information can be to the advantage or disadvantage of the problem solver. If asked which was more likely to happen, dying from emphysema or by lightning, one might try to recall all the personal cases one knows of each occurring and choose that occurrence as the most likely. A person might also try to recall all the media reports of each occurrence and choose the most likely from that perspective. Either way, a person’s own experiences, imaginability, knowledge, and memory-searching strategies will contribute to their use of availability as an heuristic and give different biases for each situation, thus ultimately affecting their subjective judgment of probability according to those biases.

Equiprobable

An overwhelming majority of probability situations discussed in school curriculum are based on an assumption of equiprobability. Students often use regular six-sided die, two-colored counters, fair coins, and equal-sized marbles to pull out of an urn for an introduction to probability. The random sampling of 100 people also assumes each person has an equal chance of being chosen. Students are introduced to probability and chance, either in school or through experiences with games of chance, as random occurrences that are equally likely to occur.

There have been several research results reporting students use of an equiprobable assumption in their approach to solving probability tasks – sometimes this assumption applies and is to the advantage of the problem solver, but other times a student overgeneralizes the assumption to situations that are not equiprobable. Lecoutre (1992) originally added the assumption of equiprobable events to the list of heuristics for solving probability problems. He found that students often associated chance and luck with events being equally likely by nature. In a study with adults, he reported that a vast majority of them believed that rolling the pair of five and six (with order not mattering) and rolling two sixes from a pair of dice as equally likely since either pair could happen by chance. Madsen (1995) surveyed 13-19 year old students and found that they, too, employed the equiprobable heuristic. Given a six-sided die with one face painted black and the other five faces painted gold, 47% of the students used the equiprobable heuristic correctly by recognizing that each side of the die had an equal chance but since there were more gold sides, gold was more likely to occur. However, 23% of the students indicated that black was equally likely as gold to appear, thus using the equiprobable bias between the two choices of color, the outcome is strictly chance.

Williams (1995) found that some 11-12 year-old children view probable and “50-50,” as an inherent characteristic of chance occurrences. When asked about the chance of rolling a four on a 10-sided die (with numbers 1-10), one student responded: “50-50. Even chance of getting a ‘6’ or a ‘4’ or a ‘3’.” In this student’s choice of language, he or she seems to use “50-50” to represent the notion of randomness – not that “4” has a 50% chance of occurring. The student clarifies this by stating that several events (and possibly implying that all events) are equally likely. The overall results of Williams’ study revealed that about 30% of the students did not use the equiprobable bias, but that about 20% of the students answered at least 1/3 of the survey and interview questions using this bias.

It appears that an assumption that events are equally probable might be used incorrectly when a student has a naive understanding of chance and randomness as well as a lack of understanding or experience in being able to use a classical approach to probability to either mentally calculate a theoretical value or subjectively estimate probabilities based on a classical framework. However, based on the limited studies done with students using an equiprobable assumption, it is difficult to tell whether this assumption is actually prevalent in students’ problem solving. The research is also troublesome since it is possible that students misinterpret tasks or misuse probability language that researchers interpret as an apparent equiprobable bias when the students might, in fact, correctly understand the probability concepts being assessed in the task.

In his 1994 position paper, Bramald blamed many of the probability misunderstandings he found in his preservice secondary mathematics teachers, especially the overused assumption of equiprobable events, on the "urge to get them to work too quickly with estimates of probabilities which assume an underlying symmetry of outcomes" (p. 85). He suggested posing tasks that are difficult, if not impossible, to evaluate theoretically (e.g., tossing a tack and figuring out the probability of it landing in each of three positions pictured in Figure 2.2.).

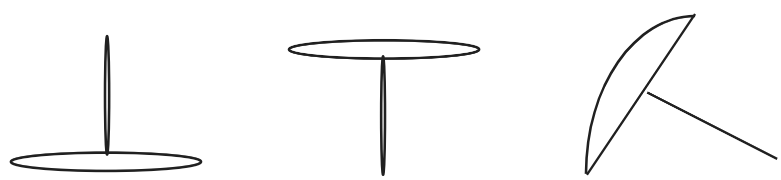


Figure 2.2. Three possible ways for a tack to land when tossed in the air.

According to Bramald, using a frequentist approach and encouraging more subjective probability judgments based on experimental results more closely resembles the real-world uses of probability (e.g., deciding to fly or drive on vacation based on the probability of having an accident in either mode of transportation). Bramald's paper is one of the few that has addressed the need for connecting probability teaching, learning, and research to the real world. The "real-world" argument often used for why we should teach probability somehow gets lost in implementing actual curriculum, instruction, and research on probability. Future work in probability teaching, learning, and research certainly needs to address this issue.

Outcome Approach

The heuristics of representativeness, availability, and equiprobable, can all cause misconceptions, some based on primary intuitions as well. However, researchers interpret the majority of these misconceptions based on the assumption that students consider relative frequencies of repeated trials. What about students who only look to predict the outcome of a single trial? The approach to solving a problem based on predicting a single outcome is certainly different than approaching the problem with a global view of the situation that considers relative frequencies and theoretical probability.

Outcome-based thinking often occurs when asking students to consider the following question: A sequence of six coin flips resulted in THTTTT. Which outcome is more likely to occur next:

A) Heads

B) Tails

C) Both A and B are equally likely

Instead of reasoning about the probability of an individual trial being heads or tails as 0.5, a student might reason that the next outcome should be heads (based on the representativeness heuristic) and because they are only concerned about correctly predicting the next outcome (based on the outcome approach). The student's response of "heads" could be interpreted by some researchers as a result of using the negative recency effect (gambler's fallacy) and that the student does not understand independence. But, if the student is interpreting the question as a request to accurately predict the next outcome, their interpretation of the task conflicts with the original intent of the researchers – making the students' response invalid.

During a study on probabilistic reasoning of college students, Konold (1987) posed tasks similar to the ones mentioned above from other studies. However, he noticed that some participants tended to interpret questions about the probability of an outcome as a request to predict the outcome of a single trial. The participants using this approach would answer probability questions as yes/no instead of giving a global index of the probability or they would evaluate predicted probabilities as right or wrong after one trial.

In one problem posed by Konold, the participants had to interpret the statement "70% chance of rain." Several students reasoned about the task based on the following continuum:

0%	50%	100%
"It will not rain"	"I don't know"	"It will rain"

Based on this continuum, several students qualified 70% as close enough to 100% and said that it should rain. If then told that it did not rain, several students decided that the forecaster's prediction was wrong. Instead of taking a global view of 70% chance of rain, they viewed the probability as an expected outcome that would be either right or wrong.

In their 1993 study, Konold, Polletsek, Well, Lohmeier, and Lipson reported many inconsistencies in students' probabilistic reasoning. However, they believe that there are patterns in these inconsistencies that are "neither a reflection of basic deficits in logical reasoning nor a result of simple carelessness" (p. 411). They observed students using many different intuitive-based misconceptions, problem solving heuristics, and approaches (i.e., representativeness, availability, equiprobability, and outcome-based) while solving the same problem and or similar problems. Thus, they concluded that "there is no simple story about how students reason about chance. Indeed, one of the major reasons that probability is notoriously difficult to teach is that students bring into the classroom not just one but a variety of beliefs and perspectives about chance" (p. 413).

Evolution of Misconceptions with Age

Throughout the literature on misconceptions, one obvious missing issue is how misconceptions change as students go through different stages of cognitive development and get more experience with probabilistic situations. In a preliminary study on the evolution of probabilistic misconceptions, Fischbein and Schnarch (1997) gathered empirical data to help them develop a theoretical framework from which to conduct future studies. They distributed a questionnaire to 20 students in each grade of 5, 7, 9, and 11 as well as 18 college-age prospective secondary mathematics teachers. The questionnaire consisted of seven problems, each related to a well-known probability misconception. The following is a brief discussion of the results for each of the misconceptions tested:

- The *representativeness* misconception decreased with age. The question used tested whether two strings of numbers were random enough to represent the parent population.
- The use of the *negative recency effect* (gambler's fallacy) decreased with age, and the *positive recency effect* was almost absent throughout all ages.

- The misconception between *compound and simple events* was frequently observed across all ages. (The question they used to test this misconception was Lecoutre's (1992) two dice question on the probability of rolling the pair five and six or rolling two sixes.) This was the only stable misconception found. I believe that the wording of this type of question is difficult and that most students probably interpreted the pairs 5-6 and 6-6 with order making a difference.
- The use of the *conjunction fallacy* was very strong through grade 9 and only observed in about half the participants in grade 11 and college.
- The misconception on the *effect of sample size* (related to the "law" of small numbers) increased with age.
- The use of the *availability* heuristic leading to misconceptions increased with age.
- The misconception of the Falk phenomenon (Falk, 1979) (where predictability is seen as one-directional and students do not realize the applicability of Bayes theorem in determining the probability of a first event given that a second event is known) also increased with age (except for the college students).

Based on these findings, the researchers plan on further studying how and why these misconceptions evolve throughout the school years. However, their use of a questionnaire with only one question to test each misconception is extremely limiting. The possibility of students misinterpreting tasks is very plausible and cannot be determined without a closer analysis of their thinking. Although they plan on using this study to develop a theoretical framework for future studies, I recommend they use more in-depth interviews, observational methods, and a variety of tasks to test a single misconception.

Connections Between Probability and Rational Numbers

Throughout the literature on the teaching and learning of probability, there are continual references to the use of rational numbers. I believe there are several common threads in the teaching and learning of both probability and rational numbers that are important for developing students' conceptions in both domains. Therefore, I would like to briefly discuss the importance of: 1) children's concept of "fair"; 2) intuitive and informal understandings of concepts; 3) the concept of equivalence; and 4) appropriate use of multiplicative reasoning.

Concept of "Fair"

The concept of "fair" plays an important role in children's understanding and performance with partitioning tasks (Pothier and Sawada, 1983) as well as in chance situations. Research on both probability and rational numbers has shown that teachers often assume children share the same meaning of "fair" as most adults have. Actually, students' concept of fair in chance situations has not been widely studied. I was only able to find one study that documented students' conception of fair (Lidster, Pereira-Mendoza, Watson, and Collis, 1995). In this study, sixth grade students were involved with several activities involving a die. Some students thought a die might not be fair because different amounts of dots (representing 1-6) could make some sides weigh more than others. Other students thought that a die's fairness could be controlled, depending on how it was rolled (e.g., starting with a "4" facing up made a "4" more likely to be rolled). When given a situation where a die actually was unfair, some students, even with evidence to the contrary, could not let go of their belief that all dice have to be fair.

In the study of both rational numbers and probability, students' ability to perform mathematical tasks involving a "fair share" or a "fair game" relies heavily on their understanding of the concept of "fair." Students enter school with a variety of life experiences where fairness has different meanings. In the real world, children, and especially parents, often share items fairly among several children with "leftovers" remaining. In the case of partitioning tasks, children need to be explicitly told that items need to be shared fairly so that there are no left-overs. Children also use the word "fair" in the context of losing a game or being punished. They often say phrases such as "that's not fair" only because circumstances are not in their favor. A young child may think a die is not fair if they do not get the result they are hoping for. Thus, in order for students and teachers to communicate about "fair" situations, teachers need to take the time to listen to students' use of the word and help them develop a shared meaning as a classroom community.

Intuitive and Informal Understandings

Researchers in both domains are interested in children's intuitive and informal understandings of concepts. Fishbein (1975) developed the idea of primary and secondary intuitions of probability. Other researchers (Hawkins & Kapadia, 1984; Schlottman & Anderson, 1994; Williams, 1995) found that children's primary intuitions can be developed into more normative ones with instruction. In the research on rational numbers, studies (Mack, 1990, 1995; Thompson, 1994) have shown that children also have intuitive and informal understandings of rational number concepts. The results from a study by Hawkins and Kapadia (1984) connect the learning of rational numbers and probability. They found that students that do not possess the formal knowledge of operations with rational numbers to work *a priori* with probabilities can make appropriate intuitive and subjective judgments on probabilities using their number sense and counting skills. For example, given a bag of candies with four yellow, two blue, and one green, children may use a part-part comparison and intuitively recognize that the yellow candies are most likely to be drawn from the bag without formally comparing the probabilities in part-whole form.

The literature in both domains recommends tapping into students' intuitions when introducing and developing concepts. I also think that instruction should include deliberate attempts to connect intuitive notions of rational numbers and probability. Tasks such as the following could provide connections between informal understandings of the size of unit fractions and comparing the probability of two events:

> Both you and a friend have the same size chocolate candy bar. One peanut is inside each candy bar. You cut your candy bar into four equal pieces. Your friend cuts his bar into six equal pieces. You each randomly choose one piece of the candy bar to eat. Which one of you is more likely to eat the piece with the peanut, or do you both have the same chance? Explain your reasoning.

Children's reasoning on such a task might help them make informal connections between concepts and provide a teacher with insight into their development and application of concepts in situational contexts.

Concept of Equivalence

In many problems involving probability and rational numbers, students use the concept of equivalence. In most school-learned probabilistic situations, objects such as coins and dice are used and an assumption is made that the possible events are equiprobable. Usually this equivalence is established geometrically (e.g., a die is a regular cube) and is important in determining the relationship between the possible outcomes. Recall Madsen's (1995) study in which 13-19 year old students were given a six-sided die with one face painted black and the other five faces painted gold. In this study, 47% of the students correctly recognizing that each side of the die had an equal chance but since there were more gold sides, gold was more likely to occur. However, 23% of the students indicated that black was equally likely as gold to appear. These students may not have had a well-developed concept of geometric equivalence and may have been relying on an informal notion that between two choices, either one are equally likely to occur. This informal knowledge may be linked to their understanding of partitioning an object into two "fair shares." Students not able to recognize the die being split into 1/6 black and 5/6 gold may be at the first or second level of the partitioning model proposed by Pothier and Sawada (1983). At these first two levels of partitioning, equality of parts is not an issue and, thus, students may not be concerned with the two color choices being unevenly distributed on the die.

Students' ability to understand and recognize equivalence, either informally (with words or pictures) or formally (with symbols), is going to affect their ability to reason appropriately about probabilities and rational numbers. I believe that activities in both domains are needed to build a solid understanding of equivalence that can be applied to both probabilistic and rational number situations. Too often, traditional curriculum and instructional methods focus on equivalence of rational numbers in symbolic form and assume that when students begin to explore probabilistic situations they will recognize equivalent probabilities. Perhaps if students could explore the outcomes when tossing a coin five times and record all ordered outcomes, then they might be able to recognize that the equiprobability of heads and tails on a single toss can be generalized to the equiprobabilty of any two specific ordered results when tossing a coin five times. Such a robust understanding of equivalence could help students avoid employing the representativeness heuristic inappropriately.

Multiplicative Reasoning

At the upper elementary and middle school level, the concept of equivalence is important with generating equivalent fractions (e.g., 1/2 = 3/6) or ratios (e.g., 3:2 and 9:6), determining whether two fractions or ratios are equivalent, and determining whether two events are equally likely to occur. All these situations involving equivalence require multiplicative reasoning skills with proportions. For example, if bag A has one blue and two green marbles, and bag B has three blue and six green marbles, a child needs to use multiplicative reasoning to determine *a priori* that the probability of picking a blue marble for bag A and bag B is equivalent. However, I think that such situations could help students develop a sense of invariance of ratio (Harel *et al*. 1994) if they are given ample experimentation time to sample from each bag and record the frequency of blue marbles. Such a frequentist approach might lead students to conjecture that the probability of picking a blue marble from either bag is equally likely. If students do this type of sampling activity with several bags whose marbles are in proportion, they might be able to look for patterns and begin to recognize the equivalent ratio between the two colors of marbles in each bag.

At times students can misuse proportional reasoning and believe that when flipping coins simultaneously the result of two heads and one tail is equally as likely as the result of 200 heads and 100 tails. For some students, the equivalent ratios are a sufficient reason for determining equiprobability without considering that with flipping a smaller number of coins, one is more likely to get the ratio of two heads and one tail. Thus, I think it is reasonable to conjecture that misconceptions students have about "the law of small numbers" could be related to their experiences with equivalent ratios in symbolic form without contextual meaning. Teachers should explicitly connect problems like this coin problem with the study of equivalent ratios and fractions so students may learn to analyze the context of a problem that appears proportional before applying proportional reasoning.

Instructional Integration

The study of probability inherently uses rational number concepts. However, the use of rational numbers in the study of probability is often unconnected with conceptual meaning (e.g., students are often taught the multiplication rule for determining the probability of compound events without reference to why multiplication makes sense to use) and, thus, is treated as if students already have the conceptual understandings – which, based on research and experience, is usually not the case. I conjecture that curriculum and instructional methods that intertwine and connect concepts in both probability and rational numbers have the potential to provide students with better understandings of concepts in both domains.

Teaching Probability Concepts

Many researchers (Borovcnik, Bentz, & Kapadia, 1991; Fischbein, 1975; Hawkins & Kapadia, 1984; Shaughnessy, 1992) agree that conceptual development of the nature of probability must not begin with teaching mathematical axioms (from a structural view), but should rely on intuitions (from a subjectivist view) and objective analysis of repeated experimentation (from an objectivist view). The traditional approach to teaching probability typically uses the classical *a priori* model in a few *ad hoc* activities within the high school curriculum and a structural approach in advanced college level courses.

Knowledge of *how* and *why* students reason under uncertainty should guide *how* probability is taught so as to properly develop the fundamental concepts involved in probabilistic reasoning. There are various opinions (Borovcnik & Bentz, 1991; Streinbring, 1991; and Shaughnessy, 1992) on how probability is best taught. These instructional models all sharply contrast with the traditional rule-driven theoretical approach which, along with other skill-oriented mathematics instruction, is not effective in teaching students underlying concepts and how to apply probabilistic reasoning to unfamiliar situations.

Several researchers have used an experimental approach to teaching probability with K-16 students. These instructional interventions were explicitly designed to develop students' primary intuitions into more normative secondary ones. In effect, these researchers have attempted to combat students' existing "misconceptions" in the areas discussed above and help them develop appropriate probabilistic reasoning. Although all the teaching experiments have a common thread of an experimental, exploratory instructional approach, only a few studies utilized technological tools, while most did not. The following descriptions are brief highlights of some of the research and the respective results.

Teaching Probability Without Technology

Several studies have been conducted to assess the effects of instructional intervention without technology on students' probability conceptions. Green (1983), Konold (1987, 1991), Shaughnessy (1977, 1992, 1993), and Castro (1998) are among several mathematics educators who advocate confronting secondary and college students' misconceptions with experimental results and then justifying those results with theoretical calculations of probabilities. Hawkins and Kapadia (1984) suggest using a subjective approach, in addition to classical and frequentist approaches, to teaching and learning probability would build a solid framework for young children to develop their probabilistic reasoning.

In his instructional intervention experiment, Shaughnessy used the following instructional model for teaching a college-level course in probability and statistics. For each activity in his teaching experiment students had to:

1. claim a stake in the task by making a guess of the outcome;
2. carry out the experiment with physical devices, gather and organize data;

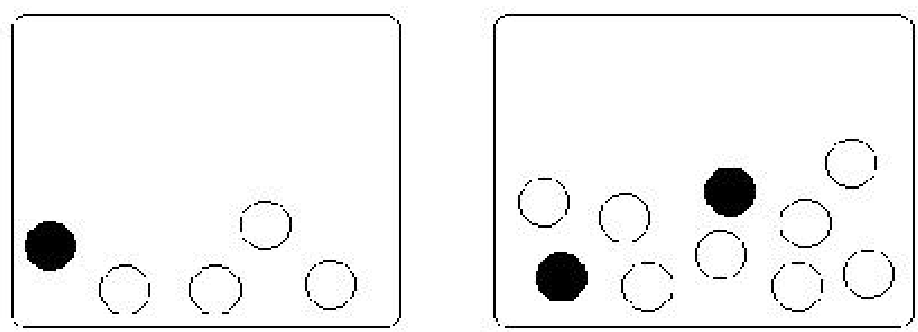
3. compare experimental results with initial guesses;
4. explicitly confront misconceptions with experimental evidence; and
5. build a theoretical probability model to explain the outcomes of the experiment.

Throughout the entire course, students had to "reconcile the dissonance between their stochastic [probability and statistics] misconceptions and their empirical observations" (p. 482). Shaughnessy assessed the students, in both this course and a traditionally taught course, pre- and post-instruction. Probability misconceptions were reduced much greater after instruction in the course mentioned above than in the traditional course. However, some students did not change their beliefs or misconceptions as a result of the course. In reflection on the results, Shaughnessy (1992) stated "It is very difficult to replace a misconception with a normative conception, a primary intuition with a secondary intuition, or a judgmental heuristic with a mathematical model" (p.481).

In Castro's 1998 teaching experiment with Spanish 14-15 year-olds, he purposefully designed an instructional model to initiate conceptual change in students' probabilistic reasoning. In comparison to the control group, which received a traditional curriculum through traditional teaching methods, the experimental group performed significantly higher on both a probability reasoning and probability calculations post-test.

Many of the studies focused on students above the age of 11. However, several researchers (Fischbein, 1975; Hawkins & Kapadia, 1984; Schlottman & Anderson, 1994; Jones, Langrall, Thornton, & Mogill, 1997, 1999a; Jones, Thornton, Langrall, & Tarr, 1999b) have studied the probabilistic understanding and reasoning skills of young children. These studies revealed that young students do possess many primary intuitions and that instruction can develop those intuitions into more normative ones.

For example, Fischbein found that after a brief instructional intervention 9-10 year-olds were able to correctly compare two probability situations and evaluate the respective chances for tasks similar to those in Figure 2.3. These children were also able to reason appropriately in probabilistic situations and in general operate correctly with the concept of probability. In addition, Hawkins and Kapadia found that even students that do not possess the facility with rational numbers to work *a priori* with probabilities could make appropriate intuitive and subjective judgments on probabilities using their number sense and counting skills.



4 White and 1 Black — 8 White and 2 Black

You want to choose a black ball. Which urn gives you the best chance of choosing a black ball, or are the chances for each urn equal?

Figure 2.3. Urn task to test children's ability to compare probabilities.

Most recently, Jones *et al*. (1997; 1999a, 1999b) have conducted many task-based interviews and delivered instructional intervention programs with elementary children. In their work, the authors have developed, validated, and refined a framework for assessing children's probabilistic thinking. In this framework, they have developed descriptors that reflect the level of children's understanding across several constructs. According to the framework, children's understanding progresses through four levels: 1) subjective thinking; 2) transitional between subjective and naive quantitative thinking; 3) informal quantitative thinking; and 4) numerical reasoning. In 1997, the framework included the constructs of sample space, probability of an event (emphasis on theoretical), probability comparisons, and conditional probability. In 1999(a), they further expanded the framework to include independence, and separated probability of an event into two separate constructs, one focusing on experimental and the other on theoretical.

Jones *et al*. (1997) found that the levels of third grade students' thinking was relatively consistent across all constructs (e.g., a student would demonstrate understandings at levels one and two for all four constructs) before instructional intervention. However, a child's levels of understanding were much more varied after the instructional intervention. It is interesting to note that the construct of sample space greatly improved after instruction while many children did not demonstrate levels of understanding above level two on the constructs of probability comparisons and conditional probability either before or after instruction. Very few of the third grade students performed at a level 4 understanding (numerical reasoning) either before or after instruction. Jones *et al*. (1997) believe that students' "lack of knowledge of fractions was a significant impediment" to developing a level 4 understanding in the four constructs (p. 122).

Overall, the many teaching experiments conducted to improve students' probabilistic understandings have employed an experimental, exploratory teaching approach and have been successful in improving students' understanding. All of the instructional interventions done without technology utilized physical devices such as coins, dice, and spinners to allow students to gather their own experimental data. However, many of the analyses done with the data only used tallying and numerical comparisons, and did not typically involve students constructing graphical representations of the data. With the widespread availability of computers in schools, it is important to study how computer environments that allow quick simulations of random events and graphing of data can affect the teaching and learning of probability.

Teaching Probability With Technology

A few software applications (e.g., *Tabletop Jr, Graphers, Probability Toolkit, and Probability Constructor*) are commercially available for teachers to use in the study of probability. However, as of yet, only a limited number of research studies have been done to study the effectiveness of using these types of tools. There are no studies known to this researcher that have actually studied the effectiveness of the above mentioned commercial software on students' probabilistic reasoning. The studies described below have all been done with software applications developed in university settings.

As part of the development of the ChancePlus curriculum mentioned in Chapter 1, Konold (1991) developed and used the *ProbSim* software with his college-level students and found that it was useful for students to see the results of the simulations and have the chance to analyze data. Most recently it was used with high school students at SummerMath at Mt. Holyoke with very positive results (as discussed in Chapter 1).

Konold uses the following general instructional guidelines for confronting probability misconceptions:

1. students must test whether their beliefs coincide with those of others;
2. students must test whether their beliefs are consistent with their own beliefs about related things; and
3. students must test their beliefs against empirical evidence.

Unfortunately, *ProbSim* is only commercially available to mathematics teachers in Australia and has been extensively used in a professional development program for enhancing teachers' understanding of probability and statistics (Watson & Moritz, 1997). Teachers in the US can purchase the software from Konold and his colleagues at the University of Massachussettes, but without commercial advertisement, I am sure the use of this software in US high schools and universities is limited.

Jiang (1994) developed a computer environment called *Chance World* (developed at the University of Georgia) and used it with middle and secondary grade students in a tutorial setting for his dissertation. He found that the small group sessions and the use of the computer environment were helpful for students apparently overcoming classical misconceptions of probability. However, the short time period and tutorial-like structure of the instruction can not predict the effectiveness of using the environment with a large class and for an extended period of time. This software only runs on older Macintosh computers, and even Jiang himself does not currently use it in his university courses (personal communication).

More recently, Vahey (1997) studied the use of another probability software application, *Probability Inquiry Environment* (PIE), in inquiry-based instruction with seventh graders. He found that the use of the computer and the inquiry nature of the activities brought out students' intuitions that were not easily characterized by the common misconceptions discussed in the literature. The 1997 study was not designed to test the effectiveness of the software in combating students' intuitions. However, the results of the study indicate that computer simulations can provide a rich domain for investigating students' understanding of probability tasks and the development of more normative probability conceptions.

In Vahey's dissertation (1998), he purposefully studied the effectiveness of the software in improving student's probabilistic conceptions. He used a quasi-experimental design with two seventh grade classes using the three-week PIE curriculum and two seventh grade classes using a three-week curriculum previously designed by the regular classroom teacher. All classes were taught by the regular classroom teacher and both curriculums employed an experimental approach to instruction. The students in both groups were given the same pre and post-test. The scores on the pre-test were not significantly different. However, the students using the PIE curriculum significantly outperformed the students in the comparison condition on the post-test. In addition, Vahey reported that the students in the PIE group became more attuned to the importance of the outcome space in probabilistic situations, and were also more attuned to interpreting randomly generated data.

Although the software used in these studies was different, they had several features in common. *ProbSim*, *Chance World*, and *PIE* were developed based on research of known probabilistic misconceptions and non-technological teaching experiments that successfully enhanced students' probabilistic reasoning. They all provided opportunities for students to experimentally generate random data and analyze it with linked multiple representations (e.g., numerical tallies, relative frequencies, lists of all trials, bar graphs). It appears that each of the software applications was successfully used to enhance probability instruction and students' learning of probability concepts.

Statement of the Problem

In reflection on his own research and the minimal amount of instruction in probability taught in many schools, Fischbein (1975) hypothesized that modifications to an individual's intuitive framework are difficult, if not impossible, "once the basic cognitive schemas of intelligence have stabilised (after 16-17 years of age)" (p. 12). If this is true, then education must find a way to transform primary intuitions into sound secondary ones during the cognitive formation years (i.e., before secondary school and college).

Shaughnessy (1992) emphasized that "probability concepts can and should be introduced into the school at a fairly early age" (p. 481). In addition, mathematics educators currently recommend that the study of chance happenings begins in the elementary school and advocate teaching probability based on intuitions and experimentation. As mentioned earlier, the NCTM and most state curriculum guides currently include probability and statistics concepts in the K-12 curriculum.

Several recent mathematics preservice methods textbooks (e.g., Schwartz & Riedesel, 1994; Riedesel, Schwartz & Clements, 1996; Van de Walle, 1997) recommend relying on students' intuitive notions as a starting point for explorations in probability.

This indicates that an effort is being made to educate future teachers on how to teach probability concepts. As Schwartz and Riedesel (1994) stated:

> A wise teacher will elicit predictions from children about how they expect events to occur. The children's predictions will be based on their intuitive notions. When events follow [from experimentation] that do not conform to the predictions, the teacher can use this opportunity to challenge the intuitive notions that led to the prediction. When a child comes face-to-face with the conflict between his ideas and observed events, there is a greater likelihood that he will be willing to modify his ideas to fit the observed events. (p. 188)

The research on students' understanding of probability has certainly provided mathematics educators with a wealth of knowledge on how students develop their probabilistic reasoning skills. An important aspect of teaching is being able to assess a student's understanding and to formulate an appropriate instructional intervention to cause cognitive dissonance so a child may assimilate and accommodate new information into their current schema. Cognitive dissonance with probability concepts occurs when a student is faced with experimental results that conflict with their intuitive notions. Assimilation and accommodation are on-going processes where individuals must make sense of experimental and theoretical probabilities.

Based on the research of students' intuitions and misconceptions of probability concepts, I have developed a software application called *Probability Explorer*. This software, along with appropriate instructional tasks and a setting designed to promote social construction of knowledge (Cobb, 1993), is designed to facilitate children's development of secondary intuitions in probability. The elements of design in the software will be discussed in Chapter 3.

There has been a considerable amount of research on children's understanding of probability concepts and several studies researching the effects of technology tools on students learning of probability in middle school and above. However, there are no existing studies that explore how technological tools can benefit elementary students in the learning of probability. I want to develop a better understanding of children's conceptual understanding and development in probabilistic reasoning in a technology-rich environment. Since the software will undergo further development over the next few years, there is also a need for a formative evaluation of the design elements in the software and its effectiveness in promoting children's construction of more normative probability conceptions.

Research Questions

The proposed research study has several questions that will guide the investigation of children's understanding of probability concepts in a technological environment.

1. What are children's understandings of probabilistic concepts (e.g., fairness, equivalence, sample space, experimental probability, theoretical probability, probability comparisons, and independence) and how do they develop appropriate probabilistic reasoning?
2. How are children's conceptions affected by their use of *Probability Explorer* as a problem-solving tool? What are the benefits and drawbacks of the instructional design and utility of tools in *Probability Explorer* for facilitating appropriate probabilistic reasoning in children?

Chapter 3 is devoted to the design of the computer microworld. I include a brief review of existing software and highlight major benefits and drawbacks of the applications that influenced the design of the *Probability Explorer*. A brief discussion of the principles of design for mathematical microworlds follows. Next, I describe how and why *Probability Explorer* was designed to help children develop their probabilistic reasoning. The specifics in the design of the proposed research study and analysis of data are discussed in Chapter 4.

CHAPTER 3

DESIGN OF *PROBABILITY EXPLORER*

The design of the *Probability Explorer* computer environment was born out of consideration of several factors: 1) the lack of research-based probability software available for elementary students; 2) a constructivist theory of designing computer microworlds; 3) research on children's probabilistic reasoning; and 4) my personal experiences with teaching probability and using computer software with children. This section will include a brief overview of the benefits and drawbacks of existing probability software and a description of a constructivist theory of designing computer microworlds. I will then describe several important features of the *Probability Explorer* in relationship to the research on children's probabilistic reasoning and my personal experiences.

Brief Review of Existing Software

Throughout the 1990s, several software applications have been developed for simulating random events. Only two software packages are available commercially – one suitable for grades K-8 (*Probability Toolkit*) and the other for grades 6-12 (*Probability Constructor*). The other three software applications previously mentioned in the review of the literature (*ProbSim*, *Chance World*, and *PIE*) were all developed in university settings and are not commercially available.

All of these computer tools facilitate an experimental approach to the teaching and learning of probability. They all have some sort of graphing capability to analyze results, keep track of frequencies and relative frequencies of experimental results, and simulate random experiments based on common physical devices such as coins, dice, spinners, and urns and balls. However, each of these software applications has benefits and drawbacks that I have learned from in creating my own *Probability Explorer*.

For example, none of the software generate manipulable representations of the results of an experiment. A few of the applications do create a list of all the outcomes; however, these outcomes are not moveable. Just as children benefit from tactile actions with concrete objects (e.g., rolling dice, flipping coins, or sorting blocks by various attributes), I believe the tactile actions of moving an iconic object in a computer environment can help instantiate the results of a random event. In addition, moveable icons can also help children develop sorting and organizational skills, as well as create their own pictographs.

Graphical representations of experimental results can be a powerful analysis tool. Only the software created for middle school and above provide dynamic links to graphs while a simulation is running. I believe that elementary students can benefit from such linked multiple representations. The ability to change parameters of an experiment (e.g., changing the likelihood of an event occurring) is a powerful benefit of computer simulation that can not been done without technology. However, only two of the computer environments (*ProbSim* and *Probability Constructor*) give the user control over parameters.

Only *ProbSim* allows teachers and students complete flexibility in designing their own experiment. However, the abstract nature of this software tool makes its usability limited to middle school and above. The *Probability Constructor* does contain many usable features; however, the language and symbolic notations used are clearly only appropriate for middle school and above. Thus, there is an obvious gap in the development of probability software that is research-based and usable by elementary students.

Principles of Design of Mathematical Microworlds

Students learn meaningful mathematics by making reflective abstractions as they accommodate their current cognitive structures to deal with a realization that something does not work or is unexpected (Cobb, 1994; Steffe, 1988; von Glasersfeld, 1995). Understanding mathematics, however, requires a conscious process of re-presenting experiences, actions, or mental processes and considering their results or how they are composed (von Glasersfeld, 1995). This process of abstraction and reflection contributes to a student's ability to construct meaningful mathematical knowledge. Based on this constructive process of developing knowledge, several researchers have worked with children and developed theories of how computer microworld environments can facilitate this process.

Papert (1980) originally used the term microworld to describe a self-contained world in which children "learn to transfer habits of exploration from their personal lives to the formal domain of scientific construction" (p. 177). It is important to note that a true computer microworld does not stand in isolation from social interactions from peers and teachers. Many researchers have commented that appropriate instructional tasks and social interactions among students and among students and teacher are vital components in successful uses of a microworld (Biddlecomb, 1994; Steffe & Wiegel, 1994; Battista, 1998; Olive, 1999). In fact, Battista believes that for a computer microworld to be "fertile" the environment should:

1) support problem-centered inquiry;

2) be based on research of students' mathematical learning;

3) cultivate mental models of abstract ideas; and

4) induce reflection and abstraction.

Biddlecomb emphasizes that "computer environments must be very flexible in order to make them as open as possible for the teacher and students to construct their own individual and shared mathematical environments" (p. 91). The open-ended nature of the computer environment is critical in fostering appropriate and sustained learning experiences. Land and Hannafin (1996) state that open-ended learning environments such as microworlds should "support experiences for learners to identify, question, and test the limits of their intuitive beliefs" (p. 38).

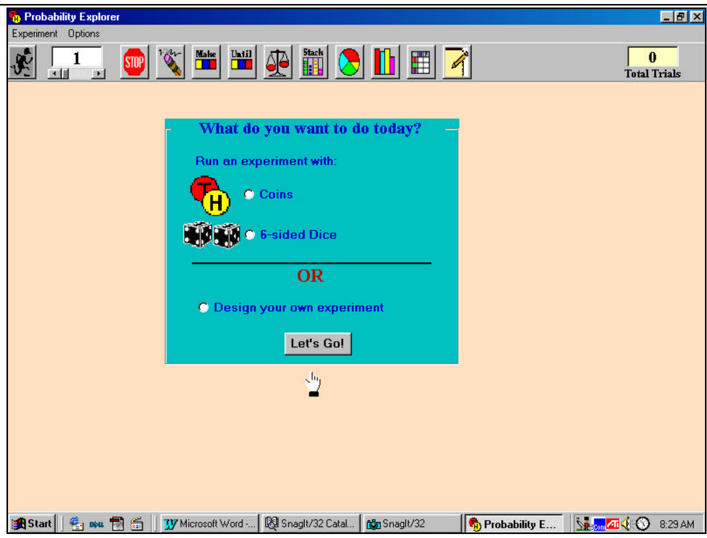
The principles of design for a computer microworld are in sharp contrast with the many tutorial and game-like software applications that dominate the commercial market in K-12 mathematics software, especially at the elementary level. The design and development of *Probability Explorer* stems from a constructivist theory of learning and design principles for a computer microworld. The environment does not "teach" students about probability. Instead, children can develop their intuitive notions about probability as they use the tools available in the microworld to develop appropriate probabilistic reasoning. The instructional tasks, and computer and social interactions will, of course, be pervasive elements in the children's process of constructing probability knowledge.

Description of Software Features

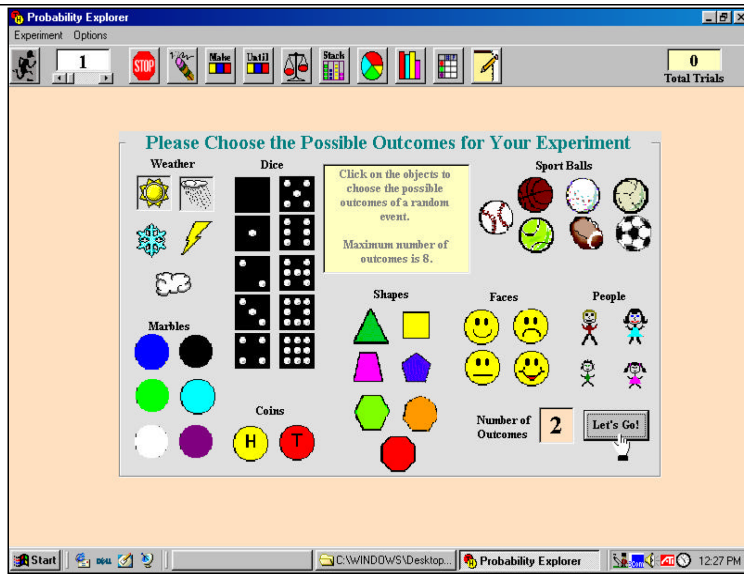
The overall goal in designing *Probability Explorer* was to create a relatively open-ended environment that could easily be used by children to simulate random phenomena and explore interesting chance situations. The chance situations could be in the context of a game (e.g., dice games) or real world uncertainties (e.g., weather). The tools and actions available in the computer environment have been purposefully designed to invoke perturbations in a child's current schema of probability concepts, encourage active reflection and abstraction to refine those conceptions, and facilitate the development of appropriate probabilistic reasoning. What follows is a description of several features purposefully designed in the computer environment and a justification of why those features are appropriate for helping children develop probabilistic reasoning. Since the design of the software has been an iterative process throughout the research study, the description in this chapter describes the features and tools available in the prototype version (as of May 1999) that was initially used in the pilot study. The enhancements and additions to the software made during the research process will be described in Chapter 4.

Designing Experiments

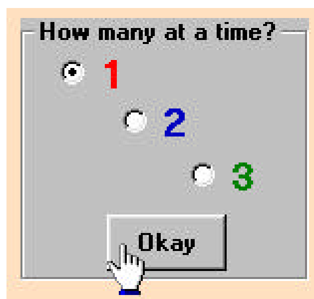
A salient feature of the computer environment is the required actions for defining the type of experiment to simulate. The prototype software contained two "preset" options for children to run simulations with flipping a coin or rolling a regular six-sided die (Figure 3.1a). However, even if a child chooses a "preset" simulation with a coin or die, he or she must decide how many coins or dice to "flip" or "roll" at a time (Figure 3.1c). The action of deciding how many events in the simulation provides students with a moment of reflection to think about how they are using the tools in the computer environment to model a chance situation. In addition, students can also design their own experiments by choosing from approximately 50 icons that will represent the possible outcomes of a single random event (Figure 3.1b). Once the possible outcomes are chosen, the child must also decide whether to simulate one, two, or three events at a time.



3.1a. Menu to choose a preset experiment or to design your own



3.1b. Options available for designing your own experiment



3.1c. Choosing a simple (1) or compound (2 or 3) experiment

Figure 3.1. Defining the type of random experiment.

With these options, children can create experiments that could be used to model many typical textbook probability situations, or they can imaginatively design experiments of interest to them (whether they are playfully contrived or meant to model real world phenomenon). In a description of a wish list for future development of probability software, Biehler stated that "it would be valuable to have more experiences with software where students can design random devices on the screen" (1991, p. 189). For children, I suggest that the action of clicking on iconic representations to choose possible outcomes will have a playful orientation and encourage students to want to explore various random situations. This is similar to Steffe's and Wiegel's theory that cognitive play can serve as a precursor to mathematical play (1994).

Connections With Physical World

Typical activities in probability involve the use of devices such as coins, dice, marbles, and spinners. Just as children benefit from concrete experiences with manipulatives for conceptual development, they also benefit from experiences with physical devices for generating random data. It is not my intention to replace physical experiences with digital simulations. In fact, I believe that without prior use of such physical devices, children will not fully comprehend the randomness of the computer simulation or make meaningful connections between the two-dimensional icons and their three-dimensional counterparts. In fact, Shaughnessy (1992) suggested that "it is important for us to continue developing connections between concrete simulations and computer simulations in our teaching and investigating the effects of the transition between the two in our research" (p. 485).

Transition between concrete and computer experiences. I have attempted to make the transition between concrete materials and computer simulations as seamless as possible. For example, the action of clicking on the "Run" button in order to simulate a random event represents a conscience action by the child to induce a chance event. This action is similar to the purposeful act of rolling a die or flipping a coin and anticipating the outcome. Once a random result appears on the screen, the child can act upon the object to move, sort, organize, or "stack" it in a playful and potentially meaningful manner. Actions on these objects can help the child instantiate the experimental results and build simple (piles) or complex (venn-like sorted groups) re-presentations of the data (see Figure 3.2). The ability to have moveable iconic representations of randomly generated data substantially extends the capabilities of experimentation with physical devices where results are usually only listed or tallied. In this regard, the computer microworld not only connects with the physical world, but extends the potential actions available in this new mathematical "world."

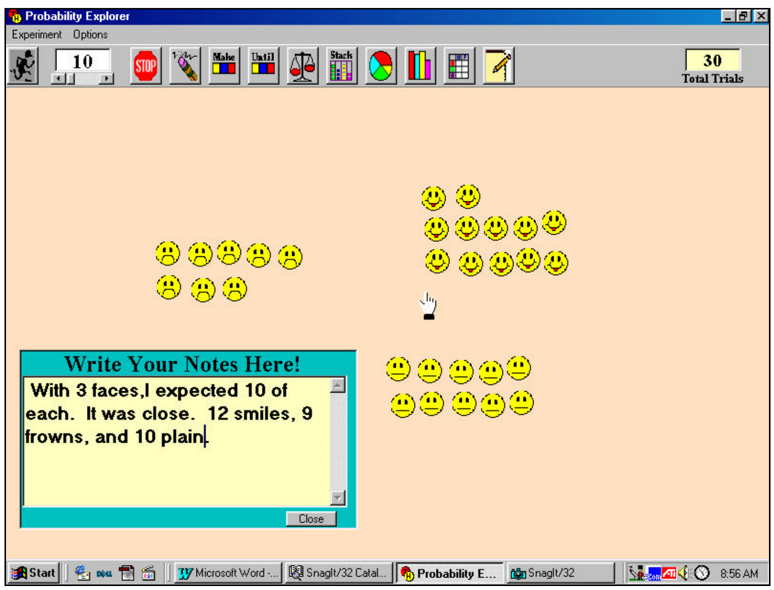
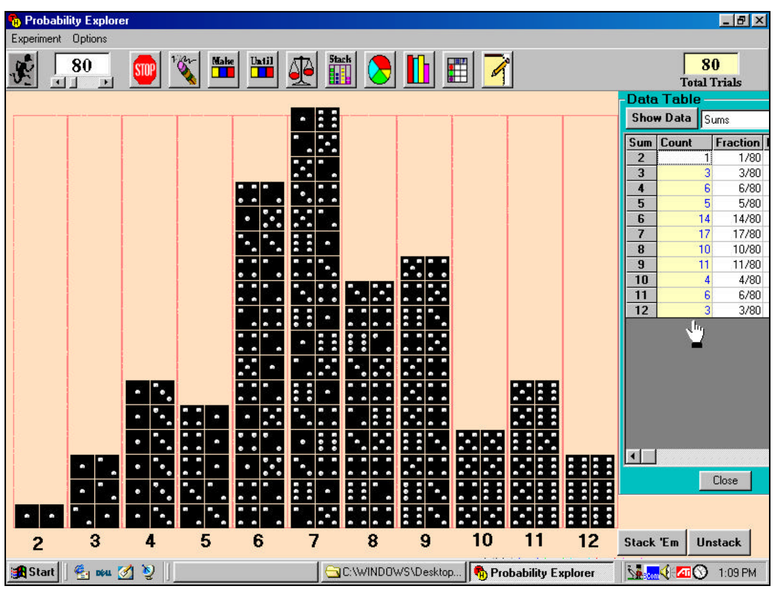


Figure 3.2. A re-presentation of data and comments in the Notebook.

Modeling typical probability tasks. Although the microworld can facilitate explorations of real world chance situation, the reality of available curriculum materials is consumed with activities involving games of chance using typical devices of coins, dice, and spinners. I do not want to take away the usefulness of games in the study of probability. Children, especially, tend to play games involving chance in their real world. Thus, studying probability in the context of games is a very viable means of enhancing their understanding. Therefore, the computer microworld contains tools and options to help children explore games of chance within the environment. For example, many commonly used games of chance involve rolling two six-sided dice. One such game might ask students to sum the results on each die and state that player A wins a point if the sum is a 5, 6, 7, or 8 and player B wins a point if the sum is 2, 3, 4, 9, 10, 11, or 12. By initially hypothesizing the “fairness” of this game and experimenting to gather evidence about their hypothesis, students will be involved in a genuinely interesting and engaging task (Vahey, 1997, 1998). Figure 3.3 shows the stacked results of sums of two dice and the table of results also displayed by possible sums. Although there are no specific games built into the environment, it was designed to be flexible enough for teachers and children to use as a modeling tool for a wide variety of game situations.



<u>Figure 3.3</u>. Results of the sum of two dice after 80 trials.

Making an Outcome

As noted in the review of the literature, many students have difficulty creating a sample space for a given experiment. I conjecture that the actions of creating their own experiment will give children a better understanding of sample space and enhance their awareness of the elements of a random situation. For simple one-stage experiments, students are actually creating the sample space as they choose the elements of the experiment. For two- or three-stage experiments, other features are available in the microworld to help them develop complete understandings of sample space. In addition, many of the instructional tasks will include contextualized problems that suggest students construct all possibilities for an experiment to make sense of the situation.

The “Make It” tool allows students to make specific outcomes based on the possible outcomes for each event. For example, the experiment in Figure 3.4 is designed with four possible choices for each stage (or event) in a two-stage (or two-event) experiment. When the “Make It” tool first appears, the two slightly indented boxes at the top of each column are blank. In order to make a specific outcome, the child must click on the icon of the desired outcome for the first event in column 1. When the icon is clicked, the image is copied into the first empty box at the top of the column. Similar actions are needed to choose the desired outcome for the second event. Once the desired two-event outcome appears in the top slot, if the child clicks on the “make it” button, the outcome is moved out on the “table.” This two-event outcome can then be moved around the entire table area for sorting and organization. Thus, a child can repeatedly create outcomes, place them on the table, and sort them in whatever way desired (see Figure 3.4). Although the vertical arrangement of the icons in the “Make Outcomes” box is meant to encourage an orderly generation of possible outcomes, children will not usually spontaneously use a systematic approach. However, I believe that the actions of sorting the outcomes on the table and the challenge of “making sure” that all possible outcomes are accounted for will eventually prompt students to look for more systematic strategies.

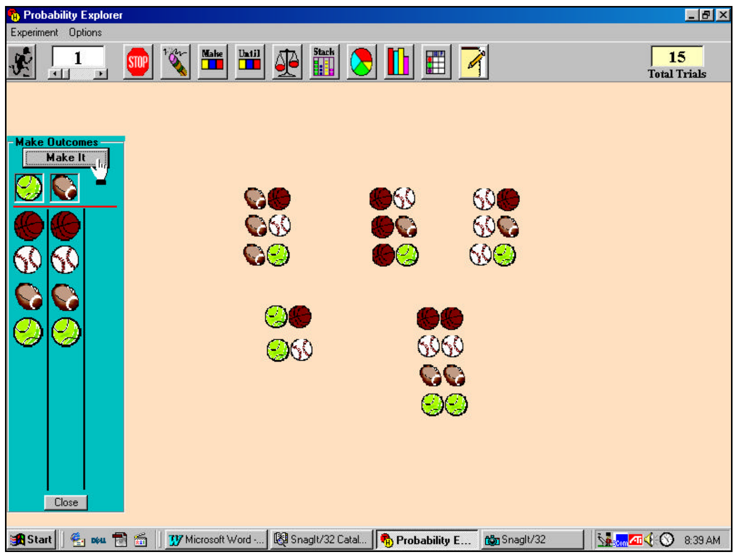


Figure 3.4. Using the “Make It” tool to construct all possibilities of an experiment.

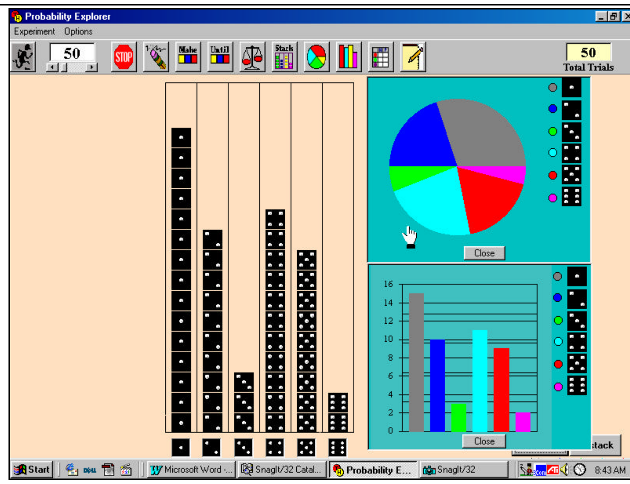
Dynamically Linked Representations

One of the most promising uses of computer technology in mathematics education is the ability to view multiple representations (i.e., verbal, numerical, graphical, algebraic, and geometric) of phenomenon. The most dominant representations used without computers are numerical and algebraic. However, the advent of powerful computing machines with advanced graphics capabilities made quick, accurate graphical representations possible. Biehler (1991) thought that, with respect to teaching and learning probability, “graphs can become exploratory tools for problem solving which is hardly possible without computer support, but this potential is too rarely exploited by current educational software” (p. 180). Many software applications for middle and secondary school students include the ability to see a mathematical situation in more than one representation. In addition, many of these representations are dynamically linked so that when the user changes an aspect of one representation, all the other representations simultaneously change. This connection between representations, and the use of a graph as a problem-solving tool, has been critically lacking in elementary mathematical software. Although several software packages exist that allow students to represent data in graphical form, the graph is often only used as a unit of display rather than a unit of analysis.

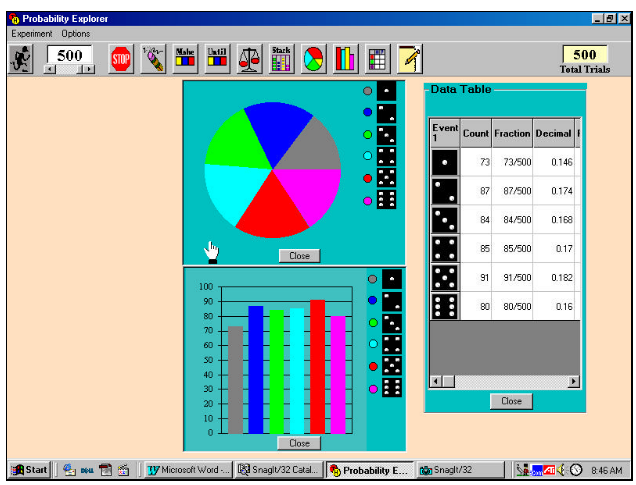
My intent was to create multiple linked representations that are updated simultaneously as random events are simulated. In this regard, the representations are not only a unit of display, but become a unit of analysis *during* experimentation. The representations available in *Probability Explorer* include:

1. Iconic representations of every trial which can moved around the screen;
2. “Stacking Columns” to create pictographs of results of small sample sizes;
3. A table which displays the experimental results as a frequency (count) or relative frequency (fraction, decimal, or percent);
4. A pie graph which displays the relative frequencies; and
5. A bar graph showing the frequency distribution.

Each of these representations is linked and changes dynamically while a simulation is occurring. For example, for a one-die experiment, Figure 3.5a depicts the results as a pictograph, pie graph, and bar graph. As the simulation is running, the child can watch different bars or columns “grow” and make observations of “fairness” or “likelihood” of certain outcomes occurring. The 50 trials of the one die experiment shown in Figure 3.5a might suggest that the die favors the number one and that three and six are not likely to appear. Such wide variability might surprise children and prompt them to run an experiment of 50 trials over and over again to test a hypothesis that three and six are less likely to occur. Upon noticing that three and six are not always the "losers," and that they are not able to predetermine which number will "win" a given set of 50 trials, students will reach a point of perturbation.



3.5a. Results from small sample (50)



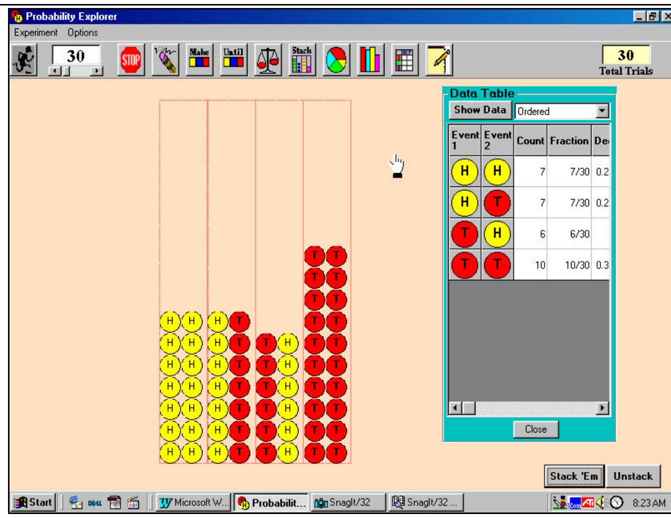
3.5b. Results from large sample (500)

Figure 3.5. Linked representations available for analysis of die toss.

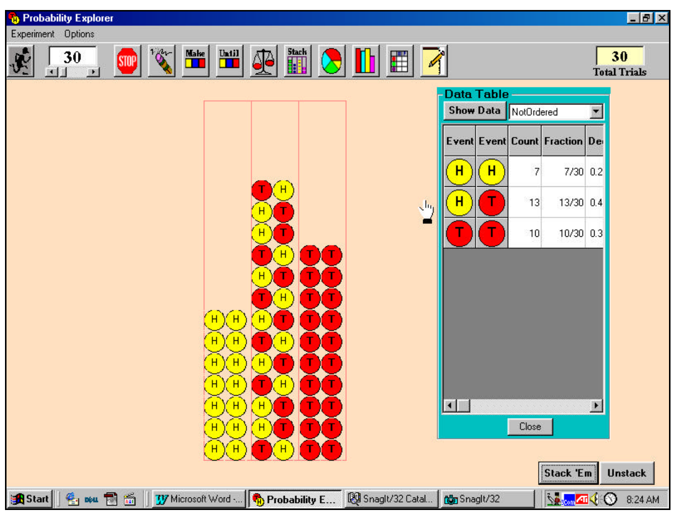
The microworld has the speed and ability to allow students to quickly generate a large amount of trials (there is an option to not show the icons on the table for such large simulations) to further explore rolling one die. Figure 3.5b shows the pie graph, bar graph, and data table for 500 trials of the same one-die experiment. Notice the similarities between the results for each possible outcome. Children can see first hand (and rather quickly) that the experimental results closely resemble what we would expect in theory from a "fair" die. The true power in this visualization of the law of large numbers comes from children watching the graphical representations *during* the simulation process and observing the wide variability in small samples and how the results (in this case) "even out" as the number of trials increases. This use of the graph as a unit of analysis can help develop appropriate uses of the representative heuristic and avoid such fallacies as the "law of small numbers."

Ordered and Unordered Events

Fischbein and Schnarch (1997) found that students across all ages had difficulty comparing the probability of rolling the pair five and six or double sixes with a standard die. Most students stated that the two events were equally likely to occur. In addition, when asked to flip two coins, I have had many children, and inservice teachers, tell me that there are three possible results – two heads, two tails, or one head and one tail. When asked about the likelihood of each event, two common responses have been "all three possibilities are equally likely" (employing an equiprobable heuristic) and "one head and one tail is more likely because we should expect half heads and half tails" (using a representative heuristic). Although it is more likely to obtain one head and one tail when flipping two coins, many people do not consider the two coins as differentiable. Thus, they do not consider the sample space of all ordered events (HH, HT, TH, HH) and determine that two of the ordered events make up the unordered event of one head and one tail (see Figure 3.6).



3.6a. Ordered outcomes in a two-coin toss.



3.6b. Unordered outcomes in a two-coin toss.

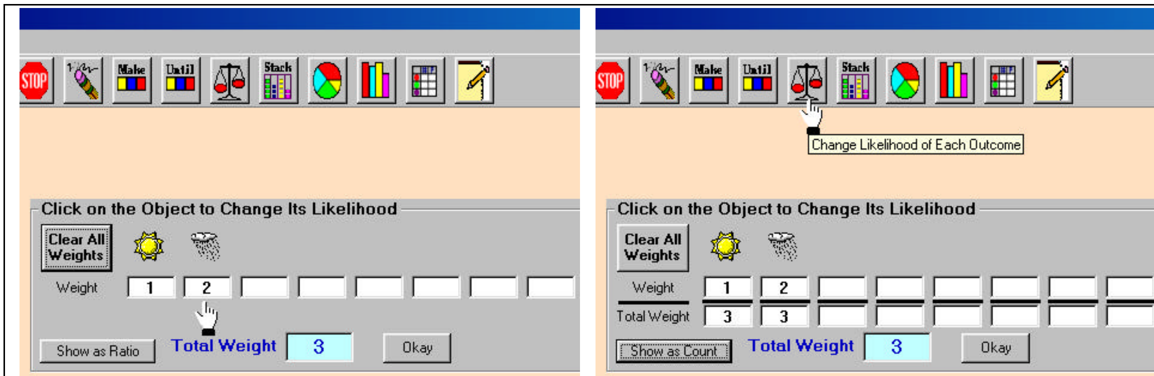
<u>Figure 3.6</u>. Organizing results of a two-coin experiment ordered and not ordered.

To help children avoid this type of fallacious reasoning, the *Probability Explorer* has the ability to “stack” two-event experiments ordered and unordered so children can compare the differences. When results from a two-event experiment are stacked with order mattering, the table of results is also displayed by order (see Figure 3.6a). Similarly, if the results are stacked “not ordered,” the table is also organized in this manner (see Figure 3.6b). I believe it is important for children to consider both ordered and unordered events in the context of real world situations. For example, if a basketball player has been fouled and is given two free throws, the order in which he or she makes or misses the basket does not effect the points earned. However, assuming a 50% shooting average, by considering all possible ordered results of the free throws, the probability of scoring only one point is twice as likely as either no points or two points. By considering these types of problem situations, I think children will be more attuned to the difference between ordered and unordered events, and when it is appropriate to consider either one or both.

Changing the Likelihood of an Outcome

Real world physical devices such as coins and dice cannot be easily and accurately altered to affect the probability of an outcome occurring. Thus many students who use such physical devices to model probability situations only experience theoretically equally probable outcomes. With only this type of experience with random phenomenon, students may inappropriately apply the equiprobable heuristic to real world situations that are not equiprobable (e.g., it is highly unlikely that two sports teams are truly equally likely to win a sporting event). In addition, I believe children’s conception of “fair” needs to be developed by purposefully experimenting with “unfair” situations. Hence, children need experiences with chance situations that are not equally likely to occur. They also can benefit from actually acting upon objects to determine the chances of a random event.

The computer microworld contains a “weighting” tool that can be used to alter the chances of an outcome occurring. I have chosen to use a metaphor of “weight” to help children understand the process of assigning probabilities to an outcome. “Heavier” outcomes are more likely to occur, while “lighter” outcomes are less likely to occur. Weight is measured in units of whole numbers. To facilitate the instantiation of the “weighting” process, children can click on an object in the “Weighting Box” to increase its weight. Each click corresponds to an increase of one in the weight. Figure 3.7 displays the “weighting” tool and the two options of viewing the chances. By default, students view the distribution of weights as a count (Figure 3.7a). This view of the chances will allow them to think about the part-to-part relationship between the outcomes. This is usually the initial way that young children think numerically about probability situations (Jones *et al*, 1997, 1999a). This level of thinking is also aligned with children’s early fractional thinking when they only consider the “parts” of a fraction (numerator) rather than the “part” in relationship to its “whole” (denominator). A part-to-part display is also similar to the concept of odds and can be useful for distinguishing between the odds and probability of an event. Because theoretical probabilities rely on both the “part” and the “whole,” children also have the ability to view the distribution of weights as a fraction (Figure 3.7b). Additionally, children can give all the “weight” to one outcome and explore “certain” situations as well as giving an outcome a weight of “0” to model an “impossible” situation. I conjecture that the ability to view the chances as both odds and probability will enhance students' ability to reason numerically about random situations.



3.7a. Part-to-part display of chances 3.7b. Part-to-whole display of chances

Figure 3.7. Tool to change the likelihood of an outcome.

The "weighting" tool gives teachers and children the power to explore many probabilistic situations that are difficult, if not impossible, to model. Thus, the study of probability can finally be connected with real world chance phenomenon and closely reflect a child's physical world. Finally, the real world argument of why students should study probability can be reflected in how probability is taught in schools.

Connections to rational number reasoning. Because the teaching and learning of probability inherently uses rational numbers, and because students notoriously have difficulty understanding and using rational numbers, it is important for a probability microworld to also have elements that promote appropriate rational number reasoning. The ability of the weighting tool to display chances as a part-part or part-whole relationship provides a contextualized way for children to use rational number reasoning. The weighting tool, data table, stacking columns and graphs can be used together to help students reason and problem solve with rational numbers while exploring probability tasks. For example, in a one-stage experiment with five possible outcomes, a student can use the weighting tool to make each of three outcomes have a probability of 2/8 and two outcomes with a probability of 1/8 each. The students can then hypothesize what they expect from experimentation, run a simulation, and explore the results using the stacking columns, data, table, and graphs. With the results in Figure 3.8, students' could be encouraged to notice the similarities in the number of squares and septagons and that the other three outcomes occurred about twice as much. This observation can be furthered explored with the count, fraction, decimal, and percent display in the data table. The depth of exploration and tools used in such activities would, of course, depend on students' familiarity with rational number representations. However, even the graphical displays and frequency counts could be used appropriately by students without sophisticated knowledge of rational numbers.

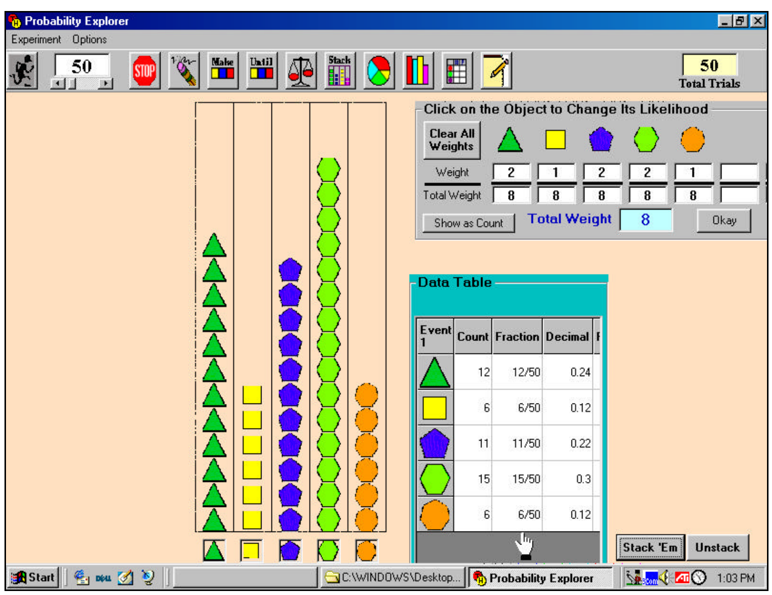


Figure 3.8. An experiment to develop rational number reasoning.

Instructional Design Issues

The development of the computer microworld environment was based on pedagogical implications of constructivist learning theory. With this in mind, *Probability Explorer* can be used in instruction to promote students' construction of probability concepts. However, the potential of the software is only as good as the instructional materials accompanying the software and the instructional beliefs and actions of the teacher. Materials for this microworld will be developed as part of the on-going research process. The teaching tasks used in the study will eventually be refined and further developed into an accompanying instructional guide that will be distributed with *Probability Explorer*.

CHAPTER 4

RESEARCH METHODOLOGY

The review of literature in the previous chapters outlines several issues pertinent to developing children's understanding of probabilistic concepts. This study builds upon previously conducted research and extends it to include the use of a research-based computer environment to help children develop probabilistic reasoning. In this chapter, I will describe the theoretical framework for the research study, the design of the study, as well as aspects such as the pilot study, participants, data collection, and analysis procedures.

Theoretical Framework

The design of the computer microworld is based on a constructivist theory of learning. The tools available in a computer environment, meaningful instructional and playful activities, students' schemes, and social and computer interactions all operate interactively as potential meaning-making agents for students' construction of concepts. The interactions between these agents will be part of a complex process of each child constructing their own knowledge about probabilistic situations.

Although there are mathematically accepted norms for defining and reasoning with probabilistic concepts (e.g., Kolomorgov's axioms, law of large numbers), children do not intuitively use normative probabilistic reasoning. The social and computer interactions in a "fertile" (Battista, 1998) computer environment should provide children with many opportunities to explore probabilistic situations and make sense of experimental results. These interactions can help students develop more taken-as-shared (Yackel, Cobb, & Wood, 1993) interpretations of probability concepts such as sample space and theoretical probability in a way that brings them closer to more normative probabilistic conceptions and allows them to communicate socially about their understandings.

To interpret the interactions which are part of the meaning-making process, this research draws upon an interpretivist approach to inquiry (Schwandt, 1994). For studying and interpreting children's meaning-making processes, Graue and Walsh (1998) state:

> The goal of interpretive research is to understand the meaning that children construct in their everyday *situated actions*. ... Individual action is generated out of social interactions and the meanings they create. It is enabled and constrained by the tools and resources (including other individuals) that compose the context. (p. 41)

Interpretive research with children's constructions of mathematical knowledge relies on a constructivist view of taken-as-shared meaning and is compatible with the symbolic interactionist view. "Symbolic interactionism sees meaning as social products, as creations that are formed in and through the defining activities of people as they interact" (Blumer, 1969, p. 5). Thus, in order to understand children's probabilistic reasoning while using a computer microworld, the researcher used interpretive research methods to observe and critically analyze the children's meaning-making processes and social and digital interactions.

Several researchers have extensively studied children's problem solving and social construction of mathematical meaning both with computer technology (Clements & Battista, 1994; Steffe & Wiegel, 1994; Bowers, 1995; Olive, 1999) and without (Wood, Cobb, Yackel & Dillon, 1993). The studies done with technology involved computer microworlds developed to help students understand fractional concepts (Steffe & Wiegel, 1994; Olive, 1999), place value (Bowers, 1995), and spatial and geometric concepts (Clements & Battista, 1994). This research study expands this line of research to include the use of a computer microworld while children solve tasks in the domain of probability.

Research Design

In order to answer the research questions set forth in Chapter 2, it was necessary to gather in-depth information about the children's understanding of probability concepts as well as how they interact with the *Probability Explorer* during problem solving activities. The information needed to assess their understanding prior to and following the activities using the *Probability Explorer* was gathered by probing the participants' thought processes during task-based interviews without the use of technology (Goldin, 1998). Recall that many earlier studies (e.g., Green, 1983; Fishbien & Schnarch, 1997; Zawojewski & Heckman, 1997) used questionnaires and test data to assess students' understanding. Task-based interviews allow for more in-depth questioning of students' reasoning and help alleviate misunderstandings of questions. In this case, the use of interviews without access to the technology were used to assess if students could transfer their experiences and probabilistic reasoning with the technology tools to typical tasks used in prior research and those that might be found in textbooks or paper-and-pencil assessments. Gathering information about their interactions with the *Probability Explorer* microworld was obtained by systematic observations while they were engaged in problem solving tasks. Hence, *qualitative* methods of inquiry were used to answer the research questions.

Patton (1990) specifies three methods for collecting qualitative data: in-depth, open-ended interviews; direct observation; and written documents. For this study, interview data (pre- and post-intervention interviews) was collected as participants completed a series of probability tasks. Additionally, direct observation of the participants during the instructional sessions by the researcher and a non-participant observer provided data regarding participants' interactions with each other and the computer environment. Audio and videotaping all interviews and instructional sessions allowed the researcher to critically review each session, obtain direct quotes, and observe participants' social interactions. In addition, the children's actions within the computer microworld were recorded using a PC-to-TV converter, microphone, audio/video enhancer, and VCR. This recording system resulted in videotapes of the children's actions on the computer (as seen on the monitor) as well as accompanying audio of all verbal interactions.

Participants and Setting

The participants in this study were children between the ages of 8-10, inclusive. In June of 1999, a description of the pilot and actual study was distributed to several 2nd and 3rd grade elementary classes within a 15-mile radius of the university with information on how to contact the researcher for possible participation. All participants and their parents signed informed consent forms (see Appendices A-D).

The instructional sessions occurred in the computer lab of the Lambeth House at the University of Virginia. This lab is equipped with five student workstations and one teacher workstation with projection capabilities and an interactive whiteboard. In addition, there is another whiteboard and large table in the room. Two workstations were each equipped with a PC-to-TV converter, microphone, audio/video enhancer, and VCR One camcorder was also used to capture whole group interactions as well as focused interactions at one workstation.

Pilot Study

A pilot study was conducted during July 1999 for several purposes: (1) to refine the researcher's skills as an interviewer, instructor, data collector, observer, and data analyzer; (2) to refine the interview and instructional tasks to be used in the actual study; (3) to develop initial hypotheses about children's use of the microworld tools in developing probabilistic reasoning; and 4) to test the software with children to help find nuances, "bugs," and other problematic characteristics in the microworld. The pilot study spanned three days and included individual 1-hour task-based interviews with children and three hours of teaching sessions using the *Probability Explorer* microworld. Three children, ages 8, 8, and 9, participated in the pilot study.

Pilot interviews. Each of the children participated in a videotaped task-based interview. The tasks used in the pilot interviews were based on tasks used in previous research to assess students probabilistic reasoning (Piaget & Inhelder, 1972; Fischbien, 1975; and Jones *et al*, 1997, 1999a). The goal of the selection process was to assemble a set of probability tasks that have been shown in earlier research to convey students' probabilistic misconceptions and levels of understanding. Transcripts of the pilot interviews were analyzed to determine which tasks were appropriate for the actual research study and to help the researcher modify wording and presentation of tasks.

Pilot teaching episodes. The children in the pilot study participated in four 45-minute teaching episodes using concrete materials and the *Probability Explorer*. Each teaching episode was video-taped using both a camcorder and the internal recording system described previously. Many of the teaching tasks used were modified from problems used in research as well as workbooks and textbooks for elementary mathematics. However, the *Probability Explorer* also allows students to explore probabilistic situations not found in elementary books (e.g., investigating the chance of having a family of four boys).

The data from the pilot teaching episodes allowed me to further refine and develop more appropriate teaching tasks for use in the actual study. I critically reviewed the teaching sessions and analyzed the students' interactions with each other and with the computer environment. This analysis helped me further understand the possible benefits or drawbacks in using the tasks and microworld in fostering appropriate probabilistic reasoning.

Further microworld development. After the pilot study, several computer bugs were fixed and the marble environment (Figure 4.1) was added to the microworld. Since many of the tasked used in textbooks involved picking items from a bag, I thought it was imperative to include that capability in the microworld. In addition, the marble bag tools would provide the children with a way to design and simulate a variety of experiments that could encourage a transition from part-part to part-whole reasoning.

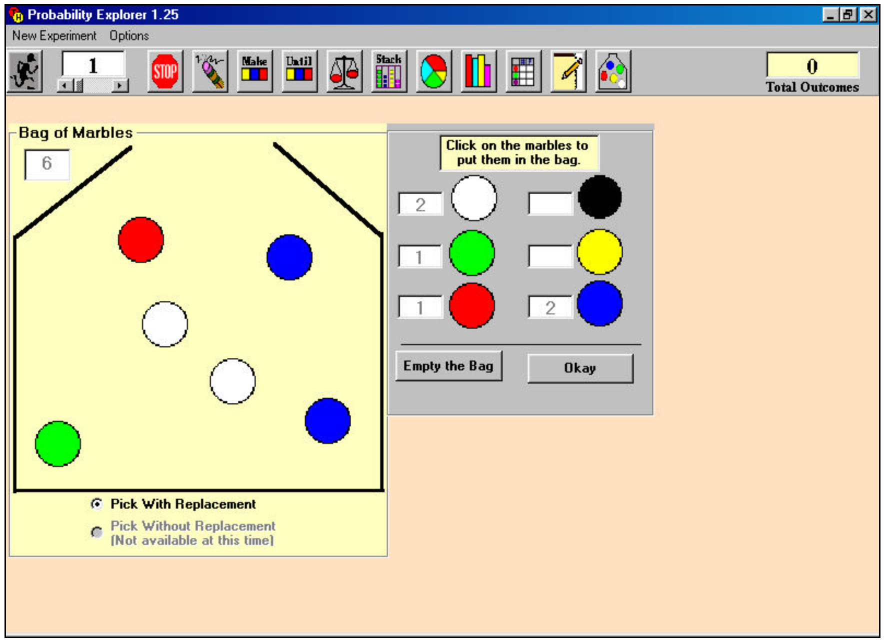


Figure 4.1. Marble bag tool added to microworld after the pilot study.

Actual Study

Three children (Carmella, Jasmine, and Amanda), all age 9 and entering the fourth grade, participated in the actual research study. The choice of using three children was a purposeful attempt to limit the scope of the study. The study was exploratory in nature and intended to provide formative evaluation of the usability, benefits, and drawbacks, of the microworld environment. By studying three children's problem solving and interactions with the software over a period of 6-8 weeks, I could critically analyze the development of the children's probabilistic reasoning. Each child formed the basis of a case study of the development of probabilistic reasoning skills with relation to meaning-making activities while solving probability tasks during the interviews and while using the *Probability Explorer*.

The research study had two separate but connected components. The first component involved task-based interviews to assess children's understanding of concepts such as fairness, equivalence, sample space, experimental probability, theoretical probability, probability comparisons, conditional probability, and independence. This assessment occurred twice, approximately 6-8 weeks apart. The purpose of these task-based assessments was to document the children's development of understanding in probability concepts prior to and following their use of the *Probability Explorer* during the teaching sessions.

The second component involved a computer-intensive teaching experiment (Steffe & Thompson, in press) in which the children worked in child-child as well as child-researcher diads to solve probability tasks. A brief outline of the teaching tasks posed during the teaching sessions is in Appendix F. The teaching sessions occurred in three 2-hour sessions and four 1-hour sessions. After the fifth teaching session, I made a decision to meet with the children individually for their last two teaching sessions. This decision was primarily based on the children's different developmental rates with probabilistic reasoning, and the difficulty that Amanda was having in interpreting and using several of the microworld tools. In essence, the last two sessions were microworld-based, task-based, semi-structured interviews with each child.

The focus of the computer-based component of the study was to gather evidence about children's peer interactions and use of the computer microworld during problem solving activities, and how those interactions reflected the children's process of constructing knowledge about probability concepts. In addition, I was interested in how, when, and why the children chose to use the various tools available in a computer environment and if those tools were effective in promoting appropriate probabilistic reasoning.

Data collection. Each interview was videotaped and audiotaped. In addition, any drawings or paper-and-pencil activity that the researcher and child did during the interview was part of the data corpus. All data was kept confidential, and each child was given a pseudonym that was used in all transcriptions and data reports.

During the instructional intervention, each instructional session was videotaped and a non-participant observer took field notes. The internal recording system was used to capture the children's actions on the computer. Any written work done by the children during instructional sessions was included as data. In addition, I kept a reflective journal in which I recorded thoughts on each instructional session, children's interactions and meaning-making processes, and which types of tasks might be developed and used with the children to further their development of normative probabilistic reasoning.

Data analysis. The analysis of data was done both during and after the research study. The analysis of each instructional session was used to inform the planning and development of materials for subsequent sessions. Following each session, I did the following: (1) wrote a reflective account of the session, including observations of interactions and initial interpretations of children's probabilistic reasoning; (2) met with non-participant observer and the other teacher/researcher to discuss the teaching session; (3) critically reviewed field notes, video-tapes, and computer video files to analyze children's meaning-making processes and made assertions about the children's probability understandings; and (4) used the analysis to inform the planning of subsequent activities for teaching sessions.

The post-analysis process was similar to the two-step process of Erickson's (1986) analytic induction model. The first step involved generating empirical assertions about what was happening in the learning environment. As stated above, those assertions were inductively generated during the instructional phase. The second step was to establish evidentiary warrants for these assumptions by systematically searching through the entire body of data, "looking for disconfirming and confirming evidence, keeping in mind the need to reframe the assertions as the analysis proceeds" (p. 146).

The audio tapes from each interview and teaching session were transcribed. While watching all video records, the audio transcriptions were then annotated to add descriptions of interactions and computer actions. All resulting annotated transcriptions were analyzed and coded using FolioViews (Version 4.2, 1998). The original coding schemes used were based on mathematical constructs (e.g., fairness, independence, law of large numbers, part-part reasoning, theoretical probability, multiplicative reasoning) and the tools used in the microworld during a particular investigation (e.g., stacking columns, pie graph, weight tool, data table). After the initial coding, several themes associated with the initial assertions were established based on the mathematical content and how the children utilized the tools in the microworld (e.g., effect of the number of trials, proportional reasoning, use of theoretical probability, and "evening out" of results). The coding and subsequent grouping of these themes provided confirming and disconfirming evidence for the development of the children's understandings with respect to these themes throughout the interviews and teaching experiment.

Establishing Validity

According to Erickson (1986), there are several threats to the validity of a study: (1) inadequate amounts of data to warrant an assertion; (2) inadequate variety of data; (3) faulty interpretation; (4) inadequate amount of disconfirming evidence; and (5) inadequate amount of discrepant case analysis.

The first and fourth threats to validity were addressed by the amount of time spent with the children gathering data. Since the data collection process spanned approximately eight weeks, the researcher was able to gather evidence of their meaning-making processes and social and digital interactions while solving a variety of different probabilistic tasks. These tasks were purposefully designed to bring out common probabilistic misconceptions and cause students to confront their beliefs and make sense of experimental evidence.

The second threat to validity was addressed by collecting data from several sources. The task-based interviews provided benchmark assessments of the children's understanding before and after the teaching sessions. This data was used in conjunction with the data gathered during instruction to provide evidence of each child's probabilistic reasoning abilities. During the instruction, evidence of social and digital interactions were gathered using two sources of video (computer video and regular video) recordings, audio recordings, non-participant observations, and my reflective journals. This variety of data collection methods facilitated a triangulation process of gathering evidence of children's understandings and meaning-making.

I addressed the third threat to validity by constantly focusing my attention on the amount of disconfirming and confirming evidence gathered, and modified subsequent instructional and post-interview tasks to facilitate the gathering of additional evidence. This process was facilitated by a peer debriefer, who also served as the other teacher/researcher during the teaching sessions.

In order to account for the fifth threat to validity, I systematically analyzed all discrepant cases to ensure that all possible reasons for the discrepancies were accounted for. In addition, the other teacher/researcher (and dissertation chair) assisted and advised during the analysis process to allow for a knowledgeable perspective on the interactions and problem-solving processes.

Concurrent Research and Development

Since this study was exploratory in nature and intended to serve as a formative evaluation of the microworld, there were several instances where further development of the microworld tools were done in between teaching sessions. Some of these developments were merely to fix “bugs” in the software. However, there were also instances when I added additional capabilities to the environment or enhanced current tools in such a way that I thought might facilitate further development of a child’s understanding. An example of the latter type of development occurred when Amanda was having extreme difficulty transferring her already weak notions of theoretical probability with physical objects (e.g., coins, dice, marbles in a bag) to the abstract numerical representations in the weight tool. To help her develop a better conception of theoretical probability and to make a transition from interpreting the theoretical probability with physical objects to that with the weight tool, I created a dynamic link between the weight tool and the marble environment (Figure 4.2). This link allows students to use the marble tool to place marbles in the bag and simultaneously watch the theoretical probability based on the contents of the bag automatically update in the weight tool as marbles are placed in the bag.

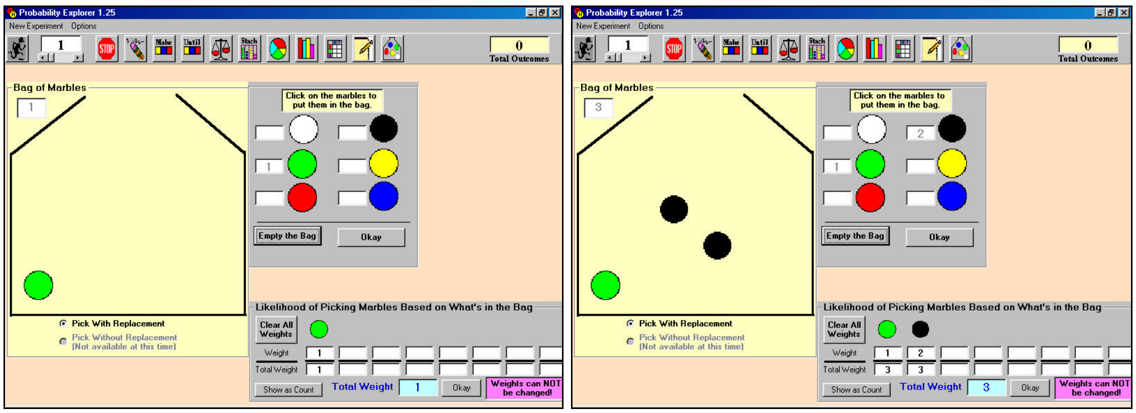


Figure 4.2. Link between marbles and weight tool to display theoretical probability.

Reporting Findings

Since one of the research questions was to understand children's development of probabilistic reasoning, each child formed the basis of a case study analysis. The case studies document each child's pre-and post-probabilistic understanding and include an in-depth discussion of three or four predominant themes for that particular child that seem to characterize their cognitive growth in probabilistic reasoning. Each case study will be discussed in-depth in Chapters 5 (Carmella), 6 (Jasmine), and 7 (Amanda).

The second research question concerns the effectiveness of the *Probability Explorer* microworld and computer-based activities in facilitating the development of normative probabilistic conceptions. Throughout the individual case studies, I will discuss how the computer environment and activities acted as an agent or deterrent in the development of that child's probabilistic reasoning. Based on a cross-case analysis of the students' work, I will make recommendations for further development of the software. In Chapter 8, I discuss the children's computer interactions and their development of probabilistic reasoning. Chapter 9 contains the summary discussion and implications for teaching, learning and future research and software development.

CHAPTER 5

THE CASE OF CARMELLA

Carmella (nine years old) is a very articulate, socially and academically confident fourth-grade student. She attends a rural public school just outside of a university town and is considered academically gifted by her school division's standards. Both parents hold Ph.D. degrees, one in biology and the other in engineering. Carmella has a pleasing personality and enjoyed being video and audio taped, often playing up to the camera and microphone during interviews and teaching sessions. She was very outspoken and excited during sessions in which she was interested; however, if she was bored or tired, she made those affects clearly known with her body language and explicit verbal statements. She was very reflective in her responses and communicated her thoughts openly during all interviews and teaching sessions. Carmella and her parents reported that she often used a PowerPC Macintosh computer at home and felt very comfortable working with different software applications (e.g., internet, drawing programs, word processing, various game-like math software).

Pre-Interview Analysis

The pre-interview occurred on August 17, 1999 for about one hour and consisted of the tasks listed in Appendix E. Throughout the interview, Carmella used the materials available to her (e.g., coins, bags with black and white marbles, spinners, buckets with cubes) and often used paper and pencil to organize her thoughts and record data. The language she used throughout the interview suggested that she has a strong sense of rational numbers as well as the vocabulary typically used to describe probabilistic situations (e.g., impossible, probably, likely, unlikely). I will briefly discuss our interactions during each task and summarize my understandings of her conceptions.

To begin the interview, I asked her to interpret the following hypothetical situation:

> Suppose, you and Jasmine were playing a game, and half way through the game Jasmine said, 'you know I don't want to play this game anymore. I don't think it's fair.' What do you think she means by the game is not fair?

In her response, Carmella provided an example of an unfair game that was designed for the first player to always win and she noted that for a game to be fair "each person would have an equal chance of winning." To explain "equal chance of winning," she replied: "So if you like played it 10 times, each person would win five times … so it really wouldn't matter who would win because it could just be a game of luck each time." Her concept of "fair" relies on an assumption of equal chances. In addition, she has an expectation that experimental results from a fair situation would result in an equal distribution of wins and includes the notion of luck as indicative of equally likely chance occurrences.

<u>Bucket of Cubes</u>

During the interview, several tasks were based on a bucket of colored cubes. The purpose of these tasks was to assess how Carmella described the chance of a certain color being randomly chosen, as well as her understanding of vocabulary such as "most likely" and "least likely."

Nine cubes. For the first task, I presented Carmella with a bucket containing four green, three red, and two yellow cubes. Carmella used a part-part comparison (e.g., four green is more than three red) of the cubes to correctly state that green was more likely to be chosen and yellow was least likely to be chosen "because there's only two of them." She then continued to discuss the chance of getting a yellow. (Note: C is Carmella, T is myself, the teacher/researcher)

> C: You wouldn't have a very good chance…it's like winning the lottery.
> T: Winning the lottery! Why wouldn't you have a good chance of winning the lottery?
> C: Because there are so many people doing it, and there is probably only like ten things you could actually win. So it would be like 10 out of a million chances.

She spontaneously related not having a very good chance of something happening to a real-world unlikely occurrence. She also used a part-whole statement with "10 out of a million" to quantify her illustration about the chance of winning the lottery. Her use of a real-world situation indicates that she has experienced, discussed, or explored probabilistic situations prior to our meeting.

To continue the task, I closed my eyes and randomly picked a cube out of the bucket.

> T: I got a yellow. You said I had the least chance of picking a yellow, but the very first time I got yellow. What do you think about that?
> C: It's just you. You wouldn't have a very good chance of it, but that doesn't mean that it's impossible…and if you try again you probably won't get another yellow (she said confidently).

Her use of "impossible" indicates that she differentiated between an event with the least chance and an impossible event. She also maintained her "least chance" theory by stating that yellow would probably not appear next time. In fact, I picked a cube (with replacement) two more times and got a yellow cube each time. The expression on her face was certainly one of surprise and puzzlement. She hypothesized that perhaps the yellow cubes were “sticking around the middle” and that I always picked from the middle of the bag. On the fourth pick, a red cube was chosen. I reminded her that she said green was most likely to be picked and that after four picks there were no green cubes drawn. She pondered this for a moment and just shrugged her shoulders saying “I don’t know.” The four results seemed to break from what she expected, and she was not able to offer a reason or conjecture as to why the results did not mirror the “most likely” and “least likely” choices. However, despite the experimental results differing from what she expected, her prediction of which color was most likely (green) did not change. This stability in her thinking demonstrates that she has a strong conception of *a priori* probabilities and that the results from our small number trials was not enough evidence for her to use subjective judgment to change a response she had formed objectively.

I continued the task by asking Carmella to describe the color with the best chance after I removed a green cube (leaving three green, three red, and two yellow). She again used part-part reasoning and correctly stated that red and green both have the best chance since they both have three cubes. I removed yet another green cube and without any prompts she stated “you probably might get a red better … it would be more likely that you get a red.” When asked which color was the least likely to get picked, she said that yellow or green were both least likely “because there’s only two.” In answering these questions, she always used the number of cubes in the bucket in a strict part-part comparison manner to determine most and least likely events.

Later in the interview, I brought this same bucket back out and asked her how she would describe the chance of picking a green cube. She immediately used a part-whole strategy and replied “four out of nine chance.” She used similar language to describe the chance of getting a red (3 out of 9) and yellow (2 out of 9). She only used references to parts when I asked her to choose the most and least likely color. However, she considered the total number of cubes in describing a chance of getting a particular color. It appears that she has had experience in stating probabilities in a part-whole format and used that format to quantify a statement of chance for an individual color; however she used part-part reasoning when comparing the chance of picking two or more different colors.

To assess her ability to use part-whole reasoning, I asked her whether the chance of picking a green cube from the original bucket of nine cubes had changed after I took out the two green cubes (modeling this with the cubes and bucket). The following protocol and my commentary show her intuitive references to the whole (total number of cubes) in her analysis of the problem.

> C: A two out of, no it would be a two out of seven chance … So is that … No, wait. Actually it would be the same amount because it would be the same chance because there’s less all and all. Because you took these two out. Wait. But now there’s more red ones. I’m confusing myself.

Her first instinct was to think that 4/9 was the same chance as 2/7 because it was two less “all in all” (i.e., less part and less whole). But she begins to rethink her response.

> T: That’s okay. Let’s try to talk it out. All right, so before when these were in here you said you had a four out of nine chance, and you also said that green was the most likely one to get picked. And now these two are out and you said the green had a two out of seven chance of getting picked.
> C: And then I got confused.
> T: All right, why did you get confused?

C: Because I wasn't sure if it was the same chance since it was less. But then I thought that it was more green, so it wouldn't be the most likely now it would be a less likely one. So then I got confused, totally, totally, totally.
T: Well let's take a look at the bucket. So if you had to tell me how the green compared to the other colors, how does the chance of picking the green compare to red?
C: It would be less likely. It would be less likely for the green to be picked than the red.
T: Because the green is less likely than the red? (she shakes her head "yes") Okay, so how does the chance of picking the green compare to the chance of picking a yellow?
C: It would be an even chance.
T: So how does the chance of picking a green from before compare to the chance of picking a green now? Is the chance the same? Is it more or is it less?
C: It's less.
T: It's less. The chance of picking a green has gone down?
C: Yes.
T: How does the chance of picking a red change? Or has it changed?
C: It hasn't changed … Well it's actually more likely now. Because now there's just two greens, so the one that is higher is now gone. Now there are three reds and two greens.
T: And so the chance of picking a red, has it changed?
C: Uh huh.
T: And how did it change?
C: Because it's gone sky high.
T: It's gone sky high. Okay. So the chance of picking a red has gone up?
C: Yes.

Although she at first did not think the chance of picking a red had changed, she used a part-part comparison to conclude that since red now had the largest part, it was the most likely, whereas green was the most likely before. Therefore she reasoned that the chance of picking a red cube had gone "sky high."

T: What are the chances of picking a yellow? These were in here. [I put the two green cubes back in the bucket] All right. So that's where we started. So now we take the two greens out.
C: Well the yellow has actually gone up. Because before there were four [green cubes] in there. And now they take those out, so it would have to be exactly the same and it's higher. But now it won't be the one that's the less. You would have one that would have just the same chance.
T: So you think the chance of picking a yellow has done what?

C: It has raised. Not totally, but it has raised a little bit.

Her initial attempt to compare the part-whole relationships for the green (4 out of 9 and two out of 7) proved difficult for her. This is no surprise since comparing fractions with unlike denominators is a difficult concept. It was easier for her to use a part-part strategy with references to the decrease in green cubes (which affects the part as well as the whole) to think through each question. The relative rank of each color in comparison to the others made the task more manageable for her. She correctly reasoned that the chance of picking a green had decreased, while the red and yellow had increased. This segment shows that Carmella came into the teaching experiment with powerful reasoning skills about part-part relationships, but less ability to reason with part-whole. Her part-part reasoning clearly helps her think about part-whole relationships. She also had some facility with both in making quantitative and qualitative judgments about theoretical probability.

<u>Four cubes</u>. Another bucket task involved three green cubes and one red cube. Again, when asked to describe the chance of picking each color, she used a part-whole statement for green ("3 out of 4") and "1 out of 4" for the chance of picking a red. I then removed a green cube and again asked her to assess the chance of picking a green cube. She replied "it would still be more [than red] because there's three cubes and the majority of them are green." For the chance of picking a red cube, she replied "It wouldn't be good." For these responses, she used informal quantitative comparisons and descriptions to justify her reasoning based on part-part relationships.

I then removed the red cube from the bucket (leaving only two green cubes) and asked her how she would describe the chance of picking out a green.

C: You have to.
T: What do you mean?
C: There’s no one else, just the green.
T: All right. Do you know any numbers to describe for me your chance of picking out a green?
C: It would be a two out of two chance, this meaning that you have two and if you took two away there would be none left. So you have to pick a green.

She continued to use part-whole reasoning to describe a certain event. Her reference to taking “two away there would be none left” seems to be a concrete reference to the physical bucket of cubes and the process of picking cubes without replacement, which would result in all green cubes picked with nothing left in the bucket. In this case, she explained “two out of two” and the concept of certain by referring to an actual physical experiment rather than merely stating that all the cubes are green.

Coin Tosses

Two types of tasks were used involving coin tosses. The first task was designed to assess the concepts of equiprobable and sample space while the second task assessed the concept of independence.

1, 2, and 3 coin toss. I gave Carmella a penny and asked her what different ways it could land if I flipped it in the air. She immediately responded “it could land heads or tails.” I then asked her if one of those results was more likely to happen. She shook her head in a strong “no” side-to-side action. The following protocol demonstrates her reference to a hypothetical experiment to justify her answer. What is interesting here is her intuition about the effect of the number of trials that she spontaneously adds to the discussion.

> C: Because there's just two sides. That's an even number. So probably if you flip it like a 100 times, it would be anywhere around 50-50.
> T: What do you mean anywhere around 50-50?
> C: Sometimes it's not exactly, but most of the time it's pretty close.
> T: Okay. So it's not exactly. So what if I flipped it 10 times?
> C: It would be a lesser chance of it getting exactly. Because the more you do it, the more of a chance.
> T: So what if I did it a 1000 times?
> C: You're pretty likely to get it even, even.

She defended her belief in equiprobable with a rationale about two being an even number and used a hypothetical experiment to explain even chances. Her use of "anywhere around 50-50" indicates that she has a sense of the variability expected with random experiments. She also believed that it was harder to get exactly even results with only 10 flips (as compared to 100) and her reasoning of "the more you do it, the more of a chance" suggests that she had intuitive ideas (formal or informal) about the phenomena of the law of large numbers. Two important ideas in Carmella's development of probabilistic reasoning emerged during this segment. First, she used a hypothetical experiment to explain theoretical probability. I will refer to this as her "hypothetical experiment strategy" (HES) from here on. Second, she brought up notions of the law of large numbers and that a large number of trials tended to result in a percentage close to the theoretical probability. Her intuition seemed to tell her that with 1000 trials you are pretty likely to get "even-even," yet it was unclear if she meant exactly even or merely close to even. This issue emerged continuously throughout Carmella's work in the teaching experiment.

For the next task in the interview, I asked her what the possibilities would be if I flipped two coins at the same time. She easily stated "you could have two heads, two tails, or a heads and a tails or a tails and a heads." Although this was a compound event, she quickly constructed the sample space. I asked her if HT was different from or the same as TH. She modeled her response with the actual coins (a penny and a quarter). “You could flip them and you could cross them, then one could be heads (quarter) and one could be tails (penny)…and then if you switched them in your hands (she turned the coins so that the quarter now showed a tails and the penny a heads) it would be the other way around.” Her ease of differentiating between the order of results was facilitated by her demonstration with the different coins.

> T: With these four choices that you have here, are any of them more likely to happen than the others? [she shakes her head “no”] No? Why not?
> C: Because it’s two coins, they both have two sides. And so then there would be four sides. That’s an even number. So the more times you would flip it, it would probably be close or even.
> T: What do you mean by close or even?
> C: Like if you flipped it 50 times, then one might have 40 and the other 60. But they are still pretty close.
> T: And so when you say 40 and 60 are you talking about?
> C: This [quarter] would flip and it would land on heads a few times. Or this [penny] would flip on tails certain times.
> T: So over here, what do you think are the chances of getting both tails compared to getting a heads and a tails? [I point to TT and HT]
> C: They would be an even chance.
> T: What about getting both heads compared to getting both tails?
> C: It could happen either way.

The reasoning she used about all four possibilities being equally likely relied on her knowledge of an even chance for heads or tails with one coin. She seemed to be extending that knowledge to this compound situation and hinting at the notion of independence. Her reference to four being an even number indicates that she may have over-generalized “even-even” to this situation and may not really be considering ideas of independence in her analysis. I followed up with these ideas during the teaching experiment.

When asked to list the possibilities for flipping three coins (penny, quarter, nickel), Carmella started with a very systematic strategy. She listed HTT, then “changed the order” to get TTH and THT. She paused, wrote HTH, then put her pencil down. When asked if she could convince me that she had found all the possibilities, she told me about her “changing order” strategy to get the first three results and then how she just switched all the tails to heads in THT and vice-versa for the last result of HTH. When I asked if the coins could land so they were all the same she replied “oops” and wrote down HHH and TTT. She again maintained that she was sure she had all results. Although she began with a sophisticated strategy, she did not apply it for the case of two heads and one tail.

To assess her ideas about independence and probability, I asked her if any of the six possibilities she had listed were more or less likely to occur than the others. She nodded a strong “no” and replied “they all have two sides and so it would be an even chance that they would land on either side … so it wouldn’t matter how many coins you have.” Again, she seems to be expressing ideas of independence here. Since each side of the coin has an even chance, she feels that any of the listed possibilities must have an even chance of happening.

<u>Flipping a coin six times</u>. Later in the interview I asked Carmella to predict the outcomes if I flipped a penny six times. She quickly replied “around three and three” but denied my request to predict the results in order “because it would be hard to get it right … it could do anything … the coin has its own mind.” She has a strong notion of the unpredictability of a random event. I flipped the penny six times and got THHTHT. I asked her what she would have thought if I would have flipped six tails. She replied “It

would be pretty unlikely, but not impossible … it would be more likely than getting 100 tails, because it would be more of a chance for the heads." Again, Carmella brought up ideas about the effect of the number of trials and hinted at the concept of the law of large numbers.

For the next series of questions, I showed her four possible results from tossing a coin six times. For the first set (HHHHTT, THHTHT, THTTTH, HTHTHT), she proclaimed that all the results were equally likely "because there's two sides so you could have it on either side." She used the "same reason" for stating that the next set of results were also equally likely (HHHTTT, HHHHHH, THTHTH, HTHTHT). She seems to have a strong notion of independence with the coin tosses. She continued this line of thinking and did not fall prey to the "gambler's fallacy" when I asked her, after flipping THTTTT, if I was more or less likely to get a heads or tails. She retorted "there are two sides and it could land on either of them." She seems to be very conscientious about considering each coin toss independently of each other.

Sampling

For one task used during the interview I presented Carmella with a black bag containing five blue, three red and two yellow tiles. The purpose of this task was to assess her sampling strategies and to determine her level of confidence in experimental results with relatively small samples. The only information I gave Carmella about the bag was that it contained 10 tiles and there were three different colors of tiles. Her task was to make a reasonable guess at what was in the bag by picking a single tile from the bag and then replacing the chosen tile back in the bag before the next pick. She wanted to choose

10 tiles "because there are 10 squares in there and I can write down what I have each time so then I can find out how many different ones there are under each color." Although it appears that her strategy initially ignores the "replacement" part of the sampling procedure, after one sample, she immediately recognized a problem with her strategy and replied "what if I pick out the same one? Am I allowed to write on the cube?" I would not allow her to write on the tiles but she continued to choose 10 tiles with replacement and sampled five blue, three red, two yellow. In reflecting on her results, she thought there were too many blue tiles compared to the number of red and yellow tiles and thought it would make more sense to have a more "even" distribution. She continued to say that in order to be more confident of what was in the bag, she would like to write on the tiles once they are drawn and to choose a sample of 30 "because then I could find out, because the most I think I could pick one [tile] three times." Her suggested sampling strategy was quite sophisticated and indicates that she recognizes the variability possible with a small sample done with replacement. Although what she sampled was actually the exact distribution of the colors in the bag, she still claimed that she had "insufficient data" and that she probably would not get that exact distribution again if she did another sample.

<u>100 Gumballs</u>

This task was taken directly from the 1996 NAEP exam and was used to assess her ability to use proportional reasoning and theoretical probability to make a prediction for a sample when the population is known. The task referenced a gumball machine with 50 red, 30 blue, and 20 yellow gumballs. If the gumballs are well mixed and someone picks 10 gumballs from the machine, how many of each color gumball is expected? Carmella predicted there would be five red, three blue, and two yellow gumballs because

> you only put in 100. So five of them are red, two of yellow, and three of them are blue. Fifty of them are red, twenty of them are yellow, and thirty of them are blue, but you are only picking out ten. So it would be more like take off the zero from 100, and the zero off from 50, and the zero off the 30, and the zero off the 20.

Although her prediction is in proportion to the population of 100 gumballs, her “take off the zero” strategy does not sufficiently indicate that she was actually using proportional reasoning. She did not mention anything proportional such as 50 being half of 100 and five being half of 10. Without further explanation on her part, it is difficult to say whether she made her prediction based on proportionality or only on a pattern recognition.

Marble-Bag Comparisons

The marble bag comparisons are similar to tasks used by both Piaget (1952/1975) and Fischbein (1975) in their research. The bags have varying amounts of black and clear marbles in them. I had pictures of the bags drawn for the interview, although I had real marbles on hand that could be used if a child needed to see a physical model of the problem. The purpose of these tasks is to assess how the child compares parts and wholes of two different bags and if any type of proportional reasoning is used to make comparisons about the chance of picking a certain color.

For the first two pictures of bag of marbles, I only asked Carmella to compare the chance of picking a clear marble to that for a black marble. Bag #1 contained two black and two clear (2B2C) marbles. When I asked her to describe the chance of picking out a black marble she replied “it would be even with the count because there would be four in there … there are four in there and two of them are black and two of them are white. So if you reached in there and mix them around, you might get a clear one or you might get a black one." She used the parts to illustrate her point about "even with the count" and her statement of "might get a clear one or you might get a black one" indicates she feels that neither black nor clear has a better chance of being picked. In her assessment of bag #2 (5B3C), she again used a qualitative description and stated that "you would have a way better chance of picking out a black one" and used the parts (5 and 3) to justify her response. As in earlier tasks, she used part-part reasoning when making comparative statements about the chances of different events occurring.

For the remaining pairs of bags (#3 & #4, #5 & #6, #7 & #8), I asked her to choose which bag she would prefer to pick from if the goal is to try to pick a black marble. With each pair of bags, I reiterated a question such as "would you like to pick from bag #3 [3B1W], bag #4 [6B2W], or does it matter which bag you choose from?" For the first pair of bags, the distributions were proportional. She used a part-part strategy and first thought having larger parts would be better (2 and 6) because there were more blacks but then reverted to bag #3 with the smaller parts (1 and 3) since there was less clear marbles. Even when she used part-whole language to describe the chance of picking out a black, she maintained that three out of four is better than six out eight because "this one [pointing to bag #4] seems like it would be better because there's more. So there would be more black marbles, but there are more clear marbles." It appears that she based her decision on the quantity of clear marbles (the undesired event) because they could detract from the chance of choosing a black marble (the desired event). This suggests she was using additive reasoning to compare the parts within each bag and then used an additive strategy to compare the within-bag differences between bags.

Bag #5 and bag #6 were also designed proportionally with 1B4C marbles and 2B8C marbles, respectively. At first Carmella said "I don't think it would matter…this is more [pointing to bag #6] … wait a minute … it wouldn't matter." Her explanation that follows shows very good proportional reasoning. Although she first may have thought that more clear marbles in bag #6 were going to affect her decision, she quickly abandoned her previous part-part reasoning and adopted a valid proportional strategy using part-whole reasoning. In response to my asking why it did not matter which bag she choose from, Carmella offered the following:

> C: Because this [bag #6] is pretty much the same as this one [bag #5]. Because this one is 10 and this one is five. You can just put two of these [pointing to bag #5] together and it makes the same one [as bag #6]. So you would have the same chance as picking out a black one.
> T: Oh? So in this bag compared to this bag you have the same chance. And you were saying there are 10 marbles in here [pointing to bag #6].
> C: And there are five marbles in there [pointing to bag #5]. And it's like the same pattern. Because there would be one out of five marbles that would be black. But this is just two out of ten marbles. So it wouldn't matter at all.

By recognizing a pattern, she was able to see the proportionality of the marbles between 1/5 and 2/10 in this task and correctly decided that it did not matter which bag she would choose from. I conjecture that the unit fraction relationship (1/5) may have been easier for her to compare proportionally to 2/10 than the 3/4 to 6/8 relationship. Perhaps her experience with unit fractions was more extensive than her experiences with non-unit fractions like 3/4 and 6/8.

The last two bags were not proportional with 2B2C and 2B3C in bag #7 and #8, respectively. This task was designed to see if children take into account the number of clear marbles when the number of black marbles are equivalent. Upon being presented with the pictures, Carmella immediately pointed to bag #7 and explained why she made her choice.

> C: Because it [bag #7] has an even chance. I admit that it doesn't sound too good if you want to pick out a black one. But this [pointing to bag #8] has a lesser chance of you picking out a black one…because this one has four [pointing to bag #7] and two of them are white, well clear, and two of them are black. And this one [bag #8] has five, three of them are clear and two of them are black. So even though having an even chance is not good, if you definitely want to get a black one, it would be better than having a lesser chance.

Carmella recognized that when the number of black marbles are constant, a bag with one more clear marble lowers the probability of picking out a black marble. Although she did not use any direct part-whole comparisons, her part-part reasoning indicates that she did consider the total contents in both bags in making her decision.

Spinner Game

The final task (taken from Jones *et al*., 1997) consisted of a game played with a spinner (1/2 red, 1/3 blue, 1/6 yellow) and eight pennies. Each player starts with four pennies and chooses a color on the spinner (the child chooses first). The players take turns spinning the arrow. If the arrow lands on a player's chosen color, the other player must give the selected player a penny (e.g., If Carmella chose the red sector on the spinner, every time the arrow lands on red, I give her a penny). The game continues in this manner until one player has all eight pennies.

Carmella chose the red sector and I then chose the blue sector. Before we began to play, I asked her if this is a fair game. She stated that she had a better chance "because if you split it in half, then one-half would be all red and the other half would be, some of it would be blue, and some of it would be yellow." We played the game a few times and got the following sequence of results (BYRYRBRB). At this point Carmella said "this is like a game of back and forth." I asked her to predict who she thought would win.

> C: I, well if you do it scientifically, it would probably be me. But it would be hard to tell at the time being because it looks like things are going back and forth.
> T: And what did you mean by doing it scientifically?
> C: Because, well really, usually you would think that it would just go I would land on it, or something, because mine is bigger than yours. But it's not really going like that. It looks like it's just going to be, it looks like you could have the whole half to yourself because it's just that we are going back and forth, back and forth.

Her reply indicates that she would normally base her predictions on the theoretical probability and sample space. This shows evidence of *a priori* thinking on her part with regard to experimental outcomes. Yet at the same time she is using subjective probabilities based on the known outcomes and the "back and forth" results to readjust her prediction and state that it is "hard to tell." Recall that she did not let experimental evidence change her opinion about the likelihood of picking a green cube during the bucket task when I repeatedly picked a yellow cube (the least likely event). She seems to be able to rely on subjective reasoning, based on experimental data, for making predictions, but does not let experimental results sway her opinion about the likelihood of an event occurring based on *a priori* knowledge. This level of thinking is quite advanced for a 9-year-old. It shows that she can intuitively use theoretical probabilities to make predictions and can critically use experimental data to update predictions.

I told Carmella that the next time we play, she would get to choose a color and I would get the remaining two colors. She said she would pick red and I should get the blue and yellow. She claimed that now the game would be fair because we both have half the circle. To further illustrate her point, she offered that red versus blue and yellow would be the only way the game could be fair because if she had picked blue, then she would only have one-third of the circle. When I asked her to show me why the blue was 1/3, she used the spinner arrow to make an imaginary line (see Figure 5.1) that would visually split the red and yellow areas combined into two 1/3 pieces. This visual proof suggests that she has had experiences with representing 1/3 with a circle model. It also shows the sophisticated level of her rational number sense and how she could easily apply it to her probabilistic reasoning.

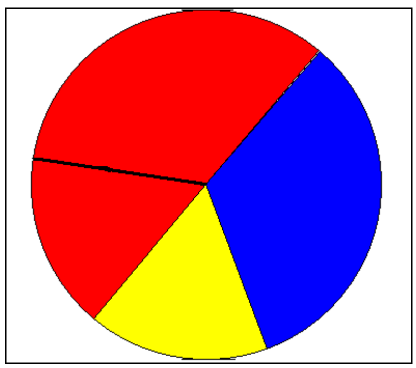


Figure 5.1. Carmella's visual proof that blue is 1/3 of the circle.

The final task in the interview involved Carmella comparing two spinners and deciding which one she would rather use if she won when the spinner landed on red. Spinner A was the same one used in the previous task. Spinner B contained 12 equal sectors, six red, four blue, and two yellow spaced in the following pattern (r, b, r, y, r, b, r, b, r, y, r, b). Carmella first tried to visually "squish" the red sectors together and said that together the red parts where slightly more than half the circle so she would choose spinner B. To see if she could reason numerically (and proportionally) about the spinners, I asked her to count the red sectors (6) and to also tell me how many sectors total (12). She then said that the two spinners had the same amount of red since half of them in spinner B were red. She then concluded that it didn't matter which spinner we used except that the game might take longer with spinner B because "there is more of a chance to land on a line." Her analysis in this task indicates that she has proportional reasoning skills that she can apply in both a geometrical and numerical situation. Her additional note about the lines also shows an awareness of how the number of times an element appears in a situation can affect the probability of that element occurring in a random simulation.

<u>Strengths and Weaknesses from Pre-Interview</u>

Carmella had a strong facility with simple probability concepts (fair, independence, theoretical probability) and used probability language appropriately. Her conception of "fair" was strongly grounded in her expectation of "equal chances." She seemed to associate random outcomes in fair situations with notions of "luck" and "anything can happen." In addition, she consistently used a hypothetical experiment strategy to explain the meaning of theoretical probabilities. Her understanding of theoretical probability was closely tied with her expectations of a distribution that closely mirrors that theoretical probability.

Carmella used both part-part and part-whole reasoning and appears to have facility with simple fractions. She tends to use part-part reasoning when comparing the chances of two or more events and uses part-whole reasoning for determining the chance of a single event; however, she relies on the part-part reasoning to help her with any part-whole comparisons. Carmella also attempted to use "proportional" reasoning when comparing chances of two or more events. However, she mostly used additive reasoning to justify equivalence and unequivalence of chances. She only correctly used multiplicative reasoning in her comparison with the 1B4C and 2B8C bags of marbles and her prediction of 10 gumballs from the gumball task. Although she displayed some appropriate multiplicative reasoning, using additive reasoning (in multiplicative situations) is a predominant strategy for her.

Carmella had a good understanding of sample space for simple events and could list a sample space for a 2-event situation. However, her combinatoric reasoning was not developed enough for listing a sample space of a 3-event situation. Although she began with a systematic strategy, she did not apply it to all combinations in order to produce all possible permutations.

A very strong understanding of independence was evidenced in many of Carmella's responses. She recognized that with every random act (e.g., flipping a coin) the chance of every possible outcome occurring is not affected by previous chance occurrences. Her response to the task with several strings of coin flip results differed greatly from the responses typically given by students of *all ages*. It is highly unusual that a 9-year-old child would have such a grounded conception of independence. In addition, she seems to have an intuition about the effect of a large number of trials on the probability of an event occurring. (Recall her comment about 100 consecutive heads being a lot less likely than six consecutive heads). She also has an intuition that it is more likely to get exactly "even-even" with 1000 flips of a coin and much less likely with 10 coin flips and is beginning to think about the effect of a large number of trials. Again, this presents evidence of her developing probabilistic reasoning and her ability to critically analyze chance situation.

Carmella's Meaning-Making Activity with the Microworld

Carmella participated in approximately eight hours of small group teaching sessions and three hours of individual sessions. The analysis of all teaching sessions brought forth several key themes that were critical in her further development of probabilistic reasoning: 1) her "total weight" approach to probability tasks; 2) her understanding of the "evening out" phenomenon; 3) her struggle between "close" and "exact;" and 4) her use of proportional reasoning. I will report my observations and analysis of Carmella's meaning-making activity and mathematical ideas, intuitions, and conceptions through descriptive vignettes to illustrate each key theme. Within each vignette, I will highlight Carmella's use of the microworld tools to demonstrate her reasoning and problem solving strategies with the aid of the digital environment as a simulation and analysis tool. Following the elaboration of each key theme, I have included a thick description of a teaching episode that highlights Carmella's problem solving and the interconnectedness of the four themes.

The Total Weight Approach

Recall Carmella's use of a Hypothetical Experiment Strategy (HES) in the pre-interview. She used a HES to describe "equal chances" and to make a proportional prediction of experimental results that mirror the theoretical probability (e.g., for a fair coin toss, she described the chances by saying “if you flip it 100 times you would probably get 50-50”). She also used this strategy on the first day of the teaching experiment when describing the chance of picking a black marble out of the 2B2W bag (this was a physical bag and not a digital representation of one). Carmella compared the chances for black and white as “50-50” and interpreted this to mean “if you have a hundred of them, 50 would be black and 50 would be white.”

During the third teaching session, Jasmine and Carmella were experimenting in the microworld with a 2B2W bag of marbles. After a few sets of 10 trials, Jasmine offered “I wonder if we just pick four out if it would come out as two and two.” Here, Jasmine used a HES strategy to predict the results of an experiment that directly mirrors the contents of the bag. She and Carmella changed the number of trials to four and ran an experiment. The following dialogue ensued. [Note: the girls had the pie graph and data table displayed during the following experimentation].

[J-Jasmine, C-Carmella, A-Amanda, T-Holly (Teacher 1), T2-Joe (Teacher 2)]

> J: Two and two. [reporting results from a trial of 4] [Jasmine then clears the screen and does another experiment of four trials resulting again in two black and two white marbles]
> T: Two and two.
> [Jasmine and Carmella run another set of four trials]
> J: Three and one. Unlikely. [the pie graph displays ¾ black and ¼ white]
> C: Well it’s actually more likely than what we were doing before.
> T: And why is it more likely to what you were doing before?
> [Carmella runs another trial of 4]
> C: See look. We got all black.
> J: We got all black! We got all black, with four of them. [directed at Amanda and Teacher 2]
> T2: Oh ... with four of them.
> T: Yeah, Jasmine had a conjecture that ... What Jasmine said was ‘if we go down and do four of them will we get a lot of two and twos?’
> J: And we did.

> C: But getting all blacks is more likely on four because ... there's less numbers, so it would be more likely that we would get that than if we had 10.

Jasmine's conjecture and Carmella's subsequent explanation provide evidence that both girls have a strong intuitive feel for the relationship between the number of trials and the chance of getting certain distributions. With only five possible combinations, Jasmine is certainly correct that there is a better chance of getting 2-2 with four trials than 5-5 with 10 trials. In addition, although Carmella did not finish her justification, she seemed to recognize that getting a 75%-25% distribution (as represented in the pie graph) with two equally likely outcomes was more likely to occur with a total of four trials than with 10 trials. She also had a strong sense that it was easier to get all of the same color marble with four trials than when with 10 trials.

Jasmine's suggestion of doing an experiment that exactly mirrored the contents of the bag seemed to prompt Carmella to expand her use of a HES to include an approach to experimentation with the number of trials equal to the total number of marbles in a bag. While exploring a 3B1W bag, Carmella and Jasmine correctly described the chances as not equiprobable in favor of a black marble. When asked to design an experiment and predict the results, Carmella suggested "I think we'll do four and three of them will be black and one of them will be white, or four of them will be black." She approached the task by designing an experiment and predicting results that exactly mirror the contents of the bag. Her addition of the possibility of four black marbles indicates her intuition that having a greater chance of choosing a black marble increases the chance of picking all four black marbles.

During the fourth teaching session, Carmella expanded her use of HES for describing theoretical probability to interpret the weights shown in the weighting tool. The first time Carmella discovered the weight tool (during an experiment where I had secretly weighted a coin toss as five heads to one tail) she interpreted the weights as “if we have six then five will be heads and one tail.” She subsequently ran six trials and got her exact prediction of five heads and one tail. When given the choice to design an experiment throughout the rest of the teaching experiment, Carmella preferred to use a Total Weight Approach (TWA) to design experiments with the number of trials equal to the total weight used in the weighting tool. In addition, she emphasized that the exact distribution of the weights used in the weight tool would be the “most likely” experimental distribution of results (e.g., with a weight of two heads and six tails, if eight trials are done, the most likely result is two heads and six tails).

Carmella continued to use a HES and TWA in interpreting weights and designing experiments. In the sixth teaching session, Carmella designed an experiment with two equiprobable outcomes, the sun and the rain. I had her display the weights as fractions (1/2, 1/2), and interpret the meaning of the fraction (see Figure 5.2).

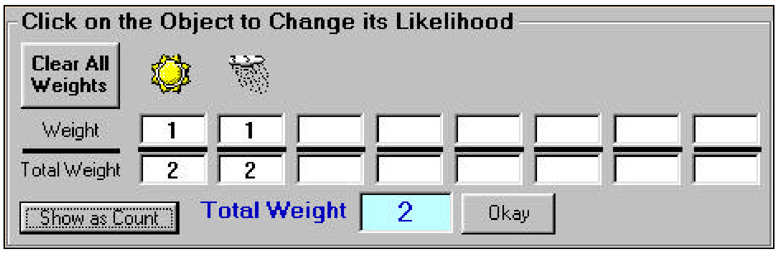


Figure 5.2. Part-whole display of weights in weight tool.

> C: It means that if you were to press this [points to the “run” button] twice, then one of them would be the sun and one of them would be the rain, most likely.

T: Most likely. Okay and why is that most likely?
C: Because the weight is one and one. And then the total weight would be two. And one is divided, and two is divided into one. And that's most likely because there is no guarantee.

Carmella interpreted the weight for each outcome to mean that two random events (because there are a total of two weights) would most likely produce results equal to the weight for each outcome (e.g., 1-1 in this case). Carmella also expressed that for 100 trials, she believed 50 sun and 50 rain were the most likely to occur. Although she considered the possibility of getting 100 suns, she correctly believed that result was "not very likely."

Later in the sixth teaching session, Carmella designed a 2-event experiment with three outcomes (sun, cloud, rain) and used the Make It tool to construct all possible ordered outcomes. After she made all nine possibilities, she opened up the data table and I asked her to interpret the one in the count column and 1/9 in the fraction column beside each of the nine possibilities (see Figure 5.3).

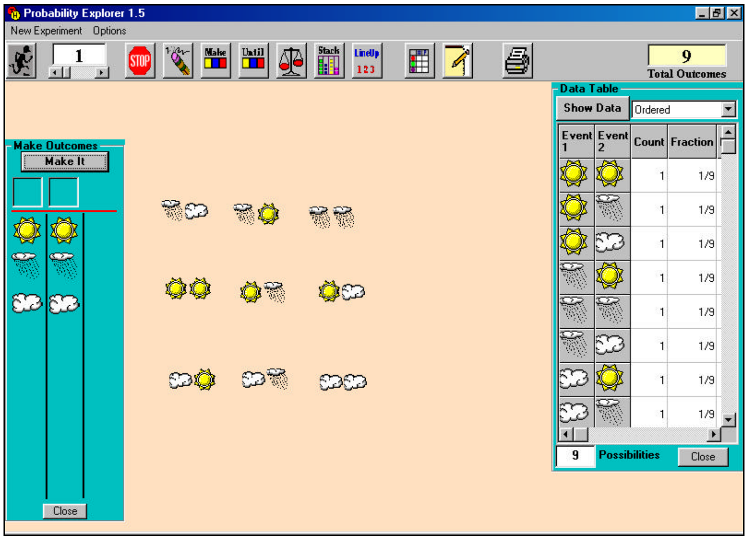


Figure 5.3. Nine possibilities in 2-event experiment with three outcomes.

Carmella interpreted the “1” and “1/9” as “the 1’s mean that it was all of them, and you have one of each. And since we have nine of the ones we have, and this is just a fraction so it’s one-ninth. That means that if we would like run it nine times then there would be one of each of them.” After I emphasized “we *will* have one of each?” she added “most likely.” In addition, she maintained that all nine possibilities were equally likely. She again used a HES to explain what the fraction 1/9 meant and used a TWA to justify that using nine as the number of trials would make getting one of each of the nine possibilities the “most likely” occurrence.

What began as a strategy for describing “equal chances” and theoretical probability without the microworld tools carried over as a strategy for interpreting weights in the microworld and led to an experimental approach with the software for predicting the most likely distribution of results. Carmella used both the HES and TWA consistently throughout the teaching experiment. In fact her reliance on the TWA for most likely producing the “right” (as reflected in the weights) distribution often conflicted with her intuitions about the “evening out” process. Recall her description of tossing a coin 1000 times being more likely to result in “even-even” than 100 tosses. This conflict led her to further struggle with difference between “close” and “even” results that began in the pre-interview. In the next section, I will describe her conceptual development of the “evening out” phenomenon (EOP). Carmella’s use of the TWA and her understanding of the EOP contribute to her struggle between “close” and “even” results. I will describe this struggle in more detail in the third section.

The Evening Out Phenomenon

Beginning in the first teaching session and continuing throughout the rest of the sessions, the students were fascinated with what Carmella termed as the “evening out” phenomenon (EOP). With every chance situation we investigated, the girls wanted to run a large amount of trials to test, watch, and confirm that the experimental results would “even out.” The EOP is comprised of two components from the students’ perspective: 1) the *process* of “evening out” that occurs during a simulation and can be visualized numerically and graphically; and 2) the working hypothesis that with a large number of trials, experimental results tend towards the distribution expected from the theoretical probability. Although the EOP was initially based on equiprobable events that tended toward an equal distribution of experimental results, the students were able to extend the idea of the EOP to non-equiprobable events in subsequent experiments. Consider the four static graphs in Figure 5.4 that represent the EOP with 80 trials of a standard coin toss.

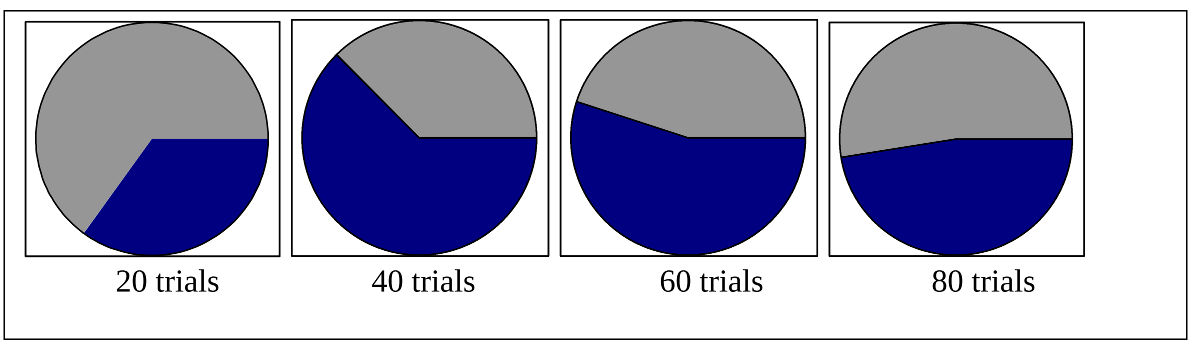


Figure 5.4. Graphical visualization of the EOP with 80 trials of a coin toss.

The discovery of the EOP occurred during the first teaching session. The students were each working at their own computer station running sets of 20 coin tosses. They were using the stacking columns, data table, and pie graph to pictorially, numerically, and graphically analyze the data during and after the simulation of 20 trials. At first they were all hoping for close to “even” results (10 heads-10 tails) and would report out their results after each simulation (e.g., 10-10, 9-11, 12-8, 7-13). At one point, the other teacher/researcher (Joe) asked if anyone could get a 20-0 result. They then focused their attention on trying to get the most dramatic differences in results. Amanda got a 17-3 at one point and Jasmine commented that she could tell if she got a 20-0 by only looking at the pie graph because “it would be all blue or all gray.” This comment sparked Amanda to continuously press the Run button to do many trials of 20 to see if she could get an all blue or all gray pie graph. Since she did not use the Clear button to erase the previous set of 20 trials, the number of trials was cumulative (i.e., 20, 40, 60,… 200). Although her goal was inappropriate, I took the opportunity to have the students gather around her computer to watch the pie graph as Amanda ran a large number of trials. The following dialogue illustrates how the visualization of the dynamic pie graph during the simulation prompted Carmella to make a hypothesis of the EOP.

> A: Well, it’s staying in the same place pretty much.
> T: Why do you think it’s staying in the same place?
> A: Because…
> C: Because she’s running it so many times, it’s like evening out.
> T: Really? Why is it evening out?
> C: Because it’s so many of them and …
> J: Look how much you’ve done it [There are about 1000 trials at this point.] – still going.
> T: So, Amanda do you think you’re ever going to get all blues or all grays?
> A: No.
> T: Why not? Why couldn’t we have a pie graph be all blue or all gray?
> C: Because it evens out with how many you do.
> T: What do you think about her theory about evening out? Why did you say that?
> C: Because the more you do, the more the chance to even out
> J: It would have to be all heads or all tails to be all one color.
> C: And look how many she has [points to number of trials]. I doubt one of them would be all the same

> T: Stop right there. Taking a look at this graph here … she's done 2,020 trials – how many heads and how many tails do you think she has in here?
> J: A lot!
> C: About even.

Carmella and Jasmine returned to their computer stations and ran a large number of trials, observing the graph, and reporting out statements such as "its still staying" and "they're not hardly moving at all." Jasmine explained that the graph has less changes because "with the more coins getting tossed, it's keeping it the same." She also justified the evening out *process* by saying "with more coins one is still going to be a little bit ahead of the another mostly, but it's unlikely that one will rise a lot above the other." Carmella added "there's so many it can't do it … it's like it evens out." The last two statements illustrate their understanding of the EOP in terms of the number of trials and the tendency towards the expected distribution. Their notions of the EOP are actually central to their understanding of the law of large numbers.

The discovery of the EOP with a fair coin toss naturally sparked a curiosity in the girls to want to test out the EOP with other experiments. During the second teaching session, we were experimenting with a regular six-sided die in the microworld (Carmella was working at computer station alone while Amanda and Jasmine shared a computer). We had done several sets of 10 trials and the girls were anxious to try a larger number of trials to add to the current set of 10 trials on their screens. In her prediction, Jasmine said the results would stay about the same as she swept the mouse pointer over the data table (she currently had zero 1's, zero 2's, five 3's, two 4's and three 6's). Carmella chimed in that the other numbers on the die would appear first and then they would stay even. It seemed that Jasmine was overgeneralizing the EOP while Carmella was able to reason that all the results should occur before the EOP would begin. I allowed them to choose how many trials to run to test their predictions. Carmella said she wanted to systematically double the number of trials and see the results. She ran several sets of 20 trials, each time using the stacking columns and data table. She then jumped to running sets of 40 trials and still continued to stack the icons and look at the table. She made several comments about the wide variability in her results. Although Amanda and Jasmine had starting doing a much larger number of trials (100), Carmella wanted to systematically test out the EOP with increasingly larger number of trials. This action suggests that she consciously wanted to test the effect of the number of trials on the EOP.

A little later in this session, I asked the students to predict what the data and graphs would look like if we did 200 trials with the graphs and table open. The girls used their memory of the dynamic changes in the graphs and table to visualize the simulation.

> J: I think that it's going to explode with dice. And they are going to stay around even.…
> C: I think that when it runs there are going to be a bunch of them everywhere. And this will start calculating [pointing to data table]
> T: That's going to start calculating. What do you think the pie graph is going to do?
> C: I think it's going to start ZWOOOM! [raises out of her chair] and spinning and spinning [rocking head side-to-side].
> J: It will go spin, spin, up and down, up and down, spin, spin. [using her forearm to mimic an up and down swinging motion]

Carmella's and Jasmine's description of the pie graph clearly indicates they expected a wide variability in the results at the beginning of the experiment. However, neither of their descriptions indicated an expectation of the EOP. When the simulation started, Carmella described the changes in the graph as "doing the hula" and Jasmine quickly copied her expression. After about 100 trials Carmella noticed the slower rate of change in the pie graph.

> C: It is slowing down now
> T: What do you mean by it's slowing down.
> C: Well when it first started it was wiggling
> T: What was wiggling?
> C: The lines that were separating because it was forming other categories and it got more and now it's slowing down a lot [at about 150 trials]. Because there's more and it's just a little bit more now.

Carmella's description demonstrates her connection between the addition of new categories, the increasing number of trials, and the EOP. At this point I asked the girls to predict what would happen if we did an additional 200 trials on top of the 200 trials already on the screen.

> J: I think that there are 200, and there's going to be a lot more, but it's still going to stay even.
> C: Well at the beginning, I guess it's going to start wiggling a little bit more, like it was before. And then when we get close to the end it will start slowly evening out.
> T: So Carmella thinks it's going to start wiggling again like it did before. And then start evening out when we get closer to the 400. What do you guys think about that?
> A: Same with me.
> [Jasmine nods in agreement]
> T: Same with you? Let's go ahead and hit Run and watch it closely.
> C: Or maybe it will hardly wiggle at all. [She says this after watching her graph during the first 50 additional trials.]
> A: It's hardly wiggling at all for us, too.
> T: So why isn't it wiggling very much? You said that it's going to wiggle and then settle down.
> J: Oh, because they are already flattened out. They are already a big number. There's already a big number.
> T: Uh huh.
> J: And then when this time it's getting to be a big number and it slowed down because it was a big enough number.

Carmella's intuition was to expect variation again and then for the results to even out. In essence, Carmella may have been overgeneralizing the EOP to mean that results even out from the time the run button is pressed until the simulation ends, rather than from no trials to a large number of trials. Jasmine intuitively thought that the graph would stay even but then later nodded in agreement with Carmella's statement. Once the simulation began, Carmella amended her prediction to account for what she saw on her screen. Jasmine first volunteered an explanation and referred to the numbers already being "flattened" at 200 and reasoned that they will stay relatively flat with additional trials. I think Carmella's experience with this last simulation helped her accommodate her concept of the EOP to account for the new experience. When I asked them about adding another 200 to the 400 trials, Carmella said "They probably would raise [pointing to the numbers in the data table], but it would probably keep the same position and it would be the same number apart but it would be higher." Her description of what she expected seems to demonstrate that she was developing both a numerical and graphical concept image of the EOP. However, if she really thinks the numbers will "keep the same position" and "be the same number apart," then her prediction is not actually reflective of the EOP.

During the third teaching session, the girls did experiments in the microworld with a 2B2W bag of marbles as well as a 5B5W bag of marbles. First with the 2B2W bag, and then the 5B5W bag, they did simulations with a small number of trials (e.g., 4, 10, 20) and then a large number of trials (e.g., 500, 1000). With both bags they commented on the wide variability with the small number of trials and the EOP with the large number of trials. They also used the similarity in the EOP with both bags to support a justification of the equivalent probabilities in both bags.

Later in this session, I had them design a 3B1W bag of marbles. Carmella intuitively knew that a large number of trials would not result in an “even” amount of black and white marbles, but rather, should approach a 75%-25% distribution. Thus, her schema of the EOP now was robust enough to account for unequiprobable outcomes. Watching the pie graph during simulations with a large number of trials further enhanced her understanding of the EOP. She was able to visualize the process as tending toward a 1/4-3/4 pie graph representation.

Carmella’s understanding of the EOP carried over into her experimentation with the weight tool and unequiprobable outcomes during teaching sessions 4 and 5. At one point during session #4, Amanda weighted the coins 89 heads to 90 tails. Carmella thought that this would give tails a slightly better chance of occurring but predicted that they would have to do“200 or so” to see the difference. With this comment, she is using her understanding of the role of the number of trials in the EOP. She also used her understanding of the EOP to test and compare weights of two heads to three tails to the previous 89-90 weights. Although she thought the weights would also result in slightly more tails with 2-3 weights, the dynamic visualization of the EOP with both situations helped her realize that the 2-3 weights were not equivalent to the 89-90 weights (i.e., with a large number of trials, the 89-90 weight resulted in an experimental distribution much closer to 50-50 than the 2-3 weights.). Although the EOP suggested to her that the weights were not equivalent, she was not able to reason why they were not equivalent.

But at least her understanding of the EOP was strong enough to suggest the unequivalence.

During the fifth teaching session, I secretly chose three outcomes (baseball, basketball, and soccer ball) and weighted the outcomes 2-2-1, respectively. I gave the girls the task of trying to figure out the weights by doing simulations and using any tool they wanted in the microworld (except the weight tool). Carmella began by running simulations with 10 trials and viewing the data with the stacking columns and data table. After about four sets of 10 trials, she thought the baseball had more weight than the other two because all her trials of 10 resulted in more baseballs than the other outcomes. However, she wanted to increase the number of trials to 100 because "it would give a real test of really how many there is." After two sets of trials with 100 Carmella conjectured the weights were 3-2-1, since baseballs had slightly more (about 8-10) than the basketballs. However with the next set of 100 trials, more basketballs occurred.

> T2: Now look at this and explain something to me. Didn't you say that you thought there were more baseballs, there would be like three of these, and two basketballs and one soccer ball. So what happened?
> C: I'm still keeping that because it's only three higher, the basketballs are only three higher [than the baseballs].
> T2: Okay. Want to try it again?
> C: I'm going to make it [number of trials] higher.
> T2: Oh okay. Now why did you decide to make it higher?
> C: Just to give it a final test.
> T2: A final test. [Carmella runs 500 trials and opens both the pie and bar graph] Now what made you decide to open up the graphs?
> C: Well numbers is one thing. But I just wanted to see really how far apart they were on the graphs to see if one was really extremely, extremely far apart or if there was a really big tie.
> T2: What do you think about this?
> C: I've definitely decided that the soccer ball has one. Because it never gets like more than hardly anything.
> T2: So one for the soccer ball.
> C: But the other kind [baseball and basketball] I'm kind of unsure about.

> [They discuss the data further noting the "closeness" and comparing that to the variability she saw with a smaller number of trials]
> C: These two are pretty close but I'm pretty sure the basketball has more.
> T2: You think running it again might be helpful?
> C: Yeah.
> T2: But before we do it let me ask you one thing. You said in the weighting you think there's one soccer, two baskets, and three baseballs. What do you think the weighting is now when you look at this?
> C: I'm not sure. It's just well before when we were doing the lower numbers it looked pretty efficient, but now it doesn't. They [baseball and basketball] look so close…I have an idea. I think that the soccer balls have one, and the baseballs and the basketballs both have three.
> T2: Now could you tell me why you think that's the case?
> C: Well because the more that we are doing it, it looks more obvious that the little tiny soccer balls aren't really getting much. Every time. See look. They hardly can even count to get to a 100. And these two [baseball and basketball] are up in the 200s. But they are always very close. The only time when they were pretty far apart was when we were only doing 10.

This dialogue and Carmella's experimentation shows how she used her understanding of the EOP to estimate the probabilities from a frequentist perspective. She intuitively knew that running a large number of trials would give her a better estimation of the pre-assigned weights.

This segment is important in establishing Carmella's rich understanding and use of the EOP with both objective (*a priori* and *a postiori*) and subjective approaches to probabilistic reasoning. She also relies on the EOP during the sixth teaching session in her investigation of a two-outcome (sun and rain), two-event experiment. After she constructed the sample space of sun-sun, sun-rain, rain-sun, and rain-rain, she noted that all four possible combinations were equally likely because the individual outcomes were equally likely. When I asked her to prove to me that these four possibilities were equally likely, she said "well the more you do it the more you can tell" and changed the number of trials to 500. She knew to use experimental data from a large number of trials to help her test her equally likely conjecture. Her understanding of the EOP has given her a problem-solving approach with probability tasks and enhanced her probabilistic reasoning while solving these tasks.

Carmella's conception of the EOP, her use of a HES and TWA, and her strong conception of independence contributed to her struggle to understand the effect of the number of trials on the likelihood that the experimental results would be "close" to the theoretical proportion or "exactly" the theoretical proportion. This interesting struggle introduces some dilemmas and moments of perturbation in Carmella's probabilistic reasoning. I will further discuss this struggle between "close and "exact" in the following section.

The Struggle Between "Close" and "Exact"

During Carmella's pre-interview, she believed there is a greater chance of getting exactly "even-even" with 1000 flips of a coin than there is with 10 coin flips. Whether she meant exactly an even distribution (500-500) or relatively close to that distribution is unclear. However, she continued to explore the effect of the number of trials on results that are relatively close to or exactly in the theoretical proportion. During the first teaching session, the girls were running coin toss simulations in the microworld. I emphasized that the pie graphs on the different computers looked very similar when we had just done 500 trials, and that the pie graphs did not seem to be as consistent when we were only tossing a coin 10 times. Carmella justified the inconsistency by noting "they were quite different from each other because there's only 10 of them and there's more of a chance for them to get mixed up." She believed that the small number of trials directly affected the likelihood for a relatively wider variability in results. Since we were discussing the pie graph displays, I do believe she was referring to the chance of results being relatively close to the theoretical proportion. In this early stage of the teaching experiment, she was beginning to think about the relative proportions and how increasing the number of trials made it "easier" to get relatively close to the "expected" proportion.

In a later example, recall Carmella's prediction in the third teaching session that with a 3B1W bag of marbles, she thought it would be easier to get the exact proportion if she only did four trials. This TWA to a probability problem is tightly connected with her intuition about the effect of the number of trials. She feels that the small number of trials equal to the total weight give the best chance for getting exactly the expected proportion. However, her intuition about increasing the chance of getting relatively close to expected with a large number of trials is supported by her developing conception of the EOP.

During the fourth teaching session, Carmella had predicted a 25-75 distribution with 100 trials of coins weighted 1-3. She runs 100 trials and gets her expected results.

> T: Just as she predicted, 25-75…So will it always be 25-75?
> C: Probably not…But close.
> T: Do you think I could get 50-50?
> C: No. Well you could, if you tried all day. But....
> T: But?
> C: It wouldn't be too likely.
> T: Wouldn't be too likely. How about this one? 29 [heads] and 71 [tails].
> C: It's close. It's four apart…Well this is four lower and that's four higher… 71 is four lower than 75…And 29 is four higher than 25.

Although she does not expect to always get a 25-75 distrinbution with 100 trials, she does think it should be close and believes that a 50-50 distribution would be unlikely. But, also of importance in this segment is her spontaneous reference to the deviation from the theoretical proportion. She continues to reference the deviation from the expected distribution in later experiments and also notes the importance of the deviation relative to the number of trials (e.g., with 10 trials, she considers a deviation of two greater than the same deviation with 100 trials). The notion of deviation from the expected distribution is critical in her analysis of relatively close and exact results.

The struggle between “close” and “exact” was brought to the forefront of her analysis during the sixth and seventh teaching sessions. With the 2-outcome (sun and rain) 2-event experiment, Carmella was relying on the EOP to prove to me that the four possibilities (sun-sun, sun-rain, rain-sun, rain-rain) were equally likely by running 500 trials. After about 300 trials she said “see, they are really close.” After the 500 trials were complete, I asked her to predict what would happen if we did another 500 on top of the existing 500. Her first instinct was “they will probably even out” but then rethought her decision and added “or they might spread out.” It seems that in this instance her understanding of independence (i.e., future results can not be predicted based on past results) could be conflicting with her intuitions about large trials resulting in relatively close results. Yet, when I asked her “if they were to spread further apart, what would that mean about the likelihood of them?” she replied “Well most people would think that one of them had more than another, on the weight.” She believed that a discrepancy in a large number of trials could indicate unequal chances (which is aligned with her conception of the EOP). However, the fact that she considers the possibility that after 500 trials the results could “spread out” indicates she was not convinced in the EOP or that a large number of trials result in relatively close results to the theoretical distribution.

During the seventh teaching session, Carmella’s struggle with “close” and “exact” appears while she is experimenting with four colors equally weighted as 10/40. She runs several sets of 100 trials and reports out her findings. Some are close (deviating by 2-3) to her expected distribution of 25-25-25-25, and others are further apart (deviating by up to 10).

> T: So we still haven't gotten all 25, 25, 25, 25. Hmmm.
> C: Maybe it's because we have 100 of them.
> T: Yeah?
> C: So it's easier for them to get mixed up and messed up.
> T: Yeah? Easier to get them mixed up. So what do you think is going to happen if we go to a 1000?
> C: They are going to get more mixed up.
> [She spends some time figuring out how to divide 1000 into quarters.]
> T: So what do you think is going to happen when we go to a 1000?
> C: It might or it might not. It probably won't because the numbers are increasing. It will probably mean it will have more of a chance to scatter.
> T: Oh more of a chance to scatter. So you think, let me go back here.
> C: There are more numbers, they have a bigger number to work with. These little, evil, devil buggies. The elves, they live in your computer and they come they mix and match and they scatter the numbers further apart.
> [She runs a trial of 1000.]
> T: So what's going on here?
> C: They are going and going. But they look kind of close together, which is kind of surprising.
> T: What's going on with that pie?
> C: It looks pretty close. And it's not moving much.
> T: Yeah. Why not?
> C: Oh man. I'm thinking in the wrong direction. The more they do it, the more they get closer together.
> T: They do?
> C: I think. I don't know. I'm getting confused.
> T: You are confused? Hmmm.
> C: Maybe … (pause)
> T: So you've said you were thinking in the wrong direction. That you think the more you do it the closer they get. Is that what's happening?
> C: Yeah, it looks like it.
> T: So is doing it a 1000 times, is that giving the elves room to make them get further scattered apart?
> C: No.
> T: No?
> C: The elves I think are taking a vacation.
> T: They are taking vacation? So you think if we did it again, that if the elves were back, we could do it so that it was scattered?
> C: No, they are going to be on vacation because they don't like all these numbers.

In this episode, Carmella first conjectured that the results would “scatter” more with 1000 trials. She seems to be focused on the exact distribution rather than a relatively close distribution. This may be due to her lengthy work during this episode to use proportional reasoning to predict the exact expected distribution. If she is thinking about getting the exact distribution, then she is certainly correct that there is a lesser chance of this occurring as the number of trials increases. However, it was the dynamic updating of the data table and pie graph during the simulation that prompted her to realize she “was thinking in the wrong direction” and to remember the EOP and the effect of a large number of trials on a relatively close distribution.

As the above episode continued, Carmella discussed the effect of the number of trials on the relative closeness to the expected proportion in the pie graph. She noted that with 10 trials we can’t get the four slices in the pie very close to quarters, but they get closer when we do 100 and even closer when we go to 1000. In addition, when I asked her what she thought would happen if we did 500 on top of the 1000 already done, she knew that the four slices would “keep getting closer the higher you go.”

It seems that her struggle with the issues of “close” and “even” actually contributed to a greater understanding of the EOP and the importance of the number of trials in the relative closeness of the results to the expected proportion. Initially it seemed that her TWA to experimentation may have overemphasized the search for the exact distribution and the emphasis on a small number of trials to obtain the exact distribution. She seemed to often interchangeably think about exact and close results in the same tasks and easily confuse herself about whether a small or large number of trials were needed.

However, I believe her repeated struggles with this issue left her with a rich understanding of the law of large numbers.

Use of Proportional Reasoning

Carmella used elements of additive and multiplicative reasoning when comparing probabilities in the pre-interview. She continued to use both types of reasoning when faced with proportional tasks during the teaching sessions. For example, during the first teaching session, I revisited the marble tasks from the pre-interview. Carmella still only used proportional reasoning with the 1B4W and 2B8W bags and not with the 3B1W and 6B2W bags. It seems that she did not recognize the doubling pattern she had used with the 1B4W and 2B8W bags.

Throughout my description of Carmella's work in the previous three themes, there were many instances where she employed proportional reasoning to predict experimental results that would reflect the theoretical probability (e.g., with a 3-1 weighting, she used her "money knowledge" of quarters to predict a 75-25 distribution for 100 trials). She also used proportional reasoning for making several different weights that reflected equal chances (e.g., 1-1, 2-2, 4-4) and a 2:1 ratio (e.g., 4:2, 8:4, 16:8).

When experimenting with a 3B1W bag of marbles, Carmella and Jasmine numerically and graphically predicted the distribution for 100 trials.

> C: I think we are going to have WAY more black ones.
> T: So if you have to guess how big, with the pie. How much of the pie do you think the black is going to take up?
> J: Probably about....
> C: There will be a quarter white.
> T: A quarter white. And why do you think it will be a quarter white?
> J: and three quarters black.

C: Because if we use each quarter kind of like a marble, there will be one that will be white.
J: It would look kind of like this. [Jasmine drew a picture of a pie graph that has about a ¼ slice shaded (see Figure 5.5). Carmella then colors in the ¾ slice with the pen.]
H: Oh, that part will be black? [pointing to the ¾ area Carmella shaded?]
[Carmella and Jasmine nod in agreement.]

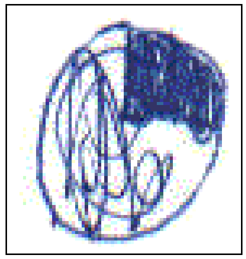


Figure 5.5. Jasmine and Carmella's ¼ - ¾ pie graph prediction.

In this case, they chose to make a graphical prediction than a numerical one. Carmella's reference to ¼ and her reasoning about "each quarter being a marble" facilitated her reasoning with rational numbers.

During the last session, I revisited the 3B1W and 6B2W bags of marbles to see if she would be able to reason proportionally about the chance of choosing a black marble from each bag. She used the marble environment to design the 3B1W bag of marbles and we used the display in the weight tool to discuss the chance of picking a black marble and for picking a white marble (see Figure 5.6). She stated the chances as "3 out of 4" for black and "1 out of 4" for white, and also predicted 75 blacks and 25 whites in 100 trials because she could divide 100 into four groups of 25 with three of the groups represented the three out of four chance for black. Her justification used direct proportional reasoning and demonstrated her ability to use theoretical probability to predict experimental results.

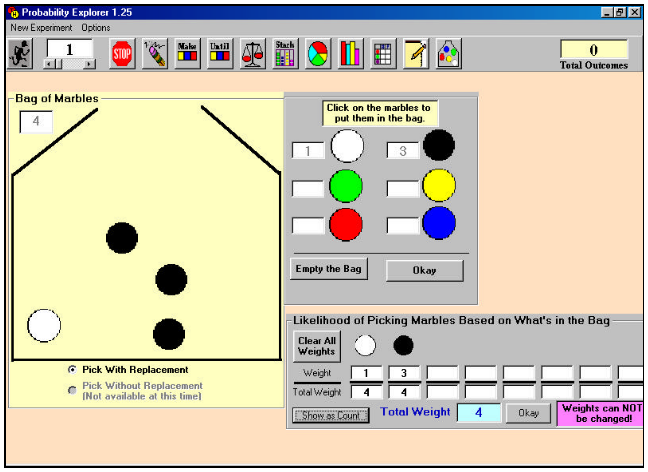


Figure 5.6. Bag of marbles designed with a three out of four chance of picking black.

She ran several sets of 100 trials and went on to predict 375 blacks and 125 whites for 500 trials. She used a "bubble chart" (as she called it) to divide 500 into four equal groups (Figure 5.7). She used this type of "bubble chart" several other times during this teaching session to help her predict distributions proportional to the theoretical probability.

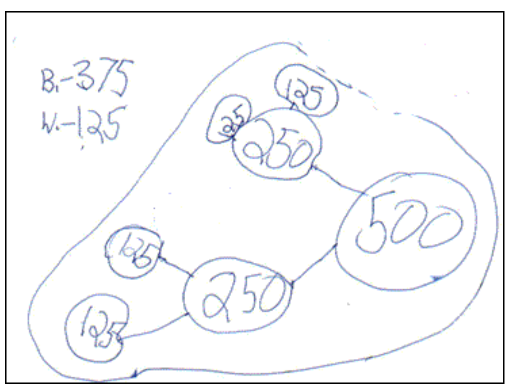


Figure 5.7. Bubble Chart for predicting proportional results with 3B1W marble bag.

When I asked her to design the 6B2W bag, she spontaneously started to compare it with the 3B1W bag. "It pretty much is the same exact thing that I did before. You just have to double with this. Because eight is twice as much as four. And this one [3B1W bag] has three black marbles, and this one [6B2W bag] has six black marbles. And this one has one white marble and this has two whites." Although she was noting the "double" relationship between the bags, when I asked I asked her to predict the results for 100 trials with the 6B2W bag, she did not immediately recognize the equivalent chances, perhaps because 100 is not easily divisible by eight. She was using part-part relationships sand did not account for the part-whole relationships. Carmella then used another bubble chart to break 100 into eight equal pieces of 12 ½, and then combined two of them to get 25 and six to get 75. She then recognized "I did all that work and I don't think I needed to because it was the same as bag 3 [3B1W]. Running several sets of 100 trials further confirmed that she should get results close to 25-75. To predict for 500 trials, she immediately used the previous prediction of 125-375 and noted that "it would still be the same chance of picking a white marble or a black marble."

I believe the multiple representations used in the marble environment, including the weight tool, as well as her bubble chart strategy, facilitated Carmella's proportional reasoning. In addition, the ease in which she could run simulations and view the dynamic multiple representations of the data helped her test and confirm the proportional relationships. As an extension of the previous bag comparisons, I asked Carmella to design another bag of marbles that would have the chance of picking out a white or black marble. She designed a 12B4W bag because "it's twice as much as this one [6B2W bag]." I then cleared the bag, placed three white marbles in the bag, and asked her how many black marbles she would need to put in the bag to make it equivalent to the 3B1W and 6B2W bags. She put in 18 black marbles because in the 3B1W bag "the black is three times as much as the white, so I gave the black three times as much as the white."

When I asked if 18 is three times as much as three, she said “oops”, cleared the bag and put in three white and nine black marbles. Although she intitially made a multiplication error, she correctly used multiplicative reasoning to solve this proportional task.

Overall, her consistency in using appropriate multiplicative reasoning greatly increased as the teaching experiment progressed. She used “money knowledge” and a bubble chart to often help her think through proportional tasks. However, she also seems to be able to think both numerically and graphically in making proportional predictions. This flexibility enhanced her proportional reasoning and allowed her to make numerical and graphical connections. The following description of her work on the “twice as likely” task further illustrates her use of additive and mulytiplicative reasoning as well as connections she makes between numerical graphical displays.

Four Themes Illustrated in One Rich Episode

During the sixth teaching session, Carmella and I worked together on the “twice as likely” task. During this task, Carmella used elements of each of the four themes previously described. In order to illustrate the interactions between the themes, I present the following thick description of her work with this task. I will highlight the occurrence of each theme as well as her use of the microworld tools during this episode.

Carmella had designed an experiment with two outcomes, the sun and rain. With the weight tool open with the default display of equiprobable (see Figure 5.2), I asked Carmella to make the weights so that “the sun is twice as likely to happen as the rain.” She kept the total weight as two and gave the sun two and rain zero, and justified the twice as likely “because there’s two of them and all are in the suns.” She may have interpreted the task to mean that she had to keep a total weight of two, in which case this distribution was her only other option. She may not have been sure how to interpret "twice as likely" in terms of assigning weights to outcomes, or may not understand the multiplicative relationship implied by the term "twice." Although this distribution does not reflect a "twice as likely" relationship, Carmella correctly interpreted the likelihood for sun as "two out of two" and the rain as "none." In addition, she correctly predicted that if we did 100 trials, we would get all suns. After running a trial of 100 and confirming her prediction, I tried to rephrase the task so that she might be able to reason about the multiplicative relationship. The following protocol illustrates her struggle and eventual illumination with obtaining a "twice as likely" relationship.

> T: Now I want you to do it so that the sun is twice as likely but that the rain does have a chance to occur.
> C: Okay. [pause]
> T: What do you think it means to be twice as likely?
> C: Uh oh. [shrugging shoulders]
> T: That's okay.
> C: It means that if you are just, pick a number 10 times, wait a minute, never mind. Let's say …oh this is what I was going to say. The total is eight. [pause] Oh yes, wait.
> T: That's okay. [pause] What are you thinking about?
> C: Well I was going to use eight as a total. But you know if you divided eight in half it would be four. So let's say there are two fours.
> T: Let's go ahead and make those two fours.
> C: Well, um, but it's not going to be even.
> T: So now if this [sun] has four and this [rain] has four, how would you describe those chances?
> C: Oh wait, I have a good idea. You make this … no wait…[then she types in six under the sun and leaves four under the rain] I think… is that right?
> T: Well how about if I explained it this way. No matter how many times we run it I want to have twice as many suns as I do rains.
> C: I have an idea. You probably wouldn't want to do this, but this would be easier for me to comprehend. [she types in 100 under the sun and 50 under the rain]
> T: Now why does that work? Or why do you think that works?
> 100 and then 50?

C: Because well the sun, well let's just say the sun makes 100 and half of 100 is 50.
T: Oh, so that makes it....
C: Well it would be a lot … well you could just do it this way. You could just take off a little zero [she deletes the last zero on each number to make the weights 10 and 5]
T: Oh, so 10 and 5.
C: Would also work.
T: What else would work?
C: 100 and 50 like I just said or 200 and 100.
T: 200 and 100, yep you are right.
C: And also there could be four and two.

Carmella obviously struggled to make sense of this task. She intuitively knew she had to increase the total weight; however, she picked 10 and then eight, two even numbers that can not be a total weight with parts in a 2:1 ratio. She may have chosen even numbers because she interpreted "twice" as divisible by two. Her thinking with the total of eight reflects her beginning understandings of the task. She is aware that she has to break the total weight of eight into two parts and starts with a 1:1 relationship, one that she has worked with extensively. She then changes the weight for the sun to a six so that the ratio is 6:4. At this point, she knew the sun had to have more weight, and she may have chosen six because it is two more than four, using an additive relationship rather than the desired multiplicative one. It appears that when I rephrased the question in terms of experimental outcomes, she was able to think of the task multiplicatively. Notice she first used 100 and 50 and commented "this would be easier for me to comprehend" even though "it would be a lot." These references indicate she may have used her HES and TWA in reverse, directly linking the results in a hypothetical experiment to the weights used to define the likelihood of each outcome. The 100:50 ratio was an easy one for her to recognize as satisfying the “twice as likely” request. Once she established the 2:1 ratio, she easily constructed several other examples that maintained the relationship.

She decided to leave the weights as 10 (sun) and five (rain) for the simulation part of this task. When I asked her to predict what she thought the pie graph would look like her intuition was that the rain would only take up “one quarter” of the pie. It appears that she did not recognize a 5/15 probability as a one-third relationship. In fact, she may have pictorially used part-part reasoning and an additive relationship of three-quarters being two quarter parts more than one-quarter. Although the weight tool was showing the weights as part-whole (10/15 and 5/15), at this point she has not made any numerical or graphical connections between a 2:1 ratio and 2/3 to 1/3 distribution.

I gave Carmella the choice of how many trials she wanted to run. She decided that the best number of trials to run was 15 since “that’s the total weight” and “most likely it would have to be right.” She was clearly using a TWA and emphasized that “right” means to have the exact distribution as designed with the weight tool. Carmella ran the simulation 15 times and did get her exact distribution (10 suns and five rains). When I directed her attention to the pie graph, the visual picture of the data prompted her to make the connection between the 2:1 ratio of weights and the whole being split into thirds with 2/3 given to the sun and 1/3 to the rain.

> C: Ohhhh … now I get it.
> T: What do you get?
> C: Well let’s just say that this line extended all the way over here. [She drew an imaginary line with the mouse pointer from the center of the circle to visually divide the gray slice into two equal pieces. See Figure 5.8 for static image of what she did.]

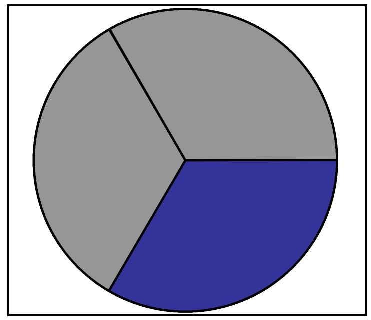


Figure 5.8. Imaginary (black) line that divides the gray area into two pieces.

T: Yeah?
C: So then this [gray slice] would be divided into two. And each of these two are five pieces of pie. And everyone knows five plus five equals 10. So this [blue slice] would be the only part left and so it's 5.
T: Oh. I see. So I want to make sure I understand here. You are imagining drawing the line right up here. [I use the mouse to retrace her imaginary line] So if that line is there, how many slices of pie would there be?
C: There are only three slices of pie.
T: Three slices of pie.
C: But that's great. Because 15 divided into three would be five slices of pie.

In this case, numerical values in a 2:1 ratio were not enough for Carmella to transition from part-part to part-whole reasoning to understand the 2/3 to 1/3 distribution. The graphical picture of the data was a more powerful tool in helping her conceptualize this relationship. It appears that the pie graph was enough of a cognitive prompt to connect with her schema of thirds as a circle divided into three equal pieces. She was able to further connect the numerical values of 10 and five by splitting 10 in half (just as she did with the gray area representing 10) and recognizing that five is one-third of the total number of trials (and weight).

We continued running simulations with the 10:5 weights and the number of trials as 15. The next simulation resulted in eight sun and seven rain, which Carmella recognized as "not very likely, but at least the sun had more so it's kind of still pushing the point." The next run of 15 resulted in 13 sun and two rain. I reminded her that after running the simulation three times, we got very different results each time (see Figure 5.9).

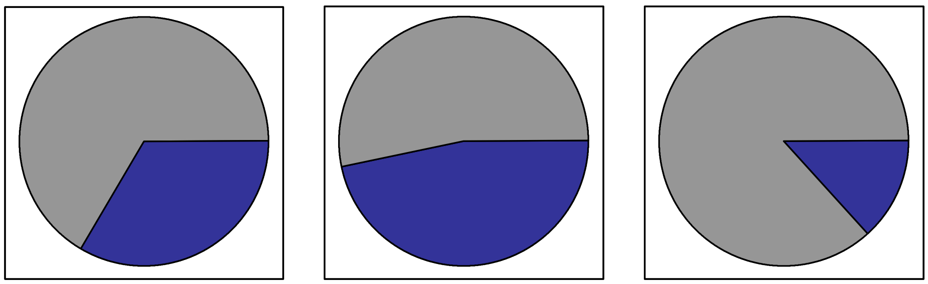


Figure 5.9. Graphs of three consecutive trials of 15 (from left to right: 10-5, 8-7, 13-2).

She summarized the visual picture of the data by saying "one of them was almost all the way half way (Figure 5.9, middle graph). This one (Figure 5.9, left graph), the one that we thought was supposed to happen, was like was one-third right here [pointing to the approximate location of the 1/3 slice on the pie graph]. And now this (Figure 5.9, right graph) is like a third of a third, which is like …I don't know." Two things are important with her summary. First, she has a visual sense of the relationship between a fraction and a circle area representation. Second, she again uses language to indicate she believes the weight represents what is "supposed to happen" experimentally. It is not clear if her notion of "supposed to happen" is based on her TWA or an indication of her reliance on the EOP.

After her summary, I asked her if she thought the pie graph of a trial of 100 would look like any of the last three graphs or if it would be something different. Her intuition was that the graph would be something different "because it's a 100, and that's way off of the numbers that we use for the total weight. It's usually more exact when you are using the total weight." To her, using trials not equal to the total weight makes it more difficult to get the exact proportional distribution. In fact, I don't think she was using proportional reasoning at all here and any notions she has previously expressed about the EOP and tendency toward the weights (theoretical probability) with a large number of trials was not evidenced in this response. Her fixation on the trials equaling the total weights (TWA) was dominating her thinking.

However, once again, the use of the graph as a cognitive problem-solving tool helped her think about the task. We cleared all previous trials and changed the number of trials to 100. Before we she pressed the run button, I asked her to leave the pie graph open so she could watch it during the simulation. The following protocol provides evidence that she was able to use her past experiences with the dynamic display of the pie graph during experimentation to think about the task in a different way.

> C: Should I shut this? [points to the pie graph]
> T: No, let's watch it. So what do you think it [pie graph] might look like?
> C: Well I think when we first start out the sun will have maybe almost all of it. And then the rain will start to get some more.
> T: About where do you think the rain is going to end up? Do you think it will be small like it was that last time? Close to even?
> C: I think it's going to be like one-third of the whole pie.

Her visualization of the dynamic pie graph gave her a powerful tool to aid her reasoning. She used the visualization of the EOP to connect with her intuition about the law of large numbers. This provides strong evidence that Carmella can solve complex probability tasks and that she relied on the dynamic multiple representations in the microworld to stimulate her mathematical thinking.

This segment continued with Carmella running trials of 100 several times and analyzing the graph and table of data to look for evidence of a "twice as much" relationship. She reasoned both numerically and pictorially and eventually reasoned that 67 suns and 33 rains was the closest distribution since "33, double that would be 66 and that's [the sun] 67." She even suggested that she wished she could throw away one of the sun icons to make her data reflect 66 and 33. (Note: As of yet, the microworld does not have a "trash can," but this feature is planned for future development.)

Overall, this episode illustrates Carmella's use of the available data representations in the microworld and how these representations aided her probabilistic thinking. She certainly used a TWA in her experimentation; however, her use of TWA initially hindered her reasoning about the results of large number of trials because "it would be way off the total weight." Thus, she was thinking about getting the exact distribution rather than a relatively close one. Her visualization of the EOP became a valuable problem-solving tool for overcoming the TWA and expecting results proportionally close to the weights of 10-5. Although she demonstrated strong proportional reasoning on previous tasks, she struggled with this particular task and needed more than numerical representations to make connections between a 2:1 ratio, the concept of thirds, and how this ratio affects the distribution of random results in both small and large samples. The dynamic pie graph certainly facilitated her meaning-making activity in this task.

Summary of Meaning-Making Activity in the Microworld

Carmella was engaged in a high level of meaning-making activity during the teaching sessions. She quickly learned how to use the various tools in the microworld (e.g., stacking columns, graphs, data table, weight tool) to design and run experiments and analyze the results in a variety of iconic, graphic, and numeric formats. She often used the multiple representations simultaneously in her analysis and recognized that she could get different information from the various representations (e.g., “numbers is one thing, but I just wanted to see really how far apart they were on the graphs”). She also discovered that although she could see numerical differences in the data table and bar graph, the pie graph would show her the relative differences (e.g., “although the numbers are increasing, the percentages in the pie are not”).

The four themes – TWA, EOP, exact vs. close, proportional reasoning—capture the essence of Carmella’s meaning-making activity, and suggest that the tools in the microworld facilitated her mathematical thinking and further development of probabilistic reasoning. Together, the four themes were interconnected and led Carmella to a fascinating realization about the effect of the number of trials on the probability of an event occurring. She realized that the chance of getting the exact theoretical distribution was largest when the number of trials was small but that the chance of getting relatively close to the theoretical distribution increased as the number of trials increased. This level of probabilistic reasoning is advanced for a 9-year-old student. In fact, the flexibility in her thinking demonstrates that, depending on the situation, a smaller number of trials can give you a better chance of getting particular results (e.g., if a couple wants to have exactly the same number of boys and girls in their family, assuming equiprobability, there is a 2/4 chance of one boy and one girl if they only have two children; however, if they have four children, the probability of two boys and two girls decreases to 6/16.).

Carmella’s meaning-making activity demonstrated that she was easily able to transition to the digital environment in the microworld and use the tools to her advantage.

She was empowered by the open-ended nature of the microworld and was able to easily use the tools to test her conjectures. In addition, there is evidence to suggest that several of the tools were critical prompts in her development of sophisticated probabilistic reasoning.

Post-Interview Analysis

The post-interview was held two weeks after Carmella's last individual session. The interview protocol (Appendix Z) contained several items exactly the same as or very similar to those in the pre-interview. In addition, several new tasks were used to assess understanding on concepts that had emerged during the teaching experiment. Overall, Carmella maintained an advanced level of thinking that was evidenced throughout the teaching experiment and her pre-interview. As I did for the pre-interview, I will recap her answers and strategies with the tasks used in this interview.

Cubes in a Bucket

As in the first interview, I asked Carmella a series of questions using a bucket containing six green, four red, and two yellow cubes. When I asked which color had the best (and least) chance of being picked, she used qualitative replies of "green because there's more" and yellow because they were the least. In a situation where the parts are easily recognized and countable, she did not need to use part-whole reasoning to find the most and least likely event. We sampled from the bucket and picked a yellow cube. I asked her "if I put this yellow back in, do you think I'm now more likely to get a yellow?" She recognized the importance of the replacement of the cube and said that the chance of picking a yellow did not change because I "put it back in." She added that I would be less likely to choose a yellow if I did not put it back in the bucket because there would be one less yellow. Not only did she recognize the value of a replacement or without replacement situation, but she clearly considered each pick from the bucket as independent from the previous result.

The task continued as I removed two green cubes, one-by-one. After the removal of the first cube, she stated that the green was still the most likely to be picked "but not by very much because the red has four, the yellow has two, and the green has five. But when this one was in [pointing to the green cube outside the bucket] that was six. It had two, now it only has one." The reference to two and one in her last comment referred to the difference between the number of green and red cubes. She used a part-part comparison to conclude that since green only had one more than red, then the chance of picking a green cube was not much more than the chance of picking a red cube. She used similar part-part reasoning to analyze the chances after I removed the second green cube. Red and green had the same chance and yellow still had a lesser chance of being picked.

I brought the bucket back out later in the interview and asked Carmella if she could use numbers to describe the chance of picking out each of the colors. She was easily able to state the chances for each color in a part-whole relationship (e.g., "6 out of 12 chance for green"). I then removed two of the green cubes and she again was able to use part-whole reasoning to state the chances of picking each color. I then asked her to compare the chance of getting a green cube now (four green, four red, two yellow) to before (6 green, 4 red, two yellow). For each of the colors, she uses rational number reasoning to compare the numerical probabilities. The following protocol illustrates our dialogue.

> T: So did the chance of picking a green cube change? Since before they were in here and when they are out?
> C:Yes.
> T: How has it changed?
> C: It changed because now it has a less likely chance…Wait a minute. This is confusing. Even though this one has, this is weird. Well this one still, I know it has a less likely chance, but they have less cubes in there all together.
> [To help guide her thinking, I asked her to state the chance of picking a green before (6 out of 12) and now (4 out of 10).]
> T: So is that the same chance as before? Or is it a different chance?
> C: It's a less likely chance.
> T: And why is it a less likely chance?
> C: Because six is half of 12. And five is half of 10. We now have four greens, so it's less than five.
> T: Oh, so what does that do about the chances of picking a green?
> C: It lowers it.

One half seems to be an easy reference point for Carmella to use in proportional reasoning. Based on five being half of 10, she reasons that four out of 10 must be less than half and concludes that the chance of picking a green cube is lowered when the ratio goes from 6/12 to 4/10. Her use of a reference point carries through in her reasoning about the chances of picking a red cube.

> T: Now what about the chance of picking out a red?
> C: Before it had a four out of 12 chance. And now it has a four out of 10 chance. Okay. This chance is actually better now.
> T: It is? And why is it better?
> C: Because let's think with our number checks. And if you take two away, that's four. And then on number five if you take one away, that's four. But you would have to take more away from six to get four than five. So it's chance is actually now better.

In this instance she uses a one-half relationship to compare 4/12 to 6/12 and 4/10 to 5/10. She reasoned that since four out of 12 is two less than six out of 12 and four out of 10 is one less than five out of 10. Thus the ratio that is closer to half must be the higher chance. She was clearly only considering the part-part relationships and did not account for the magnitude of the whole. Although this strategy worked for this example, it would not work in all cases (e.g., 48/100 and 4/10). However, for a similar problem, she uses a valid strategy when discussing the chance of picking a yellow cube.

> T: What about the chance of picking the yellow?
> C: It also is better now. Because it's easier to see, because since 12 is less, well let's think about a pie graph. With 10, each piece would have to be bigger than with 12. So it was like the yellow would have bigger pieces of pie when there's 12. I mean when it was 10. Because there's less, because like when the pie graph it still has the same amount of room that it has to cover. It just has to divide them.
> T: Oh so when we have two out of 10, that's going to take up more room on the pie graph?
> C: Well it will have to take up more room than on the two out of 12.

Here she relies on rational number reasoning and a mental picture of dividing a pie into equal pieces. She used this strategy with the pie graph several times during the teaching sessions as well.

Carmella's reasoning about the changes in theoretical probability demonstrates her understanding of the effects of a part in relationship to a decreased whole and her ability to use a common fraction (1/2) as a reference point for comparing two fractions. Although her reference point strategy for comparing 4/12 and 4/10 was appropriate, her reasoning about the differences between parts (6 - 4 and 5 – 4) could not be correctly applied to every fraction comparison problem (e.g., 48/100 and 19/40) without consideration of the magnitude of the denominators. It is unclear if she would try to use such a strategy for comparing fractions whose denominators where not numerically similar.

Coin Tosses

Two types of tasks involving coin tosses were used in the post-interview. The first task was exactly the same as the second coin task in the pre-interview. The additional coin tossing task involved using experimental data to determine if a coin is fair. The first task assessed concepts of independence while the second task assessed concepts of fairness and the law of large numbers from a frequentist perspective.

Tossing a coin six times. Upon showing her the four possible strings of results from flipping a coin six times (HHHHTT, THHTHT, THTTTH, HTHTHT), Carmella said that none of the strings of results were more likely to happen than the other strings. Her reasoning included references to the number of permutations of heads and tails for six results and the unlikelihood of getting a specific sequence in any given trial of six.

> C: There would be a lot of different ones that you could do. So it would a very unlikely chance that you would get just one of these [strings] by itself. Because you could do like all heads or all tails, or like this one at the end [points to the string THTHTH]. So there's tons of different ones you could do. So it's actually very unlikely that you would get one in particular.

Carmella has an advanced understanding of independence. That understanding, coupled with an appreciation for the number of permutations possible in a string of six results, allowed her to reason qualitatively that none of the strings would be more likely than any of the others.

She continued a similar line of reasoning when shown the next set of results (HHHTTT, HHHHHH, THTHTH, HTHTHT). In addition, she explicitly considered a particular permutation and tried to quantify the chance of getting that result.

> C: It seemed that this one would be least likely but it's not [points to the result HHHHHH]
> T: The all heads? Okay, it seems like that would be least likely but you say it's not. Why is it not least likely?
> C: Because there's just as many flips as the other ones. And it just has the same chance as all the rest.
> T: Oh, so when I flip this coin, how would you describe the chance of getting all heads?
> C: It would be like a one out of something or other chance … it would be a pretty big number.

Carmella's level of thinking goes beyond an intuitive awareness of independence for a single coin flip that she displayed in the pre-interview. Now, she treated each possible string of results as a single possibility out of the many possible strings. Although she does not know each permutation or the number of permutations, she has a sense of the magnitude of that number and can state a theoretical probability in her own informal terms (e.g., "one out of something or other chance").

For the final question about independence, I asked Carmella if, after flipping a coin and getting the results HTHHHH, I was more likely to get a heads or tails on the next flip. She promptly stated "they are both equally likely." In the pre-interview she also claimed that the results were equally likely. Thus, her concept of independence was intuitive at the beginning of the teaching experiment and continued to strengthen throughout the research period.

<u>Is this coin fair?</u> The intent of this task was to assess whether Carmella could reason from a frequentist perspective about the fairness (i.e., equiprobability) of a coin

when given experimental results. In the first situation, I told her that I flipped a coin 10 times and got eight heads and two tails. She said the coin was fair but “it’s just that’s how many you got,” implying that these results were not significant evidence about the fairness of the coin. In the second situation, I told her that I flipped a coin 100 times and got 41 heads and 59 tails. She again said that the coin was fair because the numbers were “pretty close.” With the third situation (flipping a coin 500 times and getting 175 tails and 325 heads) she still thought the coin could be fair although the result was unlikely

> C: It’s fair because it’s a pretty unlikely chance of happening, but it did.
> T: What’s unlikely that it happened?
> C: That this one gets 175 and this one gets 325.
> T: And why is that unlikely?
> C: Because it’s just that this and this, so this one has a lot [points to 325 heads].
> T: And can these results tell you anything about the fairness of the coin?
> C: No.
> T: No. Okay. What would you have to do to determine if a coin was fair?
> C: The coin isn’t fair.
> T: It isn’t?
> C: Well it can’t be fair or unfair.
> T: Oh, okay. What about the coins in the *Probability Explorer*?
> C: Well you can change their fairness.
> T: So if I had gone in there and maybe I changed it or maybe I didn’t what would you have to do to figure out whether or not the coin was fair?
> C: Do it a bunch of times and compare the results. And if one of them kept on getting a lot like this, then whatever it was it was probably ranked higher than the other one.
> T: Oh. And by ranked higher, what do you mean?
> C: When you go into that little weight thingy.

Although she recognized the unlikeliness of getting 125 tails and 375 heads with a fair coin, it seems that she had trouble conceptualizing the possibility of a real coin being unfair. This difficulty is entirely appropriate since most of her experimental experiences with real coins were with relatively fair coins with approximately 50%-50% probabilities. However, when I focused her on the possibility of changing the probabilities in the microworld environment, she immediately cited an appropriate strategy for determining whether the digital coin was fair. She used a frequentist approach by stating she would do repeated trials and see which outcome "kept getting a lot." Her reference to the weight tool and "ranked higher" indicates that she envisioned the outcome which occurred more often to have a higher probability (i.e., more weight) than the other outcome. I conjecture that her difficulty in applying a purely frequentist approach to the results with a real coin is related to her real experiences with coins. I do not think her response represents a lack of understanding of the law of large numbers and the relationship between experimental results of a large number of trials and theoretical probability. In fact, I think her response indicates a strong appreciation for this relationship.

Marbles in a Bag

As in the pre-interview, Carmella was presented with four pairs of pictures of bags containing black and clear marbles. The bags used in the post-interview were different from those in the pre-interview and were labeled with letters rather than numbers (e.g., Bag A rather than Bag #1) to avoid association with the bags used in the pre-interview. Two of the pairs of bags were in proportion to each other, while two pairs of bags were not proportional. When presented with each pair, Carmella was asked to determine which bag she would prefer to pick from, or if it mattered which bag, if she wanted to choose a black marble.

When presented with Bag A (3B3C) and Bag B (1B1C), she said that in both bags the chance of picking a black or a clear marble was "even-even" and it wouldn't matter which bag you chose from and that three out of six was the same as one out of two because "it is kinda like a reduced fraction." Again, she used a part-whole strategy for comparing the chances and her understanding of a "reduced fraction" to determine the proportionality and equal chances in each bag.

For Bag C (3B1C) and Bag D (5B2C), she used a proportional reasoning strategy that requires adding a marble to bag D to make it proportional to bag C. The following dialogue illustrates her strategy.

> C: This one has seven [bag D] and this one has four [bag C]. So it can't be the same trick as before. This one has five and that one has three [referring to the number of black marbles in bag D and bag C].
> T: So what are you thinking about?
> C: Well I think I found a way to answer this, but I'm not sure if it will even work … But I'm trying to think of it as this being eight [bag D]. But I have to figure out what color that marble will be.
> T: Oh, if you were to put an extra marble in this one?
> C: Uh huh. And this one has a three out of four chance. I'm trying to figure out what it would be. I think if you added one, it would have to be a black one.
> T: Okay. And why would it have to be a black one?
> C: Because it's just like you double them. So we already have two whites. But we only have five of the blacks.
> T: Oh and you wanted to double the 3.
> C: And that would be six. So I think that you would have a better chance to get a black one in this bag [points to bag C].
> T: In this bag. And so you said that the chance of getting a black one over here was three out of four? What's the chance of getting a black one over here?
> C: Five out of seven.
> T: And you want Bag C? [She nods "yes."]

Her strategy for comparing the bags demonstrates her proportional thinking. She reasoned that since bag D did not have enough black marbles to be proportional to bag C, bag C must have a better chance for picking a black marble.

The next two bags presented, bag E (2B1C) and bag F (4B2C), were in proportion to each other. Carmella immediately recognized this proportionality and said "it doesn't matter because you double the two and you double the one (in bag E), two and four (in bag F)." She continued by saying "it doesn't really matter because even though you have more blacks you have more whites … so it [the chance of picking out a black] doesn't really improve." Her reasoning about both parts increasing is a key point in her strategy. Although she correctly used the proportional strategy in this instance, she overgeneralized this strategy with this next pair of bags.

For bag G (2B3C) and bag H (5B6C), Carmella abandoned her "double" proportional strategy that she used in the past for a part-part strategy. After looking at the bags for a moment and writing down the contents of each bag on a piece of paper, she decided that both bags have the same chance for picking out a black marble.

> C: Well I think it's the same chance. Because if you notice in Bag G there's three whites and Bag H there's six whites. I know that it's double, but that doesn't have anything to do with it.
> T: Okay.
> C: And Bag G has two blacks. Which is one less than three. And in Bag H, the black marbles, there's five black marbles. Which you also notice is one less than the white.
> T: Yeah, I agree with that. It is one less than six.
> C: So I think it [the chance of picking a black] is going to be the same amount.
> T: So it doesn't matter which bag you pick from?
> C: I don't think so.

The proportional reasoning strategy that she used with the prior three pairs was not robust enough to help her with this pair of bags. The common difference of one between the number of black and clear marbles in each bag was strong enough for her to abandon her "double" strategy for an additive approach. Although she clearly had the ability to reason proportionally, it seems that, to her, both multiplicative and additive reasoning were valid ways to maintain ratios in proportion.

Constructing Sample Space and Theoretical Probability

All students were able to construct a sample space for one-stage and two-stage experiments during the pre-interview and throughout the teaching episodes. Therefore, the only sample space question used in the post-interview was for a three-stage experiment. The context for this task was a family with three children, ages 9, 5, and 3. Carmella was asked to list all possible arrangements of boys and girls with respect to their ages.

Carmella used a very organized recording system (see Figure 5.10) to help her find all possible arrangements. As you can see from Figure 5.10, she made three columns representing each kid in order from oldest to youngest. She first listed all girls then used a permutation strategy for one boy and two girls. For her fifth possibility, she reversed each choice in possibility #4 (GGB) to get (BBG). She then wrote GBB and BBB and announced "here's the hard part, we have to make sure we haven't done one over." In her checking process, she said "I think I found one I haven't done" and wrote BGB. She was able to use a reasonable reversal strategy to obtain the permutations with two boys and one girl as well as one boy and two girls. Although she lost track in using her strategy the second time, she was able to find her omission without prompting from me.

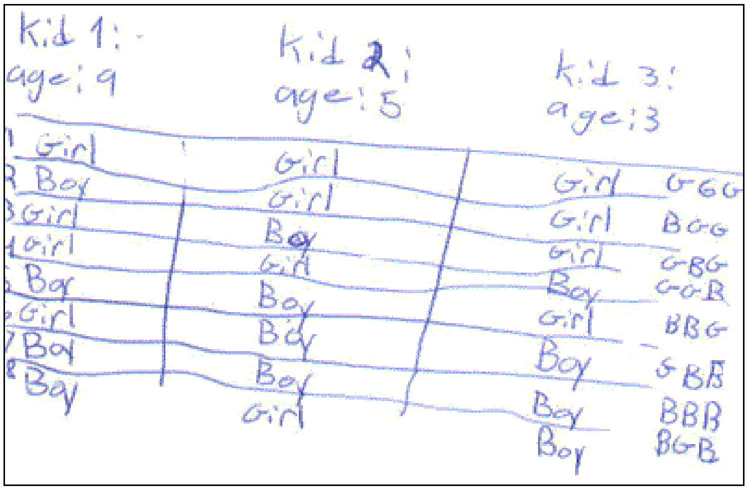


Figure 5.10. Carmella's construction of the sample space for the family task.

When I asked her to convince me she had found all the possibilities, she condensed her chart to a column of data using only G's and B's as shorthand (i.e., GGG, BGG, GBG, GGB, etc.). After studying her list and walking me through all her permutations she claimed: "I think I got them all… and I think I found I can do it using math to find out how many you have." Her reasoning was that "there are three kids and it could be either a boy or girl… so it could be three times two and that's six, but I don't think that's counting the BBB and GGG, so I think I have them all." Although her reasoning is incomplete, it shows that she believes there may be a pattern to how many choices there are and that there could be a formula using the number of stages (3) and number of outcomes (2) for determining the number of possible permutations. Her eagerness to explain her reasoning with a mathematical pattern shows her attempt at combinatoric thinking.

To assess her ability to determine theoretical probabilities from a sample space, I asked her a series of questions regarding the chance of the actual family arrangement being certain arrangements or combinations of boys and girls. Consider the following exchange as evidence of her systematic use of the elements in the sample space to determine theoretical probability of both ordered and unordered events.

> T: What is the chance that the correct arrangement is boy, girl, boy?
> C: A one out of eight.
> T: A one out of eight. And why is it one out of eight?
> C: Because it's only one chance out of eight. [She does a sweeping point over all eight choices]
> T: Oh, because there's eight possibilities there?
> C: Yep.
> T: Now what is the chance that the family has two boys and one girl?
> C: A one out of eight chance.
> T: Okay, and why is it one out of eight?
> C: Because there's eight different combinations that they could have had. And they ended up having that one [points to BBG].
> T: Now when I say two boys and a girl. It doesn't matter what age they are.
> C: Oh…[pause]
> T: So does that change the situation?
> C: Yes. So this one [points to BBG] …
> T: Two boys and a girl.
> C: This one [BGB], this one [GBB], and that's it. It's a three out of eight chance. [She puts a check beside each arrangement as she finds it]
> T: Three out of eight. Okay. You are right. What's the chance that it's two girls and one boy?
> C: So it could be this one [BGG], or this one [GBG], and this one [GGB]. Let's see, I could have done that in my head. [she again used to check marks to indicate her choices]
> T: You could have?
> C: Yes, because I know that those two combinations, it couldn't have been this one [GGG] or this one [BBB].
> T: Okay. So now what's the chance that the family has all the same gender of kids? All the same sex?
> C: Two out of eight.

Carmella used appropriate reasoning with the elements in the sample space when asked questions regarding the chance of a particular ordered event as well as an unordered event occurring. She knew that the eight possible arrangements constituted the "whole" and that the number of possibilities that fit the description of the desired event constituted the "part." Her construction of the sample space in this three-stage situation and her reasoning in determining the theoretical probability demonstrate her strong understanding of the importance of sample space and its direct relationship with determining theoretical probability.

Using Results to Design Experiment

Since the students did a lot of graphical interpretation of experimental results during the teaching sessions, I added a task to the post-interview to assess their ability to interpret and use information from both a pie graph and bar graph. I told Carmella that I had designed a bag of marbles in the microworld and ran an experiment. I showed her a graph of experimental results and asked what she could tell me about the bag of marbles.

Reasoning from a pie graph. For the first task, I showed Carmella the pie graph in Figure 5.11 (left-hand picture) and asked her if she could tell me how many times I ran the experiment. She said no because the pie graph "doesn't say." When I asked her what she could tell me about the bag of marbles, she said it has "yellow, green, and red, and red had more and the yellow and green had the same amount. Maybe not in the bag but that's how it turned out." She then drew the picture of the bag on the right in Figure 5.11.

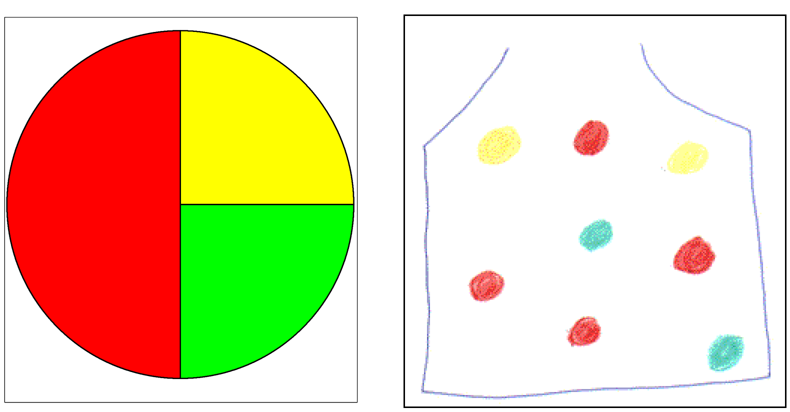


Figure 5.11. Given pie graph and Carmella's drawing of a possible bag of marbles.

Carmella reasoned that it did not really matter how many total marbles were in the bag, what mattered to her was "how many you have of each [color]." Her justification shows how she used the proportions in the graph to determine how many of each color she put in the bag.

> C: Well, I looked at the graph. And I knew that the red had twice as many as the yellow and the green. So just to make it look so it wasn't completely empty, I just like doubled it, to what it could have been. So I put two yellows because that supposedly is a quarter of it, I think.
> T: All right. Well how many marbles do you have in here total?
> C: I don't know. I haven't counted. [she counts the marbles] I have eight.
> T: Eight. And two of them are yellow.
> C: And two of them are green. And they have the same amount. And that equals four. That's why I have four of the red ones because that has twice as many as these.

She was very conscientious about keeping the colors in proportion to the results in the pie graph. Her expression of the total number of marbles not mattering as much as the amount of each color further illustrates her emphasis on the relationship between the parts and recognition that this relationship could hold true with many different bags of marbles with different totals.

<u>Reasoning from a bar graph.</u> The same questions were posed when I showed Carmella the bar graph on the right-hand side in Figure 5.12. At first she could not tell me how many times I ran the experiment but could tell me that I got "400 greens, 100 blues, and 500 yellows." Although later in the task she envisioned stacking the blue bar on top of the green bar to make 500 and added 500+500 to say that I picked marbles from the bag 1000 times. To justify her drawing of the bag of marbles (left-hand side of Figure 5.12), she again used her "take off the zeros" strategy to make four green, one blue, and five yellow marbles in her bag.

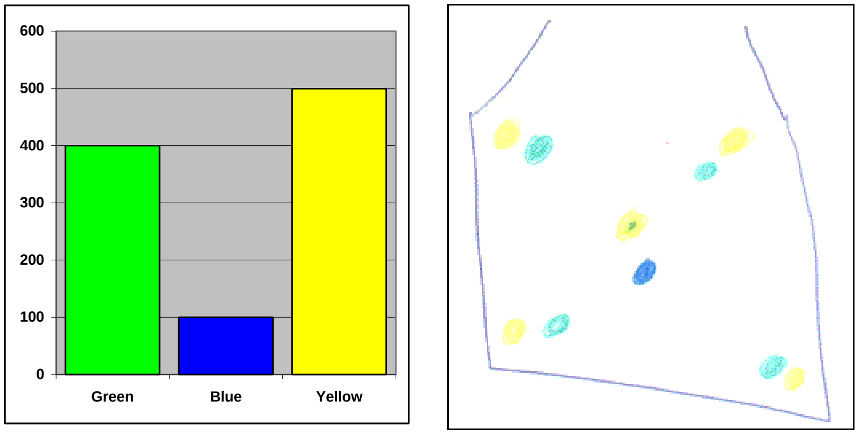


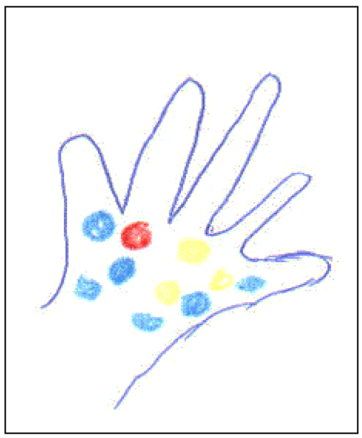
Figure 5.12. Given bar graph and Carmella's drawing of a possible bag of marbles.

Although she did not immediately realize that the total number of trials could be found by adding the individual results from each color, she eventually was able to use the individual parts to find the total. Her recognition of the "reducing" by two zeros for each color gave her a distribution of marbles that was in proportion to the results shown. There is not evidence to suggest that she actually knew that she had divided each result by 100. Nevertheless, she was consistent with her strategy using naïve proportional reasoning and did construct a reasonable possible bag of marbles based on the experimental results.

Carmella was able to use both numerical data from the bar graph as well as geometrical proportions from the pie graph in designing her bags of marbles. These two tasks demonstrate her ability to reason from a frequentist perspective and to estimate theoretical probability distributions based on experimental results.

<u>100 Gumballs</u>

The gumball task used in the post-interview was similar to the one used in the pre-interview. The gumball machine contains 100 gumballs, 30 yellow, 60 blue, and 10 red. I asked Carmella to predict how many of each color someone would get if they chose 10 gumballs from the machine. She drew a picture of a hand and then used the colored pencils to draw 10 gumballs of different colors (Figure 5.13). I asked her to explain why she drew three yellow, six blue, and one red gumball in her hand. She said, "I reduced the 100 down to 10. And so if that happened, then those 10 would have three yellows, six blues, and one red."



<u>Figure 5.13</u>. Carmella's prediction for 10 gumballs.

Although she used the language "reduced," it is not clear that she actually used a mathematical reduction using division by 10, but rather just saw the pattern of removing the zero off the end of each number. Recall her use of this strategy in the pre-interview with the gumball machine task and in the post-interview task of designing a bag of marbles given the experimental results of 1000 trials. Nevertheless, her strategy reflects naive proportional reasoning and gives her a valid and justifiable prediction.

Spinner Game

The last task in the post-interview was similar to the spinner game used in the pre-interview; however, different spinners were used. When shown spinner 1 (see Figure 5.14), Carmella chose the red sector because "its bigger" but added that the game was not fair because "you have a quarter and I have the other 75." When I showed her spinner 2 (see Figure 5.14), she said "it doesn't make a difference. If you mixed all these [sectors in spinner 2] together and made this one [points to a blue sector] pop ever here [next to the other blue sector] it would be the same [as spinner 1]." She went on to explain that each of the sectors in spinner 2 were "about 11 because they are each half of 25."

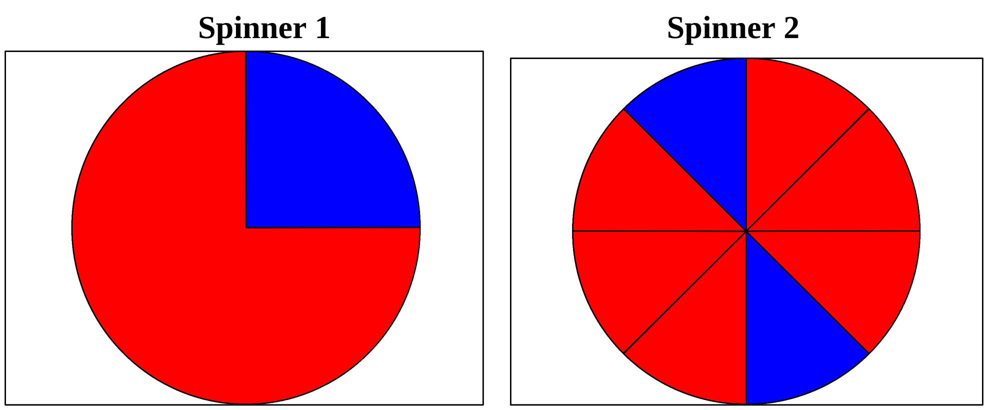


Figure 5.14. Spinners used in post-interview.

Carmella used appropriate proportional reasoning and immediately recognized the equivalence in areas in spinners 1 and 2. Her idea about each sector in spinner 2 being worth half of 25 shows an understanding of geometric and numeric relationships. Although she did not use the language of "percent," her use of 25 and 75 shows her emphasis on 100 representing the whole pie.

When I asked Carmella if there was a way for her to make this game fair, she came up with several options for changing the spinner. All her options kept the ratio of blue and red sectors equal.

> C: Then we need to have a spinner where half of it is one color and the other half of it is another …Or you could have two quarters that were one color and another two quarters that were one color …Or you could have four little slices of one color and four of another one.

Her suggestions indicate that she definitely associates fair with equal chances with corresponding equal areas in a spinner. She also has a strong sense of equivalent ratios that she uses in this task.

Strengths and Weaknesses in Post-Interview

Carmella displayed a strong understanding of basic vocabulary and *a priori* probability. Her understanding of “fair” is strongly associated with equal chances in both discrete and area models. She uses both part-part and part-whole reasoning when describing chances and comparing probabilities. She uses part-whole in describing the chance of an individual event; however she uses both part-part and part-whole in comparing the probabilities of two events.

In some of the comparison tasks, she uses part-whole relationships in combination with a few other strategies. One strategy is to use ½ as a comparison benchmark. However, she sometimes only reasons with the parts in comparison to the ½ benchmark rather than considering the whole as well. Another strategy is to use a “doubling” of both the part and whole of the first probability and to compare that result with the second probability. Again, this is a type of benchmark strategy. The third strategy she used entails an equal additive process (recall the 2B3W and 5B6W comparison) that maintains an equal difference between parts and whole. These strategies suggest that her proportional and rational number reasoning is becoming more sophisticated, but that she still inappropriately applies additive reasoning in certain situations.

Carmella was able to easily reason proportionally from both graphs of results in order to make a reasonable bag of marbles. Although she recognized that the total number of marbles in the bag did not matter, she recognized the relationships between the three colors in the results and was able to create a bag of marbles that directly mirrored the results. In this regard, she used a slightly modified version of her TWA to reason backward from a frequentist perspective.

Carmella's understanding of *a postiori* probabilities was evidenced in the second coin toss task. With a small number of coin tosses, she did not view wide variability in results as evidence about the fairness of the coin. Yet, with wide variability in results from 500 tosses of a real coin, she could not fathom that these results could tell anything about the real coin. When the task was phrased in terms of the microworld, she easily reasoned from a frequentist point of view. Altogether, though, she appropriately reasoned about the effect of the number of trials.

Carmella's understanding of independence was clearly evidenced in the first coin toss task. She also recognized the large number of possible permutations of the six coin toss results and noted that the likelihood of getting an exact string of six was low and that all possible strings were equally likely. This carried over into her work on the 3-event family task. She used a "reversal" strategy to list all possible arrangements and easily reasoned about the chance of an exact arrangement occurring by using the number of possible arrangements as the “whole.” She also used the sample space she created to state the probabilities for several unordered combinations occurring.

Overall, Carmella used probabilistic reasoning in approaching the post-interview tasks and displayed evidence of an advanced understanding of many concepts (e.g., sample space, independence, theoretical probability, experimental probability, compound events). Several times during the interview, she made reference to her experiences and tools within the microworld. In this regard, it appears that she was able to use these as part of her repertoire of problem-solving tools for the interview tasks.

Putting It All Together: Carmella’s Development of Probabilistic Reasoning

As evidenced from her pre-interview responses, Carmella had strong intuitions and developing conceptions of several probabilistic ideas prior to the teaching experiment. She used those intuitions and understandings in her social and digital interactions during the teaching sessions. During the teaching experiment, she further developed her understandings and used her intuitions to conjecture and test conjectures within the microworld. Her social and computer interactions illustrate her use of the microworld tools as cognitive prompts and elements of perturbation during problem solving (e.g., her use of pie graph during the “twice as likely” task).

Carmella’s initial understanding of “fair” was embedded in an assumption of equal chances –which she explained through a hypothetical experiment strategy (HES). Her HES developed into her TWA for explaining theoretical probability and designing experiments. I consider her use of the HES as an intuition about expected results based on *a priori* knowledge. I also conjecture that her *a priori* intuitions are linked to her strong understanding of independence. She recognizes the theoretical probability for each random outcome and has an intuition that every random event is independent. Carmella also has *a postiori* intuitions that she used in her initial understanding about the effect of a large number of trials on the decreased likelihood of an event. This intuition, coupled with her digital experiences with the dynamic data table and graphs during the simulation process, develops into her understanding of the evening out process (EOP).

Her use of appropriate proportional reasoning became more stable throughout the teaching experiment, although there is evidence of her use of inappropriate additive reasoning during final sessions of the teaching experiment and her post-interview. I conjecture that her computer experiences with the marble environment, weight tool, and the dynamic graphs helped her to develop stronger multiplicative reasoning appropriate for proportional situations. Her proportional reasoning is linked to her TWA to experimentation and her understanding of the EOP. Her TWA suggests maintaining the simplest proportion possible, yet she could also use proportional reasoning to predict the theoretical distribution based on *a priori* knowledge. Carmella's "even vs. close" dilemma and her question about whether to use a small number or a large number of trials were central to her probabilistic reasoning on many tasks.

There is evidence to show how Carmella improved her proportional reasoning, expanded her understanding of independence, increased her use of part-whole reasoning with theoretical probabilities, and developed some combinatoric reasoning for creating a sample space and using the sample space to determine probabilities. However, I think Carmella's major development in her probabilistic reasoning lies in her mathematical thinking with the “exact vs. close” dilemma and the importance of a small or large number of trials. Through the variety of problem-solving situations in which Carmella was engaged, her intuitions, strategies, concept schemes, and her social and digital interactions facilitated her development of a powerful mathematical idea. She realized that the probability of getting the exact theoretical distribution [P(exact)] decreases as the number of trials increases. In addition she realized that the probability of getting relatively close to the theoretical distribution [P(close)] increases as the number of trials increases. I suggest this discovery will help her critically think about experimental results and the purpose of experimentation depending on the intended goal. I am impressed with the depth of her thinking and must admit that her development of these ideas has enriched my own understanding of this relationship and a deeper appreciation for the law of large numbers.

Overall, Carmella improved her probabilistic reasoning skills throughout the teaching experiment and developed important mathematical ideas as well as expanded her problem-solving tools to include more graphical visualizations. I believe the microworld tools helped her to make mathematical connections between part-part, part-whole, numerical and graphical representations. Her primary intuitions have developed into strong secondary intuitions that facilitate her deep and critical analyses of probabilistic situations.

CHAPTER 6

THE CASE OF JASMINE

Jasmine is nine years old and in the fourth grade at a rural elementary school outside of a mid-size college town. Although she is considered average to slightly above average by academic achievement standards, she does have a known learning disability caused by a delay in her visual perception. Her parents cautioned that she might need additional time to process the information shown on the computer screen. Although she seemed shy at first, Jasmine showed the most enthusiasm during the teaching sessions and continually shared her thoughts with the group. She seemed empowered by the technology and continually wanted to use the software tools to model a variety of playful and real-world chance situations. The visual perception disability mentioned above did not appear to hinder her activity with the computer tools nor her interpretation of actions and images on the computer screen.

Pre-Interview Analysis

Jasmine's pre-interview occurred on August 17, 1999 for about one hour and consisted of the tasks listed in Appendix E. Throughout the interview, she used the materials available to her (e.g., coins, bags with black and white marbles, spinners, buckets with cubes) and used paper and pencil to record data. I will briefly discuss our interactions during each task and summarize my understandings of her conceptions.

To begin the interview, I asked her to interpret the following situation:

> Suppose you and a friend were playing a game and sometime during the game your friend said to you that she wanted to quit because she didn't think the game was fair. What do you think she means by the game is not fair?

Jasmine was not able to give an explanation for why a game would not be fair. In response to my inquiry about what makes a game fair, she replied "if you take turns." Since her conception of fair in the context of a game seemed to be limited, I gave her seven blocks and asked what she would do to share these fairly. She immediately used an equal partitioning strategy, gave each of us three blocks, and said "I would split them and we would just not use this one." She justified the fairness "because we each have the same amount. If this were a cookie (pointing to the one block she had put aside) we could split that one in half." Although she associated equal with sharing fairly, her concept of fair in a game context only used a limited view of equal in terms of taking turns and perhaps an assumption from her prior game-playing experience that if each player has an equal number of turns, they have an equal chance of winning.

<u>Bucket of Cubes</u>

During the interview, I used a bucket of colored cubes to assess how Jasmine described the chance of a certain color being randomly chosen, as well as her understanding of vocabulary such as "most likely" and "least likely."

<u>Nine cubes</u>. For the first task, I presented Jasmine with a bucket containing four green, three red, and two yellow cubes. She used qualitative reasoning to determine that green was the most likely "because there are more green" and yellow was the least likely "because there are least yellows" and "there are more of the other colors." Although she never explicitly referenced the number of cubes of each color, her qualitative response

indicates that she used part-part comparison to make her judgment. To continue the task, I closed my eyes and randomly picked a cube from the bucket.

> T: And I picked out a green. Now I'm going to put this back in. If I were to do it again what do you think that I would pick out?
> J: Probably another green.
> T: Probably another green? Why do you think I'll probably get another green?
> J: Because there are more greens.
> T: [I pick out another cube] I got a yellow. Now do you think with me picking out a yellow, if I were to put this yellow back in, would I be more or less likely to pick out a yellow again?
> J: You would probably get a different color.
> T: I would probably get a different color? So probably either the red or the green?
> J: Yeah.
> T: And why do you think I would either get the red or the green?
> J: I don't know. I just think that.

Jasmine referred to the quantity of green cubes in the bucket to support her reasoning for being more likely to pick out another green cube. She did not, however, use the contents of the bag to support her reasoning about being more likely to pick another color after I had chosen a yellow cube. Although her response seems correct, her lack of verbalized reasoning makes it impossible to understand exactly why she thought the other two colors would be more likely to be chosen after I picked the yellow cube.

I continued the task by asking Jasmine which color had the best chance after I removed a green cube (leaving three green, three red, and two yellow). She used a direct part-part comparison to say "red and green because they both have the same number." After I removed another green cube from the bucket she stated that she was most likely to pick out a red cube "because there is one more than the green and yellow." In addition she noted that green and yellow were tied for least likely because "they are both two [cubes]."

Later in the interview, I brought this bucket back out and asked her if the chance of picking a green cube had changed from the first bucket (4 green, three red, two yellow) to the contents after I removed two green cubes (2 green, three red, two yellow). Although she still used part-part comparisons in her reasoning, note her spontaneous use of percents to quantify the chance of picking a certain color cube.

> J: Two greens have been taken away. And now there are less greens and more reds. So you will probably pick a red.
> T: So how would you describe the chances of green when the two were in here? [I place the two green cubes back in the bucket]
> J: About 20%.
> T: How do you describe it when they are out of there? [I remove two green cubes from the bucket]
> J: 30, probably.
> T: How about the red ones. Now that these two greens are out, has the chance of picking a red one changed?
> J: Yeah. It's higher.
> T: It's higher. And why is it higher?
> J: Because they used to be one more green than red. But now that two greens are taken away, they are one more red than green.
> T: So what does that do to the chances of picking a red?
> J: It made it a little bit higher for red.
> T: [I place the two green cubes back in the bucket] What about the chances for picking a yellow cube?
> J: It's still down low. It's tied with the green.
> T: It's tied with the green. So how would you describe the chance of picking a yellow?
> J: Real low.
> T: [I remove two green cubes from the bucket] And now the chance of picking a yellow?
> J: Probably 50.
> T: So has it gone from real low to 50? It has changed a little bit?
> J: Oh now? It's changed a little bit. I'd say it was 50 before when the green was in. Now it's about [pause] the yellow and the green are tied. The red is the only thing above the two of them. It would probably be around 20.
> T: Twenty percent?
> J: And 10 for the red.

Although Jasmine compared the number of cubes for each color and correctly indicated that the chance for choosing each of the three colors had either raised (red and yellow) or lowered (green), her use of the percentages is not normative. It appears that she thinks a higher percentage means a lower a chance (e.g., chance of green lowered but she stated it went from 20% to 30%) and vice versa (e.g., she indicated that the chance of yellow increased but used 50% to 20% to quantify the change). It also appears that she does not have a sense that the chance for each color should sum to 100%. Her use of percentages also occurred in other tasks. I will further address her non-normative use of percentages to quantify the chance of something happening in subsequent tasks.

Four cubes. Another bucket task involved three green cubes and one red cube. Again, Jasmine used percents to describe the chance of picking a green cube as 10% and the chance of picking a red cube as 20-30%. She is consistent in her use of “reverse percentages” and does not show any indication she thinks the percents should sum to 100%. I then removed one green cube and one red cube and asked her to describe the chance for picking out a green cube. She replied “zero percent” and when asked what zero percent meant, “that you are going to pick out a green.” She consistently applied her reversed percentage strategy to quantify a certain event as 0%.

Coin Tosses

Two types of tasks were used involving coin tosses. The first task was designed to assess the concepts of equiprobable and sample space, while the second task assessed the concept of independence.

<u>1, 2, and 3 coin toss.</u> I gave Jasmine a penny and asked her what different ways it could land if I flipped it in the air. She easily noted that "it could land tails or heads up" and when asked if one side would occur more often if we flipped the coin several times, she said "if you flip it 10 times it's probably going to be the same for each." Her response indicates she believes heads and tails are equiprobable and that for a given number of trials the results should be equal.

For the next task in the interview, I asked her what the possibilities would be if I flipped two coins at the same time. She quickly listed the results HT, HH, TT, TH. When I asked her if TH was different or the same as HT she said "it could be the same. It's still heads and tails. What would make it different was that the quarter is the tails and the penny is heads [she models this with the coins] or … [she flips the quarter over to show heads and the penny over to show tails]." Her response shows that she is able to reason why TH and HT could be considered the same as a combination, or different as two permutations. In addition, when asked if any of the four possibilities were more likely to occur, she said "I think the head and the tail, one of those two [TH or HT] would come up more often," although she could not justify her response. It seems she has an intuition that a head and a tail, in any order, is more likely than two of the same kind (HH or TT).

When listing the possibilities for flipping three coins (penny, quarter, nickel), Jasmine used a systematic "flipping" strategy to correctly list all eight possibilities (HHH, TTT, HTH, THT, TTH, HHT, THH, HTT). She explained that while making her list she "was thinking about ways you can mix them around and turn these over [models turning HTH into THT by flipping over all three coins]." Her flipping strategy was very effective in helping her approach this task.

To assess her ideas about independence and probability, I asked her if any of the eight possibilities she had listed were more or less likely to occur than the others. She again thought of the possibilities as a combination category when she said “probably either two heads and a tail or two tails and a head.” She further said that these combinations included all the possibilities except HHH and TTT. Because she was combining several possibilities together in her response, it is difficult to assess whether she believes all eight possibilities are equiprobable. However, her grouping of possibilities shows she thinks that non-specific combinations of heads and tails are more likely than a single combination of all heads or all tails. She does not, however, justify her response based on the number of possibilities that are a combination of heads and tails (6). She further explained that HHH and TTT had an 80% chance of occurring while the other six combinations together had a 10% or 1% chance. Her use of the percentages, although non-normative, is still consistent with her previous use.

Flipping a coin six times. Later in the interview I asked Jasmine to predict what I would get if I flipped a penny six times. She expected that I should get three heads and three tails but that it would be hard to predict the exact order because “it could come out in a lot of different ways.” She seems to be applying the notion equiprobable to predict an equal distribution but also recognizes that there are numerous possible lists of six results that have three heads and three tails.

For the next series of questions, I showed her four possible results from tossing a coin six times and asked her if any of the possible results of six were more likely to occur. For the first set (HHHHTT, THHTHT, THTTTH, HTHTHT), she thought THHTHT would be most likely because “it’s mixed up more than most of the others.” In addition, she thought HTHTHT was the least likely but could not give a reason for her response. The list she chose as the most likely has three heads and three tails and at most two of the same result in a row.

For the next set (HHHTTT, HHHHHH, THTHTH, HTHTHT), she first noted that HTHTHT and HHHHHH were least likely but then added that "the others I don't think would be very likely…you mostly wouldn't see any like that…they are [all] pretty unusual." Although she could not explain what was so unusual about these results, when I asked her what would make something usual or normal looking, she responded "if it was mixed up a lot" but not in a pattern like HTHTHT. Based on her responses to both sets of results, it appears that she intuitively believes that results from a coin toss should not be in any discernable pattern and they should be sufficiently "mixed up." Her sense of what random results should look like seems to directly affect her intuition about certain combinations being more or less likely. She did not display any notions of independence in her responses.

For the final coin task, I told Jasmine that I had flipped a coin six times and got the results THTTTT. I asked her if I flipped it one more time, "do you think I'm more likely or less likely to get heads or tails or are they equally likely?" She replied "you will probably get more heads…because you already have a bunch of tails." However, she continued to support her answer based on her memory of a coin experiment on television.

> J: I was watching Bill Nye the Science Guy once and he said he had a coin flipping thing. And it said that if you already have more tails, it was something like this, and if you have more, I think it was heads, they said you were more likely to get a tail.

Whether the television character actually said that, or that was Jasmine's interpretation of the discussion about the coin experiment, is not certain. However, her memory of this experiment seems to directly affect her own judgment about the likelihood of heads or tails after four consecutive tails with a coin flip. Her response displays classical gambler's fallacy reasoning based on the representativeness of this small number of trials on what she expects from equiproble events.

Sampling

To assess Jasmine's sampling strategies, I gave her a black bag and told her that it contained 10 tiles of three different colors (5 blue, three red and two yellow). Her task was to make a reasonable guess at what was in the bag using a with replacement sampling method. She chose to draw from the bag 11 times without giving a reason for this number. After four picks from the bag, she had chosen BYBR and said she knew the three colors in the bag but did not have enough information to tell how many of each because "I need more…I'll end up guessing anyway but I'll have a higher chance of getting it right if I pick more." This statement indicates that she has an intuitive sense that sampling from a bag can not guarantee an exact prediction and that the more you sample the higher your chance of making a more accurate prediction. She eventually picked a sample of five blue, four yellow, and two red tiles.

> T: Based on your information, can you tell me what you think is in the bag?
> J: There are more blues. I'd say medium yellows and I think there are only a few reds.
> T: How confident are you?
> J: I'm pretty sure there are more blues. Unless I was just picking out the same one each time. Because if there were more blues I would be more likely to pick that out. If there are less reds I'd be less likely to pick that

> out. And if yellow is pretty much in the middle, sometimes I wouldn't pick it out as much as the blue.

Her justification is based on a part-part comparison of the contents and the notion that more parts indicate a more likely event and vice versa. When she found out that the bag actually contained more red tiles than yellow she replied "it only means that mostly it [red] would probably come up…it won't always come up." This response indicates she intuitively expects variance in results even if the theoretical probability is known.

<u>100 Gumballs</u>

This task was used to assess her ability to use proportional reasoning and theoretical probability to make a prediction for a sample when the population is known. Given a gumball machine with 50 red, 30 blue, and 20 yellow gumballs, she predicted a sample that closely mirrors the distribution in the bag but only gave possible ranges of the number of gumballs of each color (4-5 red, 3-4 blue, 1-2 yellow). She supported her response by referring to the distribution in the gumball machine and noting the likelihood of each color based on the number of gumballs (e.g., "there are more reds than anything else so you are more likely to get red"). She also noted that the gumballs would have to be "all mixed up" for this type of distribution to be picked. Her use of a range indicates her reluctance to make an exact prediction but the range of numbers she chose include values that reflect the distribution in the "population" of gumballs.

Marble-Bag Comparisons

For the first two pictures of bag of marbles, I only asked Jasmine to describe the chance of picking a black marble. Bag #1 contained two clear and two black marbles. She claimed that both the black and clear marbles each had a 9% chance of being chosen because "they are both the same. They both have the same chance because there are two of each together." Although she correctly reasoned that the black and clear had an equiprobable chance and used the same percent to quantify both chances, her use of 9% again displays her lack of understanding that the percents sum to 100% and her use of low percentages to indicate a good chance. She was not able to give any specific reason for her use of 9% beyond that each color should have the same chance.

In her assessment of bag #2 (5B3C), she used 10%-20%, but "probably closer to ten though," to describe the chance of picking a black marble and 13% chance to choose a clear marble. She also noted there was a better chance of choosing black because "there are more blacks." Again, she seems to be consistent in her use of percent to describe the chance of something occurring.

For the remaining pairs of bags (#3 & #4, #5 & #6, #7 & #8), I asked her to choose which bag she would prefer to pick from if the goal is to try to pick a black marble. With each pair of bags, I reiterated a question such as "would you like to pick from bag #3, bag #4, or does it matter which bag you choose from?" For the first pair of bags, the distribution in each bag was proportional with 3B1C in bag #3 and 6B2C in bag #4. She immediately chose bag #3 and justified her reasoning as follows:

> J: It [bag #3] may have less, least black ones. But it has the least clear ones too. That when this one [bag #4] is six blacks and two whites. And this one [bag #3] is three blacks and one white. Then one white, there is only one white so it's a better chance of picking a black.

Even though bag #4 had more black marbles, Jasmine thought the bag with the least number of white marbles (the undesired event) would give her a better chance of picking a black marble (the desired event). She used similar reasoning in comparing Bag #5 (1B4C) and bag #6 (2B8C).

> J: Bag #6.
> T: Oh, that was quick. Why do you want bag #6?
> J: There is only one [black marble] in bag #5. And there are two [black marbles] in bag #6.
> T: So the chance of picking a black one here and the chance of picking a black one here.… You think you have....
> J: No, I think it could be bag #5 because there was clear ones to fight against. Oh okay. I actually think it would be #5.
> T: Oh, so all of the white ones over here [in bag #6]…
> J: Well it would be blocking the two, because the only two black ones could be way down here [points to bottom of bag]. But there are only four clears one in that [bag #5].

Although she at first makes her decision based on the number of black marbles, she reassesses her decision based on the number of clear marbles that there are "to fight against." In both the previous comparison tasks, she shows no evidence of proportional reasoning. The last two bags were not proportional with 2B2C and 2B3C in bag #7 and #8, respectively. She used similar reasoning and chose bag #7 since it had the least number of clear marbles.

Spinner Game

The next task consisted of the penny game (as described in Chapter 5) played with a spinner containing three unequal sectors (1/2 red, 1/3 blue, 1/6 yellow) and eight pennies. Jasmine chose the red sector because "it takes up half the circle so it is more likely to get there." I chose the blue sector and said "we are going to take turns spinning the spinner, so is this a fair game." She replied "yeah" since we were taking turns and did not mention the unequal sizes of the sectors. After we played the game for quite awhile, I finally won all eight pennies. I again asked her if the game was fair and she replied "yeah." To help her focus on the size of the sectors, I asked her why she thought we did not land on yellow very many times.

> J: Because there is hardly any yellow. It's only 10%.
> T: And how do you think the red compares with the blue as far as the chance of getting the red or the blue?
> J: 50%. Because that goes down to the middle of the red.
> T: So when you say 50%, which one is 50?
> J: Oh no, this whole thing is 100 and this [red] is 50 and that's [blue and yellow] 50 because that's put together. So it [blue] would be 35.
> T: So the red you are saying is 50.
> J: Yeah, 50%.
> T: And those two together are 50%. And so the blue you are estimating to be like 35.
> J: Yeah.
> T: So are the red and the blue equal? Do they have an equal chance of coming up?
> J: When what happens?
> T: Whenever we spin, since this [red] is 50% and this [blue] is 35%. Do they have an equal chance of happening?
> J: This [points to red] is probably going to happen more.

Two points are noteworthy in our exchange. First, with an area model she spontaneously self corrected her former use of percentages. All the other situations where she used percents to quantify a chance were discrete models. With this area model, she identified the red sector as 50% and estimated the blue to be about 35% and yellow as about 10%. She also noted that the whole pie represented 100% and the blue and yellow should sum to 50% (although she did not account for 5% in her estimation of 35% and 10%). I conjecture that she has probably had prior experience in representing ½ as half of a circle and as 50%. Thus, the area model allowed her to connect those representations with her use of percents to quantify a chance of something occurring. The second point in the above conversation is that she did recognize that the red sector is "probably going to happen more." She did not spontaneously connect this observation with the fairness of the game. I, however, did not directly ask her again about the fairness of the game.

For the final task in the interview, I asked Jasmine to compare two spinners and decide which one she would rather use to play the penny game if she still won when the arrow landed on a red sector. Spinner A was the same one used in the previous task. Spinner B contained 12 equal sectors, six red, four blue, and two yellow spaced in the following pattern (r, b, r, y, r, b, r, b, r, y, r, b). At first she chose spinner B because "it's a different one and it looks more fun." I then asked to compare the chance of landing on red in each of the spinners.

> T: How do you think the chance of getting a red over here [spinner B] compares with the chance of getting a red on the first spinner?
> J: The same.
> T: The same. Why is it the same?
> J: Let me look at this. That would be the little triangles put together would equal 50%.
> T: Why would they equal 50%?
> J: I'm just guessing that they would be equal.
> T: How many of those red …
> J: Because like the yellow slices equal that [yellow sector on spinner A]. And the blue ones [on spinner B] equal the blue [on spinner A].
> T: How could you prove to me that these two yellow slices are the same as that?
> J: They have the same size. I see that that's [one yellow sector on spinner B] half of that [the yellow sector in spinner A].
> T: Okay it looks like it's half of that.
> J: It looks like it's half of that, so this [points to the other yellow sector in spinner B] is the same kind as that [the first yellow sector in spinner B] so two of those would be that [the yellow sector in spinner A].

T: And what about the reds? You said you thought the reds, that these six reds over here were the same size as this one, the large red.
J: It's in several different sizes of that.
T: How many red slices do we have over here?
J: Six.
T: We have six red slices. And how many total slices do we have?
J: 12.
T: And so six of those are red.
J: Yeah.
T: [pause] So if we were to play this again, would it matter which spinner you chose?
J: It wouldn't matter.

Her reasoning was totally based on her visual estimation of equal areas. Although I attempted to get her to focus on the numerical aspect (6 out of 12 sectors in spinner B are red), she did not use the numbers to justify her belief in the equivalence of the areas. Even though she is correct about the equivalent areas, she only needed to use a visual estimation of the areas, and did not use any numerical justification.

Strengths and Weaknesses from Pre-Interview

Jasmine had several primary intuitions about probabilistic concepts. She had a sense that "fair" involved equal parts (e.g., equal number of "turns") but did not apply this conception in evaluating whether a probabilistic situation was fair based on equal chances. She showed evidence of understanding the uncertain nature of random situations (e.g., she recognized that the "most likely" event does not always occur more often) but thought that results from a random experiment should be "mixed up" and not in any type of pattern (e.g., she thought results from a coin toss where less likely to be in a pattern such as HTHTHT or TTTTTT). Jasmine was also able to use part-part reasoning to find the least and most likely event, including events that have changed because of a without replacement situation. In comparing the bags of marbles, she used part-part reasoning and consistently used a strategy that the bag with least number of undesired events made the desired event most likely. Since it was not necessary, Jasmine never explicitly referred to the whole in making probability judgments.

Her interesting use of percents demonstrates that she has an intuition about quantifying probabilities, albeit non-normative. In every discrete case, she used her percent description consistently. Higher percents meant a lower chance while lower percents indicated a more probable event, even using 0% to characterize a certain event. However, in discrete cases she made no attempt to use percentages that summed to 100%. The continuous areas on the spinner were the only instance when she applied normative quantitative percents based on 100% as the "whole" and 50% as representing half. The spinner tasks were also the only tasks where she used proportional reasoning based on the relationships in the geometric areas. Jasmine never explicitly used numerical proportional reasoning.

Jasmine was able to easily list the four elements in the sample space for the 2-coin toss. With the 3-coin toss, she used a systematic flipping strategy to list all eight possible arrangements. However, with both the two and three coin toss experiments, it appeared she reasoned with the unordered combinations when asked if any arrangement was more or less likely or if they were all equally likely. She thought that either the combination of one head and one tail (with two coins) or two heads and one tail or two tails and one head (with three coins) were more likely to occur. Her intuition about the higher chances was appropriate since she was considering the unordered arrangements. However, Jasmine showed no evidence of understanding the independence of events in a series of coin tosses. She displayed typical gambler's fallacy reasoning in her thinking that several results of heads in a row increase the probability that the next result will be tails.

Overall, Jasmine's responses during the pre-interview demonstrate that she entered the research study with intuitions about chance that she is not necessarily able to quantify in normative ways. She relies on part-part reasoning and did not display any instances of part-whole reasoning. The only hint that she even considers the whole is with her reference to the whole pie in the spinner as representing 100%. The area model seems to be a familiar context for her and provided her with a cognitive prompt to analyze the chance of events occurring. Jasmine's strength was certainly her systematic approach to finding all possible arrangements. She also explicitly referred to the unordered arrangements in her analysis of whether the arrangements were equiprobable. This shows that she has already developed a primitive mental scheme for combinatoric reasoning.

Jasmine's Meaning-Making Activity with the Microworld

Jasmine participated in approximately seven hours of small group teaching sessions and 3.5 hours of individual sessions. Jasmine forgot to attend one of the scheduled group teaching sessions and subsequently had three separate scheduled individual sessions to make up for the unplanned absence. Thus, after the fourth teaching session, there was a lapse of 17 days before she attended the next teaching session.

The analysis of the teaching sessions with respect to Jasmine brought forth four evidentiary themes in her development of probabilistic reasoning: 1) her interpretation and use of theoretical probability; 2) the "evening out" phenomenon; 3) the relationship between the “whole” and its “parts;” and 4) her use of additive and multiplicative reasoning. What follows are my observations and analyses of Jasmine’s meaning-making activities, mathematical ideas, intuitions, and conceptions under each key theme. I also highlight how she used the microworld tools to simulate and analyze probability experiments. For cross-case comparison purposes, I have also included a thick description of Jasmine’s work on the “twice as likely” task.

Interpreting and Using Theoretical Probability

As the teaching experiment progressed, Jasmine developed her ability to interpret theoretical probability. She also developed her ability to predict and interpret experimental results based on the theoretical probability.

During the first teaching session, I used the bucket of cubes from the pre-interview (4 green, three red, two yellow) to have the girls discuss the chance for picking out each color. Carmella volunteered that green had the best chance since it had a “four out of nine chance” and justified that four out of nine meant “if you pull out all the cubes, there will be nine and four of them will be green.” Jasmine described the chance of getting a yellow as “two out of nine” and justified her reasoning with reference to the bag’s contents (i.e., nine cubes in all, two of which are yellow). This was the first time Jasmine had used any type of part-whole language to describe the chance of an event occurring. Although it appears she modeled the language after Carmella, she interpreted the statement slightly different. I also asked them to describe the chance of picking a black marble from a 2W2B bag. Carmella offered “they have the same amount … even, even…50-50” and Jasmine interpreted 50-50 to mean “they are each 50%.” This is

Jasmine’s first correct use of a percentage to describe a probability from a discrete situation.

During the first and second teaching sessions, Jasmine used the equiprobable events with both a coin and die to predict “about even” results when running experiments. The justifications she used for her predictions were mostly based on a coin and die having an “even” number of outcomes. She also referred to the evenness of the number of trials as a justification as well (e.g., 100 is an even number). Overall, her justifications at this point were qualitative in nature and only slightly hinted that she was basing the predictions on the theoretical probability of each event.

During the third teaching session, Carmella and Jasmine were simulating experiments with a 2B2W bag of marbles. Jasmine predicted they would get either 5-5 or 6-4 since the black and white were “even.” They ran several sets of 10 trials and at one point they got 8W2B marbles. Jasmine said “wow” and Carmella added “very unlikely.”

> T: So why is it so unlikely that we get so many whites?
> C: Because there’s two you can pick out of so there’s no guarantee you’ll get most of one.
> J: It’s very unlikely if you get all white or all black.
> T: And why is that very unlikely?
> J: Because there’s two of each.

Jasmine’s “two of each” shows she explicitly used the distribution of marbles in the bag as an indication of equiprobability. During this teaching session, the girls also experimented with a 5B5W bag of marbles. Carmella described the chance of picking a black marble as “it would be a five out of ten chance” and Jasmine agreed because “there are five and there are ten, there are five blacks and ten total.” Adopting Carmella’s language, she was able to identify the part and the whole in a statement of probability.

Jasmine also spontaneously pointed to the picture of the 2B2W bag and said “two out of four.”

> T: So how did the chance of picking a black marble from the first bag [2B2W] compare to this bag [5B5W]?
> J: Not really that different.
> T: Not really that different. Why not?
> J: It’s only two out of four and five out of ten.
> T: Yeah. So why aren’t those very different?
> J: Because they are still an even number.
> T: All right. So an even number. What do you think about that Carmella? The chance of picking a black is two out of four and the chance of picking a black being five out of ten?
> C: Umm, it would be the same thing.
> T: The same thing. Why is it the same thing?
> C: Because it’s pretty much there are four marbles and ten marbles. They are even numbers ... so it could be either one
> T: Could be either one? So are those chances of picking out a black marble the same? [Carmella nods “yes”] Are they exactly the same?
> J: No.... The only way they are different is because it’s, there are four marbles. And ten marbles. But the chances are the same.

Although she recognized the equivalent chances in each of the bags and stated the chances in a part-whole manner, she justified the equivalence “because they are still an even number.” Although neither girl explicitly discussed that the parts in each bag are equal, it is unclear whether an “even number” refers to the even property of four and 10 or the even (equal) distribution of marbles in each bag. This was clarified when Jasmine noted there was an “uneven” chance for black to get picked in a 3B1W bag of marbles.

> T: So why do we have an uneven chance of picking out a black?
> C: Because there’s three.
> T: But, there’s two colors in there. And there’s four marbles. And four is an even number.
> J: Yeah, but there are two more [blacks] than there is white.
> C: We were talking about even numbers of colors. It’s not an even number of colors. The white one has one and the black one has three. [To illustrate, Jasmine moves the marbles on the screen to pair up a white marble and black marble and then drags the other two white marbles to separate them from the pair.]

This illustrates that both girls continued to use a part-part comparison in determining chances and that they had been using the term “even” as synonymous with equal. Furthermore, when running several sets of 10 trials, they got many “uneven” results (e.g., 9B1W, 8B2W) but were surprised when they got an “even” (5B5W) distribution.

> J: That’s pretty unlikely.
> T: Why is that unlikely?
> C: Because it’s WAY less whites....
> J: Even though it’s an even number of marbles. It’s still unlikely because there’s only one white ... So you would have to pick out the same white each time. And there are three others against that white so that’s pretty unlikely.

This episode illustrates how Jasmine used theoretical probability to assess the likelihood of experimental results. Although she used a part-part comparison, she reasoned that the unequal chances for choosing a black or white marble made it more unlikely to get equal experimental results. Her notion of choosing the same white marble each time also demonstrates that she was connecting her reasoning with the characteristics of a real bag.

During the fourth teaching session, the students were playing a coin tossing game on the computer and I had secretly changed the theoretical probability to be 5/6 heads and 1/6 tails. Up to this point, the students had not used the weight tool and did not know that capability existed. After playing the game many times with Joe, Jasmine was very excited that she was winning so many times but was obviously perplexed about the number of times “heads” had won the game. She noted that she expected a more even distribution and attempted to explain the “problem.”

> J: I think it’s something in this computer. But I can’t explain it …
> T2: Why do you think it’s something in the computer?
> J: Well, actually I don’t. Umm, I think that I would say this is probably unfair.

Although she was not sure, the skewed results did provide data that did not match her intuitive conception of fair, thus she termed this "unfair." It seems that now her notion of fair did include a conception of equal chances and about equally distributed results. To test the unfair advantage of the heads, Jasmine ran 500 trials and predicted there would be more tails but "not many." She subsequently ran several more sets of 500 trials as an accumulation of trials and noted that the tails would take a long time to catch up. Joe then asked her to predict what would occur if she cleared all the trials and ran a new set of 500 coin tosses.

> J: Well the tails will probably be ahead of the heads.
> T2: And what makes you think that?
> J: Well actually it's more of a chance that they are going to be closer together than they are now.

In spite of all the experimentation, her intuition was to revert back to an assumption of equiprobable. After many more trials, the girls blamed the "problem" on either a computer bug or thought I was controlling the software with a machine that would give out more heads. Eventually Carmella discovered the weight tool (see Figure 6.1).

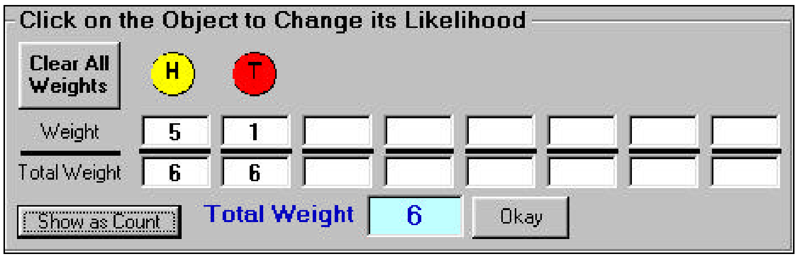


Figure 6.1. Weight tool displaying "mystery" chances for heads and tails.

Jasmine's work with the weight tool demonstrates her initial struggle to interpret the weights and her discussion of the weights in terms of chance and experimental results. At first, Joe asked her to interpret the weights displayed in Figure 6.1.

> T2: What do you think this means? What does the five and the one mean?
> J: The five means that there's going to be more heads … and that there's five out six heads and one out of six tails.
> T2: So do you think that could have caused the results?
> J: Yes. I think it's making it five and 1, because she had already turned it on. [The *Probability Explorer* was open when the students arrived.]
> T2: That's right. And what does that five and one result in?
> J: That there would be 500 or 50 more than the one each time.
> T2: 50 more than the one? Explain that.
> J: I don't know. [shrugs shoulders]

She knew the weight of five made heads more likely to occur and could state the individual probabilities as a part-whole relationship. However, when Joe asked what the five and one result in, she noted there would be "50 more than the 1" but could not explain that statement. It seems that she may be referring to an additive relationship between the number of heads and tails from experimental results.

Joe continued this dialogue by asking Jasmine how she could make the results more "even."

> J: More even? [She clicks on the picture of the tails five times.] This [tails] needs to go up. [Now the weights are 5/11 and 6/11]
> T2: You went up. [She clicks once on the heads icon and the weights change to 6/12 and 6/12.] Okay now what, is this even?
> J: Bet there won't be trouble with the tails now!
> T2: So what should happen for six and 6?
> J: They should be around even. At least closer. [She runs 500 trials and watches the pie and bar graphs during the simulation and notes how the heads and tails "even out" just as she predicted.]

Jasmine knew that to get close to even results she had to change the weights to be equal and that this equal distribution of weights would fix the "trouble with the tails." She also knew to run a simulation to test out the "evenness." Joe then asked her how she could make it so tails would have an advantage. She changed the weights to 9/22 and 13/22 in favor of the tails and ran another set of 500 trials to show Joe that the tails could "win." When the simulation was over she spontaneously opened the weight tool, cleared all the weights and made the weights 5/10 and 5/10. Joe asked her if those weights were any different from weights of one and 1.

> J: Is that different from one and 1?
> T2: You have five and 5.
> J: No, it's not different because it's the same number.
> T2: What do you mean? [She clears the weights and clicks on the heads once and tails once so the weights are ½ and ½.]
> J: One and one.
> T2: Is that any different than putting five and 5?
> J: The only difference is because of the number, but it's not very different there are still two of the same number…they will be close to the same when we run it.

Jasmine was very comfortable using the weight tool to model different situations and recognized that equally distributed weights are equivalent and will result in similar experimental distributions. This shows that she was beginning to develop an understanding that different numbers can be used to represent the same theoretical probability.

At one point during the fifth teaching session, Jasmine was telling me about a bingo barrel that her teacher used to randomly draw different numbered balls. There were 20 students in her class and every student had a numbered ball (1-20) in the barrel. If the teacher pulled out a student's number, he or she was chosen to do some task in the classroom (e.g., take attendance to the office, answer a homework problem). Jasmine noted that her number had only been chosen once since the beginning of school (about 15 days) even though she knew that each of the students had an equal chance of being chosen. She wanted to use the marble environment to model the bingo barrel. She started putting in many white marbles and then noted that she needed to "put in a black one" to represent herself. She eventually placed 19 white and one black marble in the bag. I asked her to describe the probability that her teacher would pick her ball.

> J: one out of 20.
> T: one out of 20. What does that mean, one out of 20?
> J: That's the chance of picking me out, being picked out of all the others. But it's the same for all of us. Each one has the same chance.
> T: [I pointed to a white marble.] So this marble right here has the chance of what?
> J: Has the chance of one out of 20.
> T: What's the chance of picking out a white?
> J: A white? [pause] 19 out of 20.
> T: And why is it 19 out of 20?
> J: There are 19 and there are 20 marbles.

Jasmine was not only able to use the microworld tools to model the classroom bingo barrel, but she used her understanding of equiprobable to analyze the chances for each ball to be chosen. She also was able to think about the chance of any white marble as 19/20 and later noted "that's the chance that someone else besides me gets picked…no wonder I hardly ever get picked." Her thinking during this episode indicates that she could use theoretical probability to interpret and analyze a real-world situation.

There were several other instances when Jasmine displayed evidence of interpreting and using theoretical probability with the weight tool. During the fifth teaching session we revisited the coin toss experiment with the weights of five to one. I asked her if she thought "we would be able to get it so that it's even? So we have 50 heads and 50 tails?" She replied, "if we change the scale [the weights] to even… it is possible this way [weights of five to one] but not likely." She was able to reason that the unequal weights made a 50-50 distribution unlikely and that she thought changing the scale "to even" would give her a better chance of getting "even" experimental results. Jasmine also was able to use the part-part and part-whole relationship in the weights to justify a prediction for an experiment with weights of ½ and ½. For 10 trials she predicted "five and five … because 10, half of 10 is five, and five and five are the same number, and five and five is 10." Instead of justifying her prediction with reasons about "even" numbers, she uses a more sophisticated justification that relies on the "same number" (part-part) and "half" (part-whole) relationships. Her reasoning demonstrates that she was able to use both types of relationships to think probabilistically about expected experimental results.

Jasmine also demonstrated her understanding of the concepts of unequal, impossible, and certain. In the first scenario, she had designed an experiment with four different icons (tails, circle, hexagon, volleyball) and the weights were set equal with each having a ¼ chance. I asked her if they each had the same chance and she spontaneously used the weight tool to illustrate her thinking.

> J: They each have one. But they wouldn't have the same chance if some one did [she changed the one under the tails to a zero] that. Then there wouldn't be any of those. Or how about this? [she changed the zero under tails to be a two] now it's more likely to get the tails because there are two out of five. But there's only one circle, one hexagon, one volleyball out of five.

She knew that having a weight of zero made an event impossible and that having one event with a higher weight than other events would make it more likely to occur. She used both of these situations to illustrate weights that would not represent the same chance for all four events. Jasmine also used the weight tool to model a "definite" event by changing three events to have a weight of zero and using one "or you could do any number" as the weight for the definite event. She knew to have only one event occur experimentally, she needed to use zero to model “impossible” and “any number” to make an event certain to occur. I consider her use of the weight tool to model and illustrate her thinking as strong evidence that she has developed a pretty good understanding of theoretical probability.

There is one final interpretation of theoretical probability that Jasmine made during the last teaching session. Again, she was experimenting with four events equally weighted as “quarters” and she had predicted “quarters” (both numerically and graphically) for a trial of 100. After several sets of 100 trials, she noted that only some results were close to quarters.

> T: We weighted them as quarters but we are not getting it in quarters.
> J: Oh well, close to quarters, kind of.
> T: Do you think there is something wrong with that weight tool?
> J: No. [shakes head in a strong side-to-side motion]
> T: No. Why not?
> J: Because that doesn’t mean it WILL come out in quarters.
> T: Why doesn’t it?
> J: It just means that it will come out CLOSE to quarters.

Although her experimental results did not reflect the exact distribution of the weights, she still relied on the objective a priori knowledge she had of the weights and used a probabilistic interpretation of the weights that reflects an appreciation for the random process used in the experimentation.

Overall, Jasmine’s interpretation and use of theoretical probability improved immensely during the teaching experiment. She used both part-part and part-whole reasoning and developed her analysis of experimental results with respect to theoretical probability. Her use of the marble environment and weight tool to model various situations demonstrates how the software facilitated her conceptual development and empowered her to explore probabilistic ideas.

The Evening Out Phenomenon

Recall the discussion in Chapter 5 about Carmella's developing understanding of the "evening out" phenomenon (EOP) and that the EOP is both an observable *process* that occurs during a simulation (e.g., "wiggling" then "hardly moving" motion of the pie graph), and a *working hypothesis* that experimental results tend to "even out" near the expected results. Since the students' first several experiences with the EOP were with equiprobable outcomes, their beginning understandings relied on a tendency towards an equal distribution of results; however, they eventually expanded their understanding of the EOP to include unequiprobable outcomes.

With her first experience with the EOP, Jasmine intuitively described the process she observed in the pie graph in terms of the coin toss simulation. She explained that the graph had increasingly less changes as the simulation progressed because "with the more coins getting tossed, it's keeping it the same … with more coins, one is still going to be a little bit ahead of another mostly, but it's unlikely that one will rise a lot above the other." Her notion of a "little bit ahead" suggests she expected the small variability in numerical and graphical results as part of the random nature of the coin, but that with a large number of trials she did not expect one result to "rise a lot above the other." She continued this line of thinking in the following interaction.

> J: Look, it is totally even. [pointing to Carmella's screen]
> C: It is?
> J: Yes, totally even now or at least it was. [The simulation on Carmella's screen is still running and the graph is continuously updated.]

> T: So it was even but now it's not even. So what does it mean that it was even and then it became uneven?
> J: Because some more heads got put in or some more tails got put in.
> T: So when it was even, what was happening right there where it was even?
> J: It was the same amount of heads and tails.

At this point, Jasmine's understanding of the EOP was directly connected to the observable process during the simulation. However, her statements demonstrate that she understands the process in terms of the actual experiment rather than just a dynamic visualization of the motion in the graphs.

While experimenting with a regular six-sided die in the second teaching session, Jasmine noted the wide variability in results when doing only 10 trials. At one point she had zero 1's, zero 2's, five 3's, two 4's, and three 6's. She suggested adding lots more trials to the current data and predicted that the results would "stay about the same" as she swept the mouse pointer over the table of current data. It was unclear if she was referring to the absolute differences between the numerical data or if she thought the results would stay relatively "about the same" after many more trials.

As noted in Chapter 5, Jasmine used arm motions and described the motion in the pie graph as "spin, spin, up and down" when predicting what would happen with 200 trials of rolling a die. This description emphasized the wide variability she expected but gave no indication of what she expected in terms of the EOP. However, after the 200 trials, when asked to predict what would happen if we did another 200 trials in addition to the current 200 trials, Jasmine's initial intuition was to expect "there's going to be a lot more, but it's still going to stay even." Recall, though, that when Carmella predicted the graph would "wiggle" again, Jasmine nodded in agreement. Her explanation of why the graph did not wiggle during the next 200 trials was because the pie graph was "already flattened out [and] they are already a big number." Although she had an intuitive understanding about what she expected to see, she was not confident enough to trust her intuition when Carmella had an alternative hypothesis. Yet, after visualizing that the results were aligned with her original prediction, she reasoned that once results are close to "even" and there are a large number of trials, we should not expect wide variability in future results. Although this is appropriate reasoning with regards to the EOP, this type of thinking may have hindered her developing an appreciation and understanding of independence.

When doing several sets of 40 trials with the die, Jasmine noted that the results were not always "even."

> J: See it's not always even.
> T: Ah, okay. So it's not always even.
> J: There's a bigger chance of it being even with coins.
> T: The better chance of being even is with coins? Carmella, Jasmine just said she thinks the better chance of being even is with the coins. What do you think about that?
> C: They are closer together.
> T: The numbers are closer together?
> J: They have a better chance of getting even because the numbers are closer together.
> T: And why is it a better chance with the coins do you think?
> J: Because they are less.
> T: They are less. What do you mean by they are less?
> J: Well I don't really know I am just guessing that because the numbers are farther apart with the dice.

This exchange shows Jasmine's comparison to the coin results. However, her reference to "numbers" is not clear. I am not sure if she is referring to experimental data numbers or to the number of outcomes. Her response to my inquiry was not sufficient to clearly determine the reference. In either case, she was expressing intuitions about the differences between experimenting with coins (2 outcomes) and a die (6 outcomes) and that the die seem to have a harder time “evening out” with 40 trials. However, in a later dialogue, she did refer explicitly to less “choices” with a coin, as compared to a die, as an indicator of increasing the likelihood of getting several consecutive heads (coin) as compared to sixes (die).

The work that Jasmine and Carmella did with the 2B2W and 5B5W bags of marbles in the third teaching session demonstrate their thinking about expecting wide variability in results with a small sample and the tendency towards “even” as the number of trials increased. At one point, with the 2B2W bag, they ran several sets of six trials and actually got all six white marbles. Of course the girls were very excited and Jasmine reasoned that the chance of all whites occurring was greater with six trials than it was with 10 trials but still less than the chance of all whites when they did four trials. Her reasoning illustrates a critical analysis of the effect of the number of trials on the likelihood of all white marbles. I believe her reasoning was linked to an expectation of the EOP as the number of trials increased. The more trials that are done, the more likely it is to get results close to the expected frequency (equal number of black and white marbles); therefore decreasing the chance of getting results like all white marbles.

Jasmine also reasoned about the effect of the number of trials while experimenting in the microworld with a 3B1W bag of marbles. I asked the girls if it was possible to get more white marbles than black marbles. Jasmine said it would be “possible” but that it was more likely to happen if you do a small number of trials. She changed the number of trials to three and ran several experiments until she got one black and two white marbles, and noted, “it’s a higher chance to get more whites.” Her intuition that it is more probable to get distributions that do not reflect the theoretical probability with a small number of trials demonstrates her reasoning about the converse of the EOP. In addition, when predicting results for 100 trials, I asked “what do you think about getting equal whites and blacks?” Jasmine responded, “equal whites and blacks? Woah! ... It’s possible but its surprisingly low, it would be a huge surprise.” Jasmine had an intuition that for the 3B1W bag of marbles, the likelihood of getting equal amounts of black and white marbles would be “surprisingly low” with such a large number of trials. She extended her developing conceptions of the effect of the number of trials on the distribution of results to this unequiprobable situation. She did several sets of 100 trials and noted the wide variability at the beginning and how the pie graph “settled” around the1/4-3/4 point in the pie graph as the trials approached 100.

Recall some of Jasmine’s work with the weight tool during the fourth and fifth teaching sessions that demonstrated her interpretation and use of theoretical probability. Many of her interpretations were affected by her understanding of the EOP. For example, in the fourth session, she noted that with weights of 5-1, she would have to make the tails “go up” in order to get close to equal amounts of heads and tails with a large number of trials. She also noted in the fifth session that the likelihood of getting 50 heads and tails in 100 trials increases “if we change the scale [the weights] to even… it is possible this way [weights of five to one] but not likely.” Thus, her intuitions with the EOP were contributing to her conceptual understanding of the relationship between the theoretical probability, the number of trials, and the likelihood of a certain result occurring.

Although Jasmine seems to use reasoning based on her conception of the EOP throughout the sixth and seventh teaching experiment, one of her comments in particular demonstrates that she sometimes lets her visualization of the "evening out" *process* take precedence over thinking about the law of large numbers. Her understanding of the *process* of the EOP involves an image that after many trials, the results displayed in the pie graph will not change very much. However, it appears that her image of the *process* may interfere with her development of a *hypothesis* of the law of large numbers. Consider her work in the seventh session regarding an experiment she designed with four equally weighted outcomes. After 100 trials, the results were 20, 23, 33, 24.

> T: It is close, but I don't understand why it's not more like quarters. What do you think would happen if we did a 100 on top of this 100?
> J: I think it will be about the same but close to these numbers [in the data table] doubled … and the pie graph is going to stay about the same.

Instead of reasoning about the deviation in the results getting less with more trials, she emphasized the *process* of the EOP as the pie graph staying the same. She did not yet have a strong scheme to conceptually relate the *process* with the increasing number of trials and the tendency towards the theoretical frequency. As this experimentation continued, consider our discussion during an additional trial of 100 on top of the already 200 trials displayed.

> J: It looks like they are getting closer to quarters.
> T: It does look like it's going closer. But why do you think it's going closer to quarters?
> J: Because they are adding more and more, so it can't go down.
> T: It can't go down? But how come it's not going to be something different where if I have a small part of gray and the other parts are big. Why am I going closer to quarters?
> J: Because they start out as quarters.

Her reasoning about the approximation to quarters could either be based on the process of the EOP with no reference to the number of trials, or she could have been referring to the theoretical probability. Although she did appropriately reason about the effect of a large number of trials in past investigations, she did not quantify her reasoning in this instance with reference to trials or weights. Therefore, it seems that her developing understanding of the EOP is more reliant on the process of "evening out" than it is on the law of large numbers.

Jasmine did display reasoning based on the effect of a small or large number of trials. Her thinking about the relationship between the number of trials and size of the slices in the pie graph is critical in examining her efforts to make sense of the relative display of results in the pie graph display. In the next section, I examine her intuition and developing understanding about the relationship between the "parts" and the "whole" in both theoretical probability (weights) and the pie graph display.

The Relationship Between the Whole and Parts

Jasmine did not use part-whole reasoning during the pre-interview. Her reliance on only part-part analysis decreased throughout the teaching experiment as she built a conceptual scheme between how the parts in a probabilistic situation are affected by the quantity of the whole. Tasks involving theoretical probability as well as the use of the pie graph seemed to provide the most beneficial opportunities for her to analyze part-whole relationships. During the pre-interview, the area model represented in the spinner was the only instance where she explicitly referred to the whole pie as representing 100%.

The earliest evidence that she is considering the effect of the whole occurred during the second teaching session. Jasmine thought that it was possible, although very unlikely, to get 10 heads in a row when flipping a coin. When I asked her about the likelihood of getting 10 sixes in a row when rolling a die, the following dialogue ensued.

> J: It could happen, but … you probably won't get that many of the same number. [pause] And you have more choices so it's even more unlikely.
> T: Oh, more choices than what?
> J: More choices than two [she points to some of the coins on the table].
> T: Oh, so yesterday with the coin we had two [choices].
> J: Now we have six choices.

Jasmine's comparison between the coins and dice and the increased number of choices demonstrates her intuitive understanding about the inverse effect of increasing the number of outcomes on lowering the chance of getting all 10 results of the same outcome. Her comparison indicates that she was using the theoretical probability of a single event in her reasoning and that she believed the probability of "6" (1/6) is less than the probability of "heads" (1/2). Although in both experiments, every event has a part of "1," she recognized the die has a larger "whole" than the coin. In addition, she seems to have an intuition that "6, 6" will be even less probable than "heads, heads" and the probability of successive 6's will continue to be lower than the chance for successive "heads" each time another die is rolled or a coin is flipped.

During the second teaching session, it is important to note that with the 2B2W and 5B5W bags, Jasmine recognized that the chance of picking out a black marble was the same in each bag. Although the size of the parts and the whole increased, in this "even" situation, she knew that the same relationship between parts in each bag still held true although the 5B5W bag had many more total marbles. Albeit simple, her recognition that the relationship between the parts was maintained as the whole increased is important as a precursor to her thinking in successive teaching sessions.

Recall the discovery of the “mystery weights” (5-1) during the students’ experimentation with a coin in the fourth teaching session. During that session Joe and Jasmine also investigated the effect of increasing the weight for the tails while keeping the weight for the heads constant at one. Jasmine first used weights of one and two in favor of tails and predicted “it will be pretty close … because two and one are still extremely close together. But not as close as one and one, or two and two, or three and three.” Her comparison of one to two with several equiprobable weights indicates she was doing a part-part comparison and used an additive relationship because “2 and one are still extremely close together.” However, she drew a pie graph prediction of what she expected with 500 trials (left hand picture in Figure 6.2). Her graphical prediction does not reflect the previous analysis of “it will be pretty close.” When Joe asked her to reconcile her graph prediction with her notion of “pretty close” she giggled and replied “more like this [the graph]” but could only respond “I don’t know” when asked why. It is important to note the length of time (about four minutes) she spent drawing her prediction and elaborating the story of the “sky will be eating the gray-haired guy.” Rather than directly using any part-part or part-whole relationship, she seemed to be caught up in her metaphor of the situation and may have lost sight of the original conditions of the task.

After running 500 trials with the 1-2 weights and noting the results in the pie graph (approximately 1/3 gray and 2/3 blue), she said her graph was “not really” like the graph of the results. Joe increased the weight for the tails to three and asked her what she thought would happen with these weights. She drew a pie graph prediction for weights of one to three (middle picture in Figure 6.2) and said the gray slice (representing heads) would get smaller. It seems that she based her notion of "smaller" on her previous prediction rather than the actual results. Notice she again used her metaphor of the blue area representing the "sky" and the gray slice representing a "gray haired guy." This discussion continued as she tried weights of one to four and one to eight and discussed how giving the tails more weight made "the gray haired guy skinnier." Her pie graph prediction for weights one to eight is also shown in Figure 6.2. Although her predictions did not seem to use any consistent additive or multiplicative reasoning, she did recognize and represent the trend of the decreasing chance of getting a head and the increasing difference in experimental results.

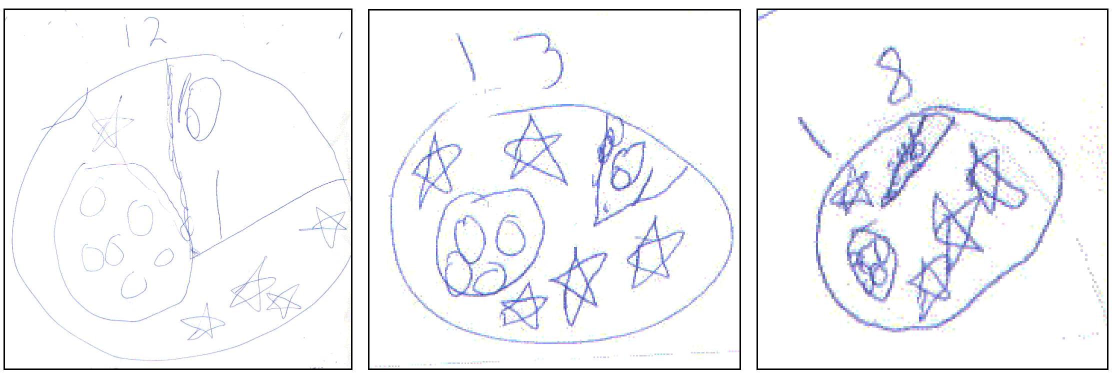


Figure 6.2. Jasmine's pie graph predictions for increasing the weight for "tails."

Throughout the previous episode, Jasmine never explicitly referred to the total amount of weight. However, I conjecture that her work with this task indicates an intuition that the slice representing the part of one (the gray slice) needed to decrease in size because increasing the other part also increased the whole. I conjecture her thinking on this task may demonstrate an initial appreciation for the effect of a larger whole on the parts. However, without her verbalization of this notion, this is a tenuous conjecture.

The other investigation that Jasmine and Joe explored during the fourth teaching session was trying several different weights with a constant additive relationship of eight (e.g., 1 to 9, 2 to 10, 0 to 8, 20 to 28, 100 to 108). Her first instinct was to think that keeping the weights eight apart would result in similar results. In comparing the one to nine and two to 10 weights, she commented that with 500 trials, the number of heads with the one to nine weights were a little smaller than the number of heads with the two to 10 weights. Joe then typed in the weights zero and eight and asked her if those would give the same results as before since the zero and eight differed by eight. Jasmine stated that it wouldn't work "if you used zero because then you would never get any heads." She intuitively knew that a weight of zero made heads an impossible event.

Joe continued by asking her to predict results from 500 trials with weights of 20 and 28. Jasmine said "I think they are going to be far, far apart …well they [weights] are eight apart … but apart doesn't seem to matter." She then used the mouse to predict that the gray slice would stop at about the 30% mark on the pie graph (see reference points in Figure 6.3). Upon seeing that the graph from 500 trials was about 40% gray and 60% blue, she noted "well it was close [to her prediction]."

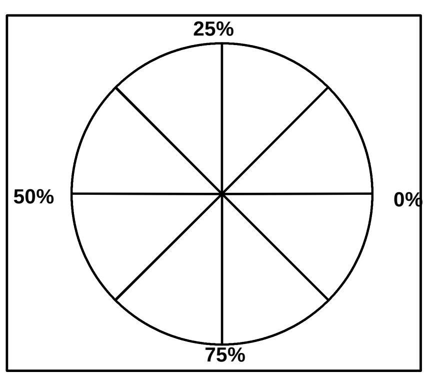


Figure 6.3. Reference pie graph marked with percent locations.

Joe and Jasmine continued to explore weights that differed by eight. For 500 trials with the weights of 100 and 108, Jasmine thought the graph should be similar to the one she got with weights of 20-28 (left graph in Figure 6.4). After running 500 trials, she looked at the graph of her results (middle graph in Figure 6.4) and tried to explain why the graph looked different from the one done with weights of 20-28.

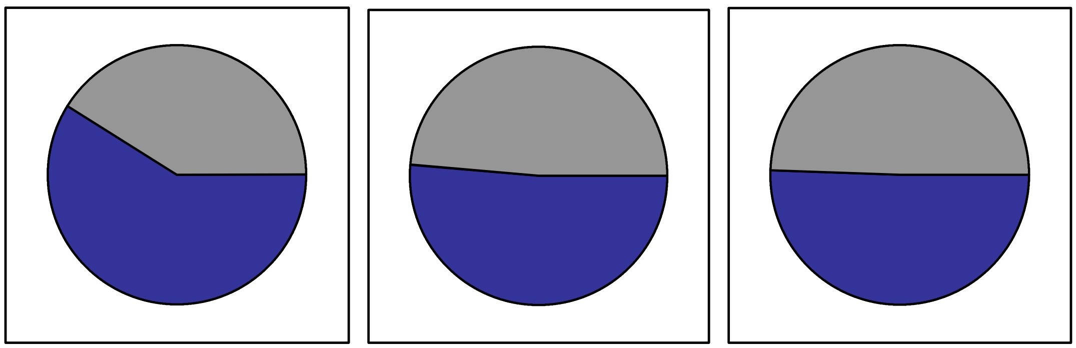


Figure 6.4. Pie graph results from weights of 20-28, 100-108, and 300-308 (left to right).

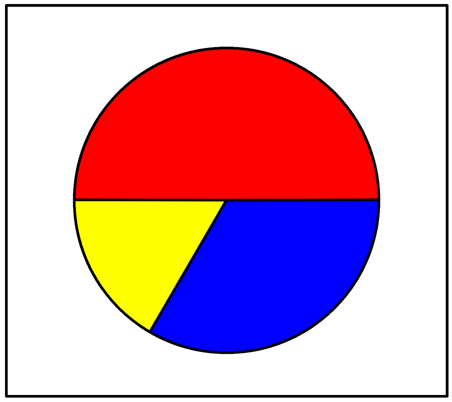
> J: It's still in the corner [sweeps mouse over the top 50% of graph]. But it it's almost out of it.
> T2: Yeah, why do you think?
> J: Well it's a bigger number, its a 100, so it will still be in that corner [the top 50% of the pie graph], but it's going to be farther out [points to area in pie graph close to the 50% mark in the pie graph].
> T2: Why is that?
> J: It's going to be farther out because it's a 100 … but because they aren't equal, it's going to be in this corner [she does a sweeping motion from 0% to 50% of the pie graph].
> T2: Why does a 100 make it come further out [toward 50%]?
> J: Well a 100 is a bigger number…But with a 100 it's farther out but I don't know why though.

Jasmine believed the "bigger number" was contributing to the differences and the trend towards the 50% mark; however, she was not able to reason thoroughly about why the bigger number made a difference. She also indicated that since the weights were not equal, the gray slice was still in the top 50% of the graph. Jasmine wanted to test her theory about bigger numbers and changed the weights to 300 and 308.

> J: They are going to be close around even.
> T2: Oh, why?
> J: Well with just 100 and whatever a 100 makes [108] it was like this. [points to the pie graph showing results from weights of 100-108]. So this one with 300 it may even be just a little out. [used mouse to indicate she expects it to be just slightly above the 50% mark on the pie graph]

The results from the 500 trials with weights of 300-308 are displayed in the right-hand graph in Figure 6.4. Although Jasmine never explicitly referred to the whole, she did recognize that having large parts with the same additive relationship affected the differences in experimental results. She summarizes all the “8 apart” weights by noting that they all have experimental results with gray slices in the top half of the pie and that small weights are in “that corner” near 0% while larger weights are closer to “this corner” near 50%. Her analysis with this task suggests that she is beginning to understand that larger parts have an effect on the relative relationship between the parts.

An important transition in Jasmine’s reasoning from part-part to part-whole occurred during the fifth teaching session. I had given Jasmine a pie graph (see Figure 6.5) and asked her to design weights for three possible outcomes that she thought would give experimental results similar to the pie graph. It appears that this particular circle representation without reference to numerical data, induced appropriate part-whole and multiplicative reasoning, especially with references to 50% as half of a circle.



<u>Figure 6.5</u>. Graph given to Jasmine for her to design appropriate weights.

Jasmine first assigned the whole pie to be worth a total weight of 10. Then she used multiplicative reasoning to determine that one weight needed to be five to represent the red slice since “5 is half of 10.” She used additive reasoning to break up the remaining half into two and three because those parts need to add to five and “one has to be shorter than the other.” However, with a total weight of 10, weights of two and three are the closest whole number approximations for this task. This is the first time Jasmine has explicitly used the whole and connected relationships between the whole and its parts. Up to this point, all the part-part reasoning was done with known weights and resulting pie graphs and numerical data. When asked to reason a *postiori* from a graph, without numerical data, she was able to think of the pie graph as a numerical whole and establish the relationships between the part and whole based on the visual relationships in the circle.

During the sixth teaching session, Jasmine used the pie graph again to construct a direct part-whole relationship. She was experimenting with equally weighted coins and predicted that the results should be “close to even.” After getting results of four heads and six tails, we had the following dialogue.

T: I think that's about as close as we can get to five and five. But this doesn't look very close to a straight line across [I point to the imaginary line representing the 50% mark].
J: Because there are only 10 in the whole thing.
T: Oh, so what does that have to do with anything?
J: So one would only be, only one point would only be about that much. [She uses mouse to draw an outline of the 1/10 slice shown in left-hand pie graph in Figure 6.6] One, let me see … One, two, three, four, five. [While she is counting she moves the mouse pointer five times about evenly spaced counterclockwise starting with the imaginary slice above the 50% mark over to the 0% mark.] See five of those equals, you need to be able to fit five of those [1/10 slice] in here [She moves mouse pointer over the top half of the pie.]. So that would be 10. [She moves mouse pointer over entire pie graph.]

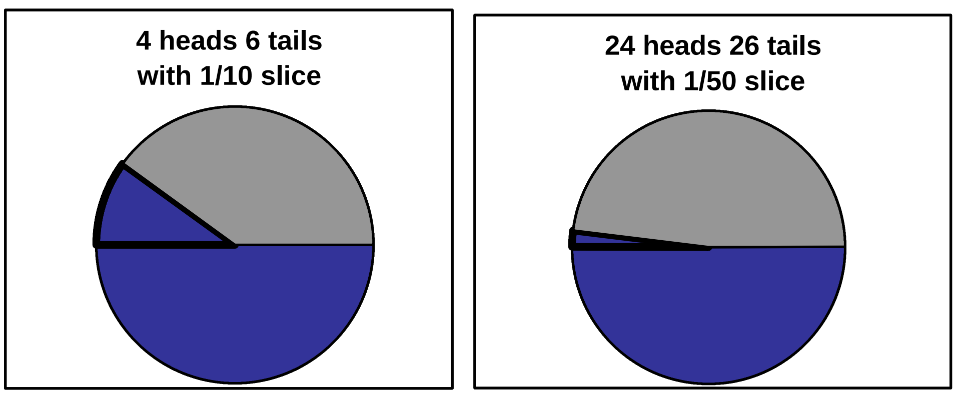


Figure 6.6. Jasmine's use of a reference slice in her part-whole reasoning.

The visual representation of 4-6 results and using 50% as a benchmark prompted Jasmine to construct the unit fraction 1/10 based on the whole pie representing 10. She was also able to iterate the fraction unit five times to construct 50% and 10 times to construct the whole pie, or 100%. Now that she constructed a part-whole relationship based on the number of trials as the whole, I asked her to predict results for 50 trials. She again predicted "close to even" and ran the simulation several times. She eventually got experimental results of 24 heads and 26 tails.

T: 24 and 26. That's pretty close to 25 and 25.
J: Yep.

> T: Now let's talk about the slices again. We are close, we are just one off from being even. And this time we are closer to that half-way mark. Why are we closer?
> J: Because you only need to take one away from this [she draws the imaginary 1/50 slice in the blue area] and add on to this [gray area] and they are even.
> T: Oh. But when we did four and six, then we only needed to take one away. But the pie slice was way up here [pointing to the 40% mark].
> J: Because it was a smaller number. This [pie] is 50. So 25 would be able to fit there [she moves mouse pointer over the gray area].
> T: Oh, 25. If I take one of these little slices right here, I need to get 25 of those in the gray area. Oh, that's pretty good reasoning Jasmine. Now let's go and let's try 100 [trials]. So if we do 100 are our slices getting larger or smaller?
> J: Smaller.

Her reasoning about the slice from the 4-6 results being larger than the slice from the 24-26 results because 10 was a smaller number indicates she is fully considering the effect of the whole on the size of the unit slices. She used the 50% part of the circle in her reasoning as she imagined 25 of the 1/50 slices fitting into the top half of the circle. This visualization with the pie graph provided her with a cognitive tool in constructing each of the unit fractions and to be able to reason that the size of a unit fraction decreases as the whole increases. Her ability to iterate a unit fraction and form larger units (e.g., 1/10 iterated five times forms 5/10 or 50%), and her comparison of the units in terms of how many iterations are needed to make an equivalent unit of 50%, demonstrates her development of a part-whole scheme.

In the seventh teaching session, Jasmine used her part-whole scheme in analyzing a pie graph from 10 trials from an experiment with four outcomes. The results were one die, four blue marbles, and five smiley faces. I focused her attention on the slice representing the one die.

> T: We got one die. Look at how big that slice is [as compared to her prediction of a smaller size slice].
> J; Oh … we were doing 10.

> T: What does that have to do with anything?
> J: Because the whole circle would be that much. For example, this would be 10 [the whole pie] and these would only be five [points to both the 1/10 and 4/10 slices that together are 50% of the circle]. So the one [slice] would have to be bigger.
> T: You are right.
> J: It depends on how big the circle.
> T: You are right. We only have the one slice representing the one here.
> J: If we did it with 100 [trials] and that [die] was only one, then it would be smaller.

She was able to use her part-whole scheme to help analyze the pie graph and also hypothesize the effect of changing the number of trials to 100 on a slice representing "1."

Considering that Jasmine did not use any references to the whole in her analysis of tasks in the pre-interview, she made remarkable progress in developing a part-whole scheme and reasoning about the effect of the whole on its parts. The weight tool and pie graph display were valuable tools in her developmental process. The exploratory nature of the microworld, the tasks she investigated, and the visualization she did with the pie graph all contributed greatly to her part-whole reasoning. As Jasmine was developing a part-whole scheme, she also increasingly used appropriate multiplicative reasoning instead of relying solely on additive reasoning. I will highlight a few instances of her use of additive and multiplicative reasoning in the next section.

Use of Additive and Multiplicative Reasoning

Jasmine used both additive and multiplicative reasoning during the teaching sessions. Some of these instances were discussed with respect to the previous themes. For example, she used additive reasoning in the fourth session when she claimed that weights of one and two would give results "close" to 50%-50% but "not as close as one and one, or two and two, or three and three" because "two and one are still extremely close together." However, she used multiplicative reasoning in the fifth teaching session when, with weights of ½ and ½, she predicted "5 and 5" for 10 trials "because 10, half of 10 is five, and five and five are the same number, and five and five is 10." In this instance, she relied on both "same number" (part-part) and "half" (part-whole) relationships to aid in her multiplicative reasoning. In addition, during the sixth teaching session, her iteration of the slices in the pie graph (1/10 and 1/50) to construct the slice representing 50% demonstrate her use of an additive process to model a multiplicative relationship (e.g., 1/10 + 1/10 + 1/10 + 1/10 + 1/10 = five (1/10) and 5/10 is 50% of 10/10).

Jasmine's multiplicative reasoning seemed to be facilitated by graphical representations; whereas her additive reasoning was usually done with discrete objects (marbles) or numerical representations (weight tool and data table). During the fifth teaching session, I had secretly weighted the baseball and basketball as two and the soccer ball as one. Jasmine's task was to experiment and use any available tool, except the weight tool, to help her determine the secret weights. She decided to run 100 trials and opened the pie graph, bar graph, and data table. Her results are shown in both the pie graph and bar graph in Figure 6.7. Jasmine used the graphs to conjecture a multiplicative relationship in the secret weights.

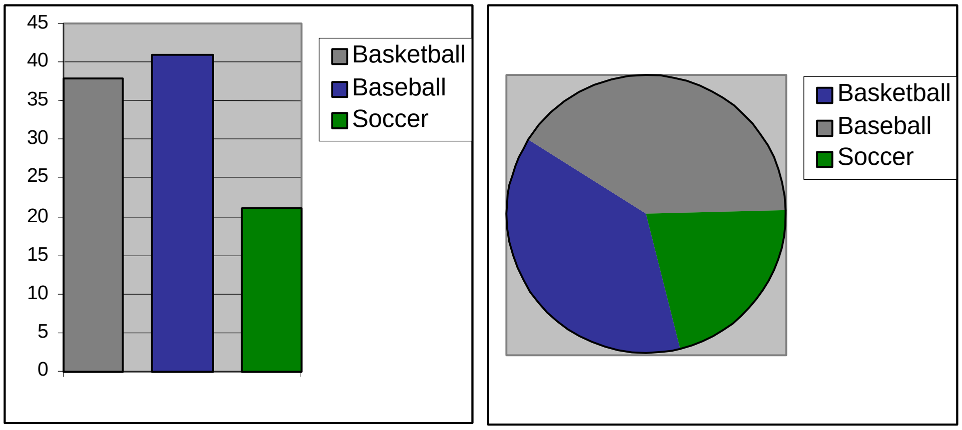


Figure 6.7. Graphical displays from 100 trials with "secret" weights of 2-2-1.

> T: What can you tell about the soccer and the baseball with the bar graph?
> J: These two [baseball and basketball] are staying close. So I think they are either one apart or tied. And this [soccer] is way down, it hasn't been above 50, [looks at data table] see 21. So I think that that [soccer's weight] is around one, two, or three.
> T: Now how does this blue bar compare with let's say the green one?
> J: About half.
> T: It looks like about half of it? Hmm…It does. Does it look that way on the pie graph too?
> J: Yeah.
> T: … So does this green look like it's about half of the blue?
> J: Yeah. I think maybe … I was going to say it was an even number.
> T: What was an even number?
> J: One of the big ones, these two [basketball and baseball] were tied at an even number. And then this [soccer] was half of that even number.
> T: … So is there anything that you can do to test that theory?
> J: Try it again and see if it comes around half. [She clears all trials and runs new set of 100 trials and watches the bar graph during the simulation.] Well it [green bar] is not around half, except it's getting there. [pause] See it's around half. Only it's getting a little up higher. It's still close to half. Oooh, it's 22, one higher [than the last result of 21]. I think that it's umm, I'll say about, those [points to gray and blue bars] could be eight.
> T: Which ones, these two [baseball and basketball]?
> J: Yeah. Those two may be tied at eight and this one [soccer] at four or three. Because that's close to four.

She was able to use a visual measurement of half with the bars and predict weights using a multiplicative relationship based on her "half" observation. She also knew that she needed to use an even number for the larger weights so she could figure out its “half.” Jasmine continued to use multiplicative reasoning to maintain the relationship between the bars for 500 trials.

> T: … What if you ran like 500 times. What do you think would happen?
> J: I think it’s [soccer] going to be around the middle?
> T: What do you mean by around the middle?
> J: Around the middle of whichever one of those two [basketball and baseball] is higher.
> [After 500 trials, she noted that baseballs had slightly more than basketballs but that the soccer balls were about “in the middle” of the baseballs. The only tool she used for this analysis was the bar graph.]
> J: I’m going to change that. I think that the baseball is one ahead of the basketball. And then the soccer ball in the middle of the baseball. So I’d say the baseball is about eight. The basketball is seven. And the soccer is four, or three.

Jasmine used an additive strategy to slightly change the weights after noting a slight difference (9 apart) between the results of basketball and baseball. It seems she was not accounting for the relative difference. Perhaps she would had opted to maintain the multiplicative relationship if she would had used the pie graph at this point.

Jasmine was pleased with her conjecture of when she opened the weight tool and discovered the weights of 2-2-1. At first she used an additive interpretation and noted that they were only “1 apart” but then recognized that “1 is half of 2” and that her original guess of 8-8-4 was “right.” She was able to use multiplicative reasoning to recognize the equivalence in the weights based on her “two of the same and one half” relationship. I then challenged her to create other weights that would maintain the relationship.

> T: So I used weights of two, two, and one. If we clear the weights, could you weight these so that it would be the same as two, two, one but with different numbers? [pause] What do you think? What numbers would you use?
> J: I’m thinking. I know. [types in three for the basketball] Oh no, that’s not good. My lucky number is an odd number. That’s not very lucky.

> T: What's your lucky number?
> J: Three.
> T: So if you put a three here [basketball], and a three here [baseball]. What do you have to put here [soccer]?
> J: One and a half.
> [I explain that the weight tool can not use fractions.]
> T: So if we can't use three what else do you want to use?
> J: Oops. I forgot.
> T: That's okay. We want to have the same relationship as two, two, one.
> T: So what's your strategy here?
> J: Twenty, twenty, ten.
> T: …And why is that the same as two, two, one?
> J: It's the same amount with, actually it doesn't have to be with the same balls. But two are the same and one that is half to the ball. [She changed the weights so soccer and basketball both had 20 and baseball had 10.]
> T: So now if we ran it this way, what would you expect to happen?
> J: The soccer ball, and the basketball would be close together. And this would be a way.
> T: How far?
> J: About half.

She was able to create two different examples of equivalent weights. With the first set, she wanted to use an odd number and recognized that it would not be evenly divisible by 2; although she knew that half of three would be 1½. (Note: I have adapted the weight tool to take decimal values as weights because of her investigation.) She also used 20-20-10 and was able to predict that the experimental results would reflect the "same" and "half" relationships.

During the seventh teaching session, Jasmine predicted that four outcomes with equal weights would result in about "quarters" in the pie graph. When I asked her to explain what a quarter was, she drew a circle, drew a horizontal "half" line, then drew a line to mark half of the top half and again for the bottom half. As she drew the quarter markings, she explained, "it's half of a half, one out of four [pointing to one quarter slice] is half of two out of four [pointing to both quarter slices that constitute a half]… because two is half of four and that's [the same as] one out of two." Her explanation of a quarter shows evidence that by using a circle, she was able to use multiplicative reasoning to construct the part-whole relationships and recognize the equivalence of 2/4 and1/2. Again, this is evidence that a circle is a powerful cognitive reference for her understanding of half.

The last episode that demonstrates her use of both additive and multiplicative reasoning occurred during the seventh teaching session when she was experimenting with a 3B1W and then a 6B2W bag of marbles. At first she thought the 3B1W bag gave the best chance for both the black and white marble because there "are less in the bag." After running several sets of 100 trials with both bags and recording the results, she conjectured that the chances were the same in both bags because the blacks were always "around 70-80" and the whites were "around 20-30." However, she was not initially able to support her "same chance" conjecture. To help her consider the relationship between the bags, I asked her to describe the chance for picking out a white marble in each bag. She used a part-whole statement to note "one out of four" and "two out of eight" and wrote each of those statements as fractions and stared at them for about 20 seconds (see Figure 6.8). She then widened her eyes, smiled at me and explained "one plus one equals two and four plus four equals eight" as she wrote the corresponding notations in between the two fractions (see Figure 6.8).

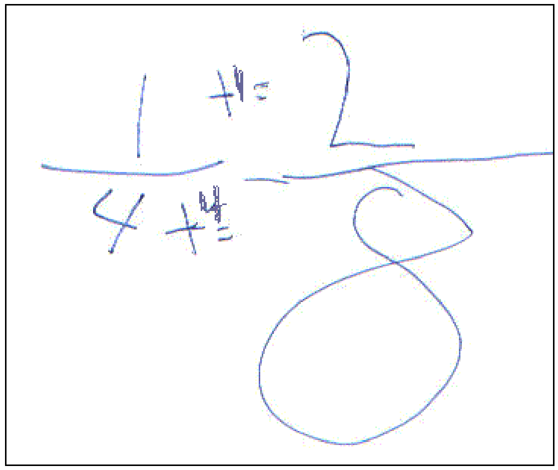


Figure 6.8. Jasmine’s numerical explanation of the equivalence of 1/4 and 2/8.

Although she used an additive process to equate the two fractions, she did not use a constant addition. Instead, she maintained the multiplicative relationship by adding each part to itself to double the quantities in each part. She used the same reasoning for the fractions describing the chances of picking a black marble (3/4 and 6/8) since “ three plus three is 6.”

Although Jasmine used additive reasoning throughout the teaching experiments, sometimes inappropriately, she made substantial progress in developing her multiplicative reasoning. Her multiplicative structures were based on additive processes and facilitated by references to 50%, “half” and parts of circle. She continues to use both additive and multiplicative reasoning in her work with the “twice as likely” task. The description of her investigative approach to this task provides further evidence of her development of probabilistic reasoning across all the four key themes and demonstrates how her approach to this task is similar and dissimilar to the approaches of Carmella and Amanda.

Jasmine's Investigation of "Twice as Likely"

Jasmine's work with this task occurred during the sixth teaching session. She designed an experiment with two outcomes, the sun and lightening bolt. I asked Jasmine to use the weight tool to design the chances so "the sun is twice as likely to happen than the lightening bolt." Her first instinct was to use an additive relationship and she entered weights of three and one.

> T: Let's go ahead and click on the button [in the weight tool] that says Show As Fraction. And tell me what these fractions mean. [the weights are now displayed as3/4 and 1/4]
> J: They are four days. There are three sunny days and one lightening day. So there's three out of four chance in the next four days that there's going to be a sunny day. And one out of four chance that there's going to be a miserable day. A thunderstorm day.
> T: And so does this mean that sun is twice as likely to happen as the lightening bolt?
> J: Yes.

Although she could interpret the weights of ¾ and ¼, the numerical part-whole relationships did not help her to recognize the relationship between the weights. With these weights, for 100 trials she predicted "it will probably be more sunny days" and predicted 44 lightening and 56 suns. Her predication did not reflect a "twice as likely" relationship and appears to more of an estimation based on the additive relationship and that suns have more weight. However, once she ran the 100 trials, analyzing the pie graph, numerical data, and the assigned weights prompted her to develop a better understanding the "twice as likely" relationship.

> J: Oh, we got 76 and 25
> T: What do you think about your results here?
> J: The sun was ahead. By a little bit too much [pause] Not really though.
> T: So let's open up that Weight Tool again, so we can see our weights.
> J: [She opens the weight tool.] Oh, [squealing] I forgot that this is supposed to be twice as what it says, it has to be twice as many. So this

[she points to the 76 suns in the data table and then the gray area representing the sun in the pie graph.] would have been twice as many.
T: Oh, is 76 twice as many as 24?
J: I don't know. It's close.
T: What would be twice 24?
[She adds 24 + 24 to get 48.]
T: So there would be 48. Seventy-six is a bit larger than that. Don't you think? [she nods "yes"] So what do you think about these weights? [I open the weight tool] Three out of four, and one out of four. I'm going to move these down so we can see our pie graph here …Well let's see. We've got three out of four, and one out of four. And this is the results that we got.
J: I know what I'll do. One, … [types in weights of one and 99]
T: Ninety-nine. So is the sun twice as likely as the lightening bolt?
J: No. [giggles]
T: No. Why not?
J: Because one two times is not 99.
T: Oh. So how could we make this so that the sun is twice as likely?
J: [She enters weights of 48 and 24] That's twice.

The experimental results gave Jasmine a basis for exploring the "twice" relationship. She recognized that 76 was more than twice 24 and knew she needed to change the weights to get a "twice" relationship. Although she playfully entered one and 99, she knew those weights did not satisfy the relationship because "1 two times is not 99." Since she had already established a multiplicative relationship, she merely used 24 and 48 as her weights to satisfy the "twice" relationship. She then made the connection between repeated addition and the multiplicative relationship by using additive iterations of 24 to justify why 48/72 and 24/72 were appropriate.

T: 48 and 24. So what do these number mean here? 48 over 72?
J: Forty-eight over 72 …Oh, there are 72 suns and lightening bolts put in the box. Forty-eight of them are suns. Twenty-four of them are lightening bolts. And children put in that many because they think out of 72 days there are going to be 48 and 24, there are going to be 48 sunny days and 24 thundering days.
T: … Is the sun twice as likely as the lightening? [She nods a strong "yes".] And how could we test it?
J: 24 and 24.

T: Yeah? What?
J: 24. [She is looking up at the ceiling while thinking.]
T: There is paper there if you need it. What are you thinking about?
J: See 24 plus 24 plus 24 equals [writes the addition vertically on paper]
T: And how come you are adding three of them together?
J: I don't know … I'm just going to find out why. [pause] Seventy-two.
T: 72. And that's the number that's on the bottom [in the weight tool].
J: I know. [smiling]
T: So you took 24 and you added it together three times and you got the number that's on the bottom.
J: I know!
T: Why? Really, why is it supposed to do that?
J: Because 24 and 24 are 48. But with another 24 it's supposed to be 72.
T: Oh, I see. So we've got three sets of 24 here. Two of which are suns.
J: Yeah!
T: And one is lightening bolts.

Using an iterative scheme, she established that there are three parts in a "twice as likely" relationship, with one element having two parts and the other one part. Her thinking demonstrates how she used additive reasoning to construct the relationship needed for her multiplicative reasoning. This also provides evidence of her part-part reasoning transitioning into part-whole reasoning.

As our investigation continued, I asked her to predict what the pie graph would look like if we ran 100 trials. She drew a circle then drew the "half" line but commented "oops I wasn't supposed to do a line in the middle … because it would not be half and half if they are not even." She then drew a line above the "half" line at about the 30%-40% point. She marked the smaller area as the lightening bolts and the larger area as suns. Although she could not give any justification beyond "that's what I think," her use of 50% as a reference was consistent with other work she has done with a pie graph. After running 100 trials, the resulting graph looked similar to her predication. I asked her if she thought the gray area was twice as big as the blue area.

J: I can’t really tell. [She measures the width of the blue slice with her fingers and keeping her fingers at that width, iterates that width twice in the gray area] About.
T: Yeah? Can you tell from the numbers? [66 and 34]
J: Yeah. Thirty-four. But it probably won’t be. Let me see something. [she adds 34+34+34] Okay, it should be 72. No way.
T: What do you mean no way?
J: It doesn’t equal 72 so it’s not…
T: So you took 34 and added it…
J: Yeah because 24, 24, 24 equals 72. And 72 was the biggest number. It has to be something, something, something, and equals.
T: But how many times did we run it?
J: A 100 …. oh, they are two off.
T: What do you mean they are two off?
J: It equaled a 102.
T: Oh. I see. So it looks like, if you take 34 and just add one more 34 to it, what do you get?
J: 68.
T: And we got 66.
J: Yeah, only two off.

Jasmine used her iterative strategy with both the pie graph and numerical results. Although she initially was using the total weight of 72 as the “whole” for the experimental results, when I focused her on the 100 trials, she was able to look for the “twice” relationship by adding 34 three times. She then noted that the results were “2 off” since three 34’s are 102 and two 34’s are 68 rather than a 66.

For the final task in this investigation, I asked Jasmine if 48 and 24 were the only numbers she could use in the weight tool to have twice as likely. She quickly replied “no, no, no” and entered weights of 200 and 100 because if the lightening was 100 the suns had to be “another 100, so that’s 200…a 100 and 100 and 100 is 300.” I asked her to compare these weights with the previous ones of 48/72 and 24/72.

T: So is this the same then as what we had before? We had 48 out of 72?
J: [nods “yes”] Different numbers though.
T: And 24 and 72? Is the chance the same?
J: Yeah.

T: Yeah? Why is it the same?
J: Well because there's still twice as many chances.

Jasmine recognized that maintaining the multiplicative relationship made the weights equivalent. She ran 100 trials and predicted that we should get close to the same results as before. She actually got 66 suns and 34 lightening bolts again. She noted that was "because there are the exact same number of chances… it may not be the same number but it will be close, probably." It seems that experimental evidence helped her justify that the chances were the same and that although she got the exact numbers this time, she would only expect the results to "be close, probably." This statement hints at her application of the EOP with experiments based on several sets of equivalent weights.

Jasmine's reasoning with the "twice as likely" task gives further evidence of her transition from part-part to part-whole reasoning, and her developing multiplicative reasoning based on additive iterations. Although she began with an additive approach to this task, her use of the microworld tools facilitated her understanding of the multiplicative structure embedded in the problem. She used the weight tool, data table, and pie graph together to interpret and analyze the "twice" relationship. She connected the 2:1 ratio with a 2/3-1/3 relationship and was able to use that discovery to create the 200 to 100 weights and justify her choice by adding 100 three times to get 300.

<u>Summary of Meaning-Making Activity in the Microworld</u>

Jasmine was very enthusiastic, highly motivated, and engaged in meaning-making activity during the teaching sessions. She quickly learned how to use the various tools in the microworld (e.g., stacking columns, graphs, data table, weight tool) to design and run experiments, and analyze the results in a variety of formats. She often used the multiple representations simultaneously in her analysis and recognized that various representations could help her make sense of the data (e.g., “we have four ways to see everything”). Jasmine often used the tools to model real world problems of interest to her (e.g., her simulation of the bingo barrel) and playfully described displays in the microworld in an imaginative way (e.g., “the gray is eating a piece of the sky”). For the most part, her playful orientation to the tasks helped to sustain her engagement in the tasks and gave her a creative outlet for describing her mathematical ideas in a literary manner. However, on a few occasions, she did lose sight of the original goals of a task because she was adding a playful orientation to the task.

The vignettes described within the four themes – theoretical probability, EOP, part-whole relationships, additive and multiplicative reasoning—capture the essence of Jasmine’s meaning-making activity, and suggest that her use of the tools in the microworld facilitated her mathematical thinking and further development of probabilistic reasoning. Together, the four themes were interconnected and demonstrate that Jasmine’s meaning-making activities relied on many aspects of probabilistic thinking. Her increased use of theoretical probability and her developing understanding of the EOP were influenced by her eventual transition from part-part to part-whole reasoning and her developing multiplicative reasoning.

Jasmine based a lot of her reasoning throughout the teaching sessions on the visualization of the pie graph and references to 50% and half. A circle and her scheme of “half” seemed to be useful thinking and building block tools for her transition to part-whole reasoning and development of multiplicative schemes. Jasmine’s meaning-making activity demonstrated that she was easily able to transition to the digital environment in the microworld and used the tools to her advantage. She was empowered by the open-ended nature of the microworld and used the tools to continually test her conjectures. However, it is important to note that several times during the teaching sessions, she wanted to conduct parallel physical experiments to compare her results (e.g., picking marbles out of a real bag). It seemed very important to her to connect the microworld work with real-world activities.

Post-Interview Analysis

Jasmine's post-interview was held 3 1/2 weeks after her last individual session (see Appendix G for post-interview protocol). Her work during the interview and my analysis of her responses are organized by the different tasks.

Cubes in a Bucket

Similar to the pre-interview, I asked Jasmine a series of questions using a bucket containing six green, four red, and two yellow cubes. She used strict part-part reasoning to determine that green was most likely "because there are more" and yellow was least likely "because there are less." After I randomly chose a green cube and then replaced the cube in the bucket, I asked Jasmine if I was more or less likely to pick another green next time. She said, "you are still more likely because there's still the same amount." She did not let a previous event influence her analysis of the chance for picking a green and noted that the contents of the bag had not changed. In addition, when I asked her to use numbers to describe the chances, she easily used part-whole statements like “6 out of 12.”

The task continued as I removed two green cubes one-by-one. She again used part-part reasoning to note that after one green cube was removed, green was still more likely “by a little bit because there are still more, but only one more than the rest.” After removing the next green cube she explained “then the chance of green is tied with the red.” Her part-part reasoning on this task was appropriate for answering the questions.

I brought the bucket back out later in the interview and asked Jasmine to compare the chance of getting a green cube now (4 green, four red, two yellow) to before (6 green, four red, two yellow).

> T: Has the chance of picking a green cube changed since when I took those ones out?
> J: Yes.
> T: Yes. How has it changed?
> J: It’s now less likely to pick a green cube out.
> T: Oh it is. And why is it less likely?
> J: Because there are less green cubes.
> T: Less green cubes. What about the chance of picking a red from when they were in here to when they were out? [I model placing the two green cubes back in the bucket and then removing them.]
> J: It hasn’t changed.
> T: The red hasn’t changed?
> J: Oh, it has!
> T: It has? Why has it changed?
> J: Because it used to be little but now it’s tied with the green.
> T: Oh, so there’s four reds and four greens in there.
> J: Yep, yep, yep …
> T: So how do you think the yellow has changed?
> J: It has less against that are in the bucket so it’s a little bit more likely to get picked.

Although she does not explicitly use part-whole reasoning, her part-part reasoning accounts for the decrease in the green “part” and subsequent decrease in the total number of cubes in the bucket (e.g., yellow has “less against that are in the bucket”). Even though the yellow and red “parts” did not change, after my leading question (i.e., the red hasn’t changed?”) she was able to recognize that the chance of picking each color increased by removing the two green cubes.

Coin Tosses

Two types of tasks involving coin tosses were used in the post-interview. The first task assessed concepts of independence while the second task assessed concepts of fairness and the law of large numbers from a frequentist perspective.

Tossing a coin six times. I showed Jasmine the four possible strings of results from flipping a coin six times (HHHHTT, THHTHT, THTTTH, HTHTHT) and asked whether any of the results are more likely to happen than the others. At first Jasmine replied “no” but then said she thought the last one might be less likely since it is in a pattern. However she noted that the other three lists “all have the same chance.” It seems that her understanding of independence has improved, but her conception of random as “mixed up” conflicts with the last string in the HTHTHT pattern. This seems too orderly and to her is less likely than the other “more mixed up” strings of six. She continued a similar line of reasoning when shown the next set of results (HHHTTT, HHHHHH, THTHTH, HTHTHT). She thought that all of these strings were unlikely to happen because they all had some sort of pattern in them. In addition, she noted that HHHHHH was not likely “but the higher you go with the number that you do it, it gets more and more unlikely.” She did not employ any notions of independence in her reasoning but still showed evidence that in a random situation, she expects results to look more random than they do "in a pattern."

For the final question about independence, I asked Jasmine if, after flipping a coin and getting the results HTHHHH, I was more likely to get a heads or tails on the next flip. She promptly stated "no" because "just because you got more heads and less tails, they are still a head on one side and a tail on the other." Her response demonstrates some use of independence and understanding that the theoretical probability did not change on the coin.

Is this coin fair? The intent of this task was to assess whether Jasmine could reason from a frequentist perspective about the fairness (i.e., equiprobability) of a coin when given experimental results. In the first situation, I told her that I flipped a coin 10 times and got eight heads and two tails. When I asked her if she could tell if the coin was fair from the data, she immediately said "no" then continued to justify her reasoning.

> J: Well it would probably be fair.
> T: It would probably be fair?
> J: Yep.
> T: And why do you think that I couldn't tell whether it was fair or not?
> J: It's still possible that it could come out eight and two … And it could be unfair because maybe it's magic. But it doesn't want to land on heads all the time or tails all the time. But it always might get more heads. Because it doesn't want people getting too suspicious.
> T: Oh, I see. But if it did have that magic and it would land on heads more often would that be fair?
> J: Huh huh. [nods "yes"]

In the second situation, I told her that I flipped a coin 100 times and got 41 heads and 59 tails. She again noted that she couldn't tell if the coin was fair but that these results "are close together" and that "half of a 100 is 50 and those are close to 50 ... and it would probably be close together if it's something you can split in half and if you can't split it in half then it would be around the 50." She used "half" as an indicator of what she would expect for results from a fair coin. However, she also maintained that these results were not an indicator of fairness. I then asked her what she thought about getting 175 heads and 325 tails with 500 trials. She immediately said "I wonder if that would really happen" but still maintained that she could not tell if the coin was fair although it was "really unlikely" but "could happen."

> T: So if I have a fair coin this [125 heads and 375 tails] could happen?
> J: Yeah.
> T: So all of this can't really tell us anything about the fairness of the coin? What if I did it a 1,000 times? And if I did it a 1,000 times and what if I got 900 tails and only a 100 heads?
> J: That doesn't tell you anything.
> T: And so if the coin was fair, and we do it a 1,000 times. What kind of results do you think we might get?
> J: I don't know … Something around the middle ... But it doesn't HAVE to be.

It seems that her understanding of randomness includes a strong notion that "anything is possible" and that she doesn't feel this large amount of data can provide information about the fairness of the coin. Her response does not show evidence of any thinking about the law of large numbers. Her experiences with the EOP in the microworld do not appear to be transferring to this situation. This may be due to her inability to fathom the possibility that a real coin with both a head and a tail could actually be unfair.

<u>Marbles in a Bag</u>

As in the pre-interview, I presented Jasmine with pictures of four pairs of bags containing black and clear marbles. Two of the pairs were in proportion to each other while two pairs were not proportional. When presented with each pair, I asked Jasmine to determine which bag she would prefer to pick from, or if it mattered which bag, if she wanted to choose a black marble.

When presented with Bag A (3B3C) and Bag B (1B1C), she immediately recognized the equivalent chances for picking out a black marble.

> J: It doesn't really matter … because there are three black and three whites in Bag A and one black and one white in Bag B and it's the same amount … if I put my hand in there, in Bag B they have the same chance. And over here [bag A] they would have the same chance.

The one-to-one relationship between the black and white marbles in each bag seems to help Jasmine justify the equivalent chances. Although she thinks about the bags experimentally, her reasoning is entirely based on part-part relationships.

When shown Bag C (3B1C) and Bag D (5B2C), Jasmine immediately chose Bag C based on the small number of undesired events (white).

> J: [bag C] because there are four in all. And there's only one that's not a black one. And over here [bag D] there are two [white marbles]. There are more black marbles, but there are two of them.
> T: Oh, and so having the two against it makes a difference?
> J: Uh huh.
> T: So you want Bag C. How would you describe with numbers the chance of picking a black out of Bag C?
> J: Three out of four.
> T: And what about this one over here?
> J: Five out of seven.
> T: Five out of seven. And so you think the three out of four is going to be better?
> J: Yeah.

Although a 3/4 chance is better than a 5/7 chance, Jasmine's justification is solely based on a part-part analysis in favor of a small number of white marbles and does not rely on any multiplicative or proportional reasoning. However, it is important to note that a comparison of ¾ and 5/7 is a difficult task for many students, let alone a fourth grade student.

The next two bags presented, bag E (2B1C) and bag F (4B2C), were in proportion to each other. After studying the bags for about 20 seconds, she chose bag E as having the better chance for picking out a black marble for the "same reason" that she used with the last pair. Although bag F had "more" black marbles it also had "two [white marbles] against." Even when she stated the chances in a part-whole manner (2/3 and 4/6) she did not employ any multiplicative reasoning to compare the chances. It seems that her part-part reasoning with respect to the additive difference between parts superseded any part-whole analysis that she began to use during the teaching sessions.

For the last pair of bags, bag G (2B3C) and bag H (5B6C), Jasmine studied the bags, counting each of the parts separately before making any verbal comments. She then again employed her part-part difference strategy and said she was having difficulty choosing the better bag because "they each have one more than their number against it so I think it's an even chance." In this case, I forgot to ask her to state the chances in a part-whole manner. She never considered the whole in her initial analysis, and I am uncertain if stating the chances as part-whole would have made her rethink her decision. Recall that using part-whole statements did not induce any perturbations in her comparison of bags E and F.

Constructing Sample Space and Theoretical Probability

The only sample space question used in the post-interview was for a three-event experiment. The context for this task was a family with three children, ages 9, 5, and 3. I asked Jasmine to list all possible arrangements of boys and girls with respect to their ages. She began using a systematic flipping strategy (see Figure 6.9) to help her find all possible arrangements (as she did in the pre-interview). Once she recorded an arrangement, she would “flip” each item in the list to obtain another arrangement (e.g., BGB, GBG). However, she only initially found the first six possibilities and only later added the last two arrangements (GBB, BGG).

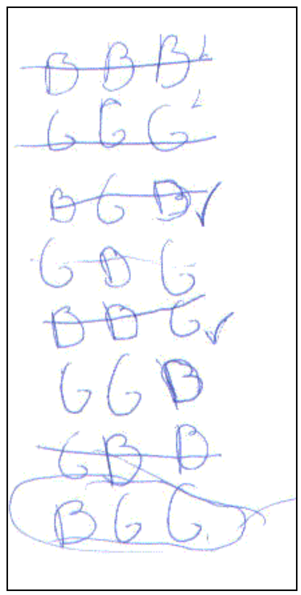


Figure 6.9. Jasmine’s list of all possible family arrangements.

After Jasmine listed the first six possibilities, I asked her whether BGB was different than or the same as BBG. She claimed they were different because “these ones [points to first element in list] are the oldest, these [second element] are the middle and these [third element] are the youngest. That [BGB] means the oldest would be a boy, the middle would be a girl, and the youngest would be a boy.” Her reasoning about the order mattering was tied directly to the birth order in the problem. She recognized why order mattered in this situation.

To assess her ability to determine theoretical probabilities from a sample space, I asked her a series of questions regarding the chance of the actual family arrangement being certain arrangements or combinations of boys and girls. At first I asked her if any of the family arrangements were more or less likely to happen than the others. She immediately said “no” but then changed her mind and noted that perhaps the BBB and GGG arrangements were less likely although “it’s not really too less likely, it’s just a little bit because it’s a small number [of kids].” However when I asked her what the chance was for having a family of all boys, she easily said “one of six because there’s one BBB out of all the others.” She also said that GGB would have a “1 out of 6” chance and that the chances must then be the same; although her expression indicates she was not positive about this conclusion and she could not explain her thoughts any further. When I asked her about the chance for having a family with two boys and one girl in any order, she found the two instances on her paper and remarked “2 out of 6” and continued “and with two girls and one boy it’s two out of six also.” Although she did not have the complete list of all arrangements at this point, she was able to use the sample space to construct the theoretical probability of both ordered and unordered families.

With several attempts to get her to try to convince me that she had all the possible arrangements, she kept stating that she did not know how to convince me but she wanted to eagerly know what the “answer” was so she could see if that was on her list. She seemed tired of explaining her thinking (this is about 25 minutes into the interview) and only wanted to check if she was right. I asked her if the family could have a girl as the oldest and then two boys. She looked at her list, said “oh, I missed that one” added GBB to her list and then immediately reversed it and wrote BGG. She then noted that all the chances now changed to “1 out of 8.” She also note that chance for the combination of two girls and a boy was now “3 out of 8.”

Jasmine was able to use a strategy in making her initial list of six possibilities but was not able to use any justifications to convince me that she had found all possible ways. This shows that she could only employ limited combinatoric reasoning but that she was able to construct and compare the chances of arrangements by using a part-whole scheme. She easily grouped the unordered arrangements and noted the equiprobability of two boys and one girl, and two girls and one boy. Although her combinatoric reasoning was weak, her use of the sample space in determining theoretical probabilities was strong.

Using Results to Design Experiment

The next two tasks were used to assess her ability to interpret and use information from both a pie graph and bar graph. I told Jasmine that I had designed a bag of marbles in the microworld and ran an experiment. I showed her a graph of experimental results and asked what she could tell me about the bag of marbles.

Reasoning from a pie graph. For the first task, I showed Jasmine the pie graph in Figure 6.10 (left-hand picture) and asked her if she could tell me how many times I ran the experiment. She noted that I could have “run it a lot, a lot of times to get this because you wanted to get this [pie graph] or maybe you just ran it once.” She clarified that “once” was not just one marble, but maybe “10, once.” She thought about the task as it could be done in the microworld and the capability of hitting the run button once to simulate 10 trials. Her response also indicates that I may have had to run the experiment many times to get that exact graph if that was what I was expecting. This response hints at her understanding of likelihood of getting the exact results based on the chances in any given set of trials.

Figure 6.10. Given pie graph and Jasmine's drawing of a possible bag of marbles.

Jasmine described the contents of the bag as having three different colors with "more reds and there were about the same green and yellows." She drew the contents of the bag (right-hand side of Figure 6.10) and explained her reasoning.

> T: Why did you design it that way?
> J: I did it that way because the circle is six, three and three equals six, and six and six is 12. So the circle would equal 12. This whole circle. And these are half [She traces the 50% line on the graph]. So half of 12 is six and so one of these was in the middle, so it would have to be half of whatever my number is, and half of 12 is six. So this part would have to be six, and then I already made six so I have to do half of six.

She used appropriate proportional reasoning to complete this task by assigning the whole pie to be worth 12 and reasoning about "half" of the whole and then "half" of that number. Again, this task illustrates how powerful the reference to a circle and half is prompting Jasmine to use appropriate part-whole and proportional reasoning.

Reasoning from a bar graph. The same questions were posed when I showed Jasmine the bar graph on the right-hand side in Figure 6.11. She easily used the scale to determine that I ran the experiment 1,000 times since that was "500 yellow, 400 red, 100 blue all together." She then noted that the bag contained "probably more yellows, less blue, and kind of in the middle green." Her informal quantitative estimation of the contents suggests she used a part-part difference strategy to estimate the quantity of each color.

Figure 6.11. Given bar graph and Jasmine's drawing of a possible bag of "marbles." (Pumpkins represent yellow, black culdrons represent green, and the tree represents blue.)

As you can see in her drawing, Jasmine creatively used Halloween and Christmas symbols to represent the marbles. When I asked her to explain why she drew the number of each item, she again used a part-part difference strategy.

> J: Well on here [the graph] it's five. And four, well if we were just making it five, four and one and not in the hundreds, then four is one below five. And one is only one. So I only did one of those [tree to represent blue]. Ten pumpkins. And nine is one below 10 so I made nine cauldrons, and one [blue] is one. So I made one Christmas tree.

Although she was able to simplify the numbers on the scale to 5, 4, and 1, she did not maintain any proportionality when increasing the number of yellow (pumpkins) to 10.

Instead, she kept the “one apart” relationship between five and four to make 10 and nine and then chose to keep the blue (trees) at one rather than increasing that number as well. She did not use any multiplicative reasoning like she used with the pie graph. The bar graph and the more difficult relationship between 5, 4, and 1 (although five is half of 10) seemed to prompt Jasmine to approach the task with additive reasoning.

100 Gumballs

The gumball task used in the post-interview was similar to the one used in the pre-interview but contained 30 yellow, 60 blue, and 10 red gumballs. I first asked Jasmine how many red gumballs she would predict if someone picked out 10 gumballs. She initially said “not many” then “2 because there’s only 10 in there and that’s against 60 and 30…actually it’s going to be one, or zero, no there were only one.” She then said there would be “a lot” of yellows, “seven,” and “three” blues, but then quickly changed her mind and said “opps, six blue, three yellow, and one red.” She justified her choice because “there are more blues, there are medium yellow, so it’s a medium number, and there’s less red so there’s a less number.” She also noted that her answer was just like the problem if you “take away the zeros.” It appears that she used informal quantitative reasoning to begin with but may have employed naïve proportional reasoning for her final answer. Regardless, her thinking on this task demonstrates that she certainly based her reasoning on the gumball distribution in the machine.

Spinner Game

The last task in the post-interview was similar to the spinner game used in the pre-interview; however, different spinners were used. With the first spinner (1/4 blue and ¾ red), Jasmine chose the red area "because it's more." We then had an interesting discussion about the fairness of the game.

> T: Is this a fair game?
> J: Yeah.
> T: It is?
> J: [Looks at spinner] No!
> T: It's not. Why isn't it fair?
> J: Well it's kind of fair, but....
> T: So tell me about the fairness of the game.
> J: It's not fair because there's more red.
> T: There is more red, and so what does that mean as far as the fairness?
> J: Oh, it's not fair because of the red.
> T: But why does having more red make it not fair?
> J: Because the blue is less and behind.
> T: The blue is less and behind. Does that make it not fair?
> J: Because there's not much blue. But it is fair because [pause] did you make the game up yourself?
> T: Did I make the game up myself? Yeah.
> J: Because it's for teaching children and stuff. And that's why you did it.
> T: Okay. So it's fair?
> J: In a way.
> T: In a way. What's fair about it?
> J: I don't know but if it wasn't fair you wouldn't be playing it with me.
> T: (We both laugh.) You think I would never play an unfair game?
> J: No, I think you would. I just think it's kind of unfair.

Although her first instinct was that the game was fair, upon my reaction (i.e., "it is?), she quickly reassessed her judgment based on the unequal areas on the spinner. However, she thought that the game might be fair merely because I had made it up and I was willing to play the game with her. Her thoughts about the fairness of the game indicate that she could not imagine why someone, who presumably knows something about probability, would make up a game that was unfair and agree to play it if the chance of winning was unfavorable. Her subjective opinion was certainly reasonable and should be considered an equally valid use of probabilistic reasoning as her objective analysis based on the areas on the spinner.

I showed her spinner B with eight equal sectors, two of which were blue, and six were red. She immediately recognized the equivalence and used a measuring technique with her fingers to measure one of the red sectors in spinner B and iterate that measurement six times in the red area of spinner A. She further noted that it did not matter which spinner we used because the chances were the same. I then asked her to use numbers to describe the chance of landing on blue in both spinners. For spinner B, she counted all the sectors and announced “two out of eight” and then said “one out of two” for spinner A because “there are two things and there’s one [pointing to blue area].” Although she correctly used part-whole language for spinner B, her use of one out of two for spinner A indicates that she does not have a strong understanding that each part must be equal to establish a part-whole relationship.

She continued to claim that ½ and 2/8 were equal because of the visual relationship between the spinners. I wrote her chance statements as fractions on a piece of paper to see if she would recognize “1 out of 2” as one-half. Even when I asked her to focus on the numbers, she kept looking at the spinners and telling me they were the same. She then spontaneously drew two short segments and then eight short segments underneath those two (see Figure 6.12, a recreation of her original work).

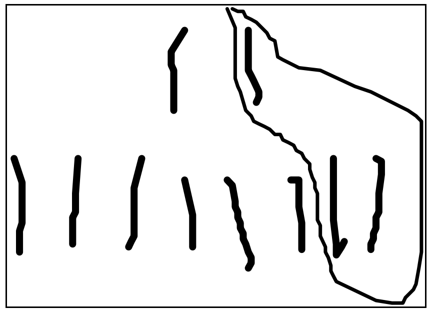


Figure 6.12. Jasmine's illustration of her comparison of 1/2 and 2/8.

She used the stick drawing to illustrate her thinking about why the two fractions were equivalent. She related the sticks to the spinners in her explanation.

> J: They [the eight sticks and eight sectors in spinner B] are smaller because the larger the number [of sectors] the smaller it [sectors] gets. So these [8 sticks] are smaller. That means the pieces [of the pie] are smaller and so two of these [from the eight sticks] would equal one of those [from the two sticks]. [She draws a circle around two of the eight sticks and one of the two sticks.] So that equals that and all of these [sweeps hand over remaining six sticks] equals that [points to one stick on top].

Jasmine's descriptions indicates she was thinking about the effect of a larger number in the whole on decreasing the size of the sectors. She also was able to establish the 2:1 ratio between spinner B and spinner A. However, she did not extend the constant ratio when she said the other one "stick" (from spinner A) equaled the remaining six sticks from spinner B. Of course, the six red parts from spinner B are indeed equal to the one red sector in spinner A. Although an accurate description, her reasoning does not maintain proportional parts because her two "sticks" and the two areas in spinner A are not equal parts. Although she tried to demonstrate the equivalence, and was using appropriate reasoning when she established the two blue parts from spinner B as equivalent to the one blue part from spinner A, her total analysis was flawed due to the initial unequivalence of the parts in spinner A.

To try to draw upon a familiar reference for her, I drew a circle and shaded half of it blue. I then asked her to describe the chance of landing on blue in this new picture. She looked puzzled and said “I don’t know.” I then drew another circle and shaded approximately 1/10 of it blue and asked her if that was the same as spinner A and a one out of two chance for picking blue. She said “I think so.” I then used my last drawing (about 1/10 blue) and asked if “the chance of landing here [in the blue] is the same as landing anywhere over here [in the large white sector].” She then replied “No … actually I messed up. It depends on what this circle equals. It depends on what the fraction is.” She then used the last circle I had drawn and made 10 almost equal sectors said “this [the original blue slice I had drawn] is one out of ten.” Since she was now focusing on equal sectors for establishing the part-whole relationship, I asked her to look at spinner A again and asked if the chance for blue was one out of two. She compared spinner A with spinner B and noted that she could see the lines in spinner B that made it two out of eight. She then drew imaginary lines in spinner A to make it have eight equal sectors and said that spinner A was also two out of eight and not one out of two [as she scribbled over the ½ written on the paper and wrote 2/8 beside it]. Although Jasmine initially did not consider that equal parts was important in making part-whole statements of chance, she eventually, with my guiding questions, made the connection with ideas she had used previously about the whole pie being worth a certain number and then each equal slice being one out of that number.

For the last part of this task, we revisited the idea that the game was not fair. Jasmine was unable to verbally describe what it means for a game to be fair, but she demonstrated how she could make spinner A fair if she could make another sector the same size of the blue one (1/4) that was currently red and color it blue. That visually would give her a spinner that was half red and half blue; however, she was not able to verbalize why this arrangement in the spinner made it fair. It seems she had an intuitive understanding of fairness and it's dependency on equal, but she could only demonstrate that concept and not describe it her own words.

Strengths and Weaknesses in Post-Interview

Jasmine was easily able to make *a priori* probability statements based on known possible outcomes (e.g., bag of marbles) as well as outcomes she constructed as part of a task (e.g., the family arrangements). She associated the concept of fair with notions of equal chances, although she often had difficulty verbalizing her conception of fair. She could use a part-whole statement to describe the chance of an individual event, but she mainly relied on part-part reasoning throughout the post-interview.

In some of the tasks, she used a part-part difference strategy to compare probabilities (e.g., 2B3W and 5B6W bags have equal chances for picking out a black marble) and to predict a theoretical distribution based on experimental results (e.g., her use of an additive relationship from the bar graph to predict a bag of marbles). She did employ part-whole reasoning when predicting the bag of marbles from the pie graph but had difficulty transferring this to the spinners in the last task. Her use of part-whole reasoning in that task did not initially account for unequal parts in spinner A. She had to be engaged with the task for quite awhile with several on-the-fly questions from me (e.g., the questions relating spinner A and her notion of "1 out of 2" with the additional circle graphs I drew) before she established appropriate part-whole relationships based on equal sectors. The only evidence of proportional reasoning was seen during her work with predicting a bag of marbles from the pie graph.

Jasmine's conception of chance and randomness seems to contain a strong notion that most results are possible, albeit unlikely. This notion appears to supercede any use of the EOP when considering a large number of trials. Although she expected a fair coin to get results "around the middle" with a 1000 trials, she reasoned that even 900 tails and 100 heads were "possible" and couldn't tell you anything about the fairness of the coin. Thus, her understanding of *a postiori* probabilities and her ability to analyze data from a frequentist's perspective is very weak. Although she did reason from frequentist perspective during the teaching sessions (e.g., using experimental data to hypothesize the secret weights of 2-2-1), it appears she may not have had enough experiences with tasks that used *a postiori* reasoning.

Jasmine's understanding of independence seems to be developing somewhat. She expressed notions of independence in noting that "just because you got more heads and less tails [HTHHHH], they are still a head on one side and a tail on the other," and in her idea that most sets of six results from a coin toss where equally likely. She also recognized that all possible arrangements of the three children were equally likely. However, Jasmine did express ideas about results with "patterns" in them being slightly less likely to occur. It seems that her scheme of randomness includes a strong notion of "mixed up" that is conflicting with her developing understanding of independence.

Jasmine used a "flipping" strategy to list possible arrangements in the family task but was unable to use any type of reasoning to convince me she had them all. She did easily reason about the chance of an exact arrangement occurring by using the number of possible arrangements as the "whole." She was also able to use the sample space she created to state the probabilities for several unordered combinations occurring.

Overall, Jasmine showed evidence of developing a more sophisticated approach to probabilistic reasoning. She displayed evidence of understanding several concepts (e.g., sample space, theoretical probability, ordered and unordered compound events) and beginning to develop conceptualizations for concepts such as independence and experimental probability. Her reliance on part-part reasoning and additive relationships shows she has not developed a stable understanding of part-whole reasoning, nor is she able to establish and maintain multiplicative structures.

Putting It All Together: Jasmine's Development of Probabilistic Reasoning

Seeing Jasmine work with the computer microworld was like watching her open a toolbox and start using the tools to build a playhouse. She was highly motivated and engaged in all the teaching sessions. When given the opportunity, she playfully contrived experiments and made up stories to help verbalize her data analysis. She also learned to use the tools to her advantage and often chose to use a certain tool because she knew it would help her in a certain situation. Jasmine was empowered by the tools in the microworld, and her expressions and lively interactions certainly showed her enthusiasm.

At the beginning of the teaching experiment, Jasmine had several primary intuitions about chance occurrences and was able to quantify her belief in the chance of an event occurring with her non-normative use of percents. After she used percent language appropriately with the spinner task in the pre-interview, she only used percents on occasion and it was always in reference to 50%. Her strengths from the pre-interview included her combinatoric reasoning with the 3-coin task and her reasoning with the continuous areas on the spinners and references to “half.” She was actively engaged and used her intuitions in her social (e.g., sharing her predictions or reactions to experimental results) and computer interactions (e.g., designing playful experiments and using the tools to test her conjectures) during the teaching sessions. Several times during the teaching sessions, Jasmine used the microworld to model something of interest to her (e.g., bingo barrel). She also used many metaphors when designing an experiment or interpreting experimental results (e.g., “gray guy eating a piece of the sky”). Her social and digital interactions certainly illustrate her use of the microworld tools as cognitive prompts and elements of perturbation during problem solving (e.g., her use of pie graph).

Jasmine made significant progress in her understanding of theoretical probability. She could use part-whole statements to describe the chance of an event occurring in both discrete and continuous cases. The use of the weight tool, graphical displays, and experimental data seemed to help her interpret theoretical probability and make judgments of the likelihood of certain events, like getting more white marbles from a 3B1W bag, with respect to the *a priori* probability, as well as the number of trials.

During the teaching sessions, Jasmine eventually transformed her part-part reasoning into part-whole reasoning. Most of her progress with part-whole reasoning involved the use of the weight tool and pie graph. In addition, she began to use her additive reasoning to develop multiplicative structures through iterations (e.g., recall her work with the twice as likely task). The only time she displayed evidence of a multiplicative scheme in the post-interview was when she was comparing spinner A to spinner B. Her development of the effect on the whole to its parts from the teaching experiment also did not transfer to her reasoning with comparison tasks in the post-interview. In making marble bag comparisons, she used similar strategies as she had in the pre-interview, mainly focussing on the effect of the number of undesired events on the desired events. Her part-whole scheme seems to be heavily reliant upon the visualization of a circle and did not transfer to discrete situations. Instead, she reverted to using part-part and additive reasoning for comparing the chances.

Jasmine's conception of independence improved slightly throughout the teaching experiment. However, she consistently believed that experimental results in "patterns" would be less likely to occur. Her intuitive understanding of random as "mixed up" seems to directly influence her development with independence as well as the law of large numbers. Although she showed evidence of understanding the EOP during her work with the *Probability Explorer*, this notion did not come in to play during the post-interview task about experimental results from 500 and 1000 coin tosses. It seems that perhaps Jasmine's experiences with the EOP in the microworld were not sufficient in developing an intuition about the law of large numbers that would transfer to her analysis of tasks not associated with a dynamic visualization process that she experienced in the microworld. She seems to have made progress in using *a priori* reasoning, but her reasoning with *a postiori* situations is still weak. I conjecture that an additional reason for her lack of *a postiori* thinking is her unstable multiplicative structures.

Jasmine showed evidence of developing multiplicative structures during the teaching sessions. However, her development was very reliant on her use of the pie graph and weight tool. This is positive in the sense that the tools where able to help her use appropriate multiplicative reasoning. However, her use of additive iterations with circle slices and repeated addition did not appear to help her create multiplicative relationships with discrete objects such as the bag of marbles and her stick drawing that she used in the post-interview. Jasmine needs many more experiences with constructing multiplicative relationships before a stable scheme will develop.

Overall, Jasmine made progress in probabilistic reasoning throughout the teaching experiment. I believe the microworld tools helped her to make connections between part-part and part-whole reasoning, as well between numerical and graphical representations. Her primary intuitions about chance have developed into a better understanding of several probabilistic concepts. Although some of the reasoning she used while engaged with microworld activities did not transfer to her analysis of tasks in the post-interview, one cannot discount the development of powerful ideas during the teaching sessions. The lack of transfer suggests that Jasmine's work was not sufficient for sustaining the probabilistic thinking she was using. This also suggests that Jasmine's probabilistic thinking was directly linked to her experiences in the microworld and might not have occurred with non-computer based activities.

CHAPTER 7

THE CASE OF AMANDA

Amanda is nine years old and in the fourth grade at a rural elementary school outside of a university town. She has a vibrant personality and is a very social child. However, Amanda struggled with verbalizing her thoughts and explaining her reasoning on mathematical tasks. She seemed easily intimidated when the other girls gave explanations for their reasoning and she either did not have a cognizant reason or was unable to accurately verbalize her thoughts. She got frustrated quickly with tasks that challenged her and passively participated in these tasks and did not offer many comments or questions during whole group sessions. After the third teaching session, Amanda expressed concern to her parents that she felt she did not know as much as the other children and felt inadequate during the whole group sessions. Amanda worked much more consistently and persevered on her attempts to reason through a task when working in a child-researcher pair. However, she needed constant reassurance from the teacher/researcher that she was doing a good job.

Pre-Interview Analysis

Amanda's pre-interview occurred on August 17, 1999 for about one hour and consisted of the tasks listed in Appendix E. Throughout the interview, she used the materials available to her (e.g., coins, bags with black and white marbles, spinners, buckets with cubes) and used paper and pencil to record data. I will briefly discuss our interactions during each task and summarize my understandings of her conceptions.

To begin the interview, I asked her to interpret the following situation:

> Suppose you and a friend were playing a game and during the game your friend said she wanted to quit because she didn't think the game was fair. What do you think she means by the game is not fair?

Amanda suggested “we are not playing by the rules.” I asked her to describe the kind of rules that made a game fair and she replied “having fun, taking turns, and sharing.” I further asked her what she thought could be wrong with a game to make it unfair. She again noted “that you are not playing by the rules” and “this is a game board or something and you are cheating and not letting the other person do what they need to do.” Her conception of fair in the context of a game was limited to characteristics of game-playing (e.g., rules, taking turns) and did not seem to include any notion of equal chances of winning. To assess her understanding of fair in a sharing context, I gave her nine blocks and asked what she would do to share these fairly. She immediately replied “split them in half” but when I asked if splitting things in half made them fair, she replied “not always.” She then used an equal partitioning strategy, gave each of us four blocks, and said that for the remaining block “whoever wants the extra one should have the extra one” or that we could “cut it in half.” Her conception of fair in terms of sharing certainly is based on equal parts; however, she does not seem to apply her conception of sharing fairly to establishing a fair game in terms of equal chances. It seems that in a game context, “fair” may indicate that the game is worthwhile to play.

Bucket of Cubes

During the interview, I used a bucket of colored cubes to assess how Amanda described the chance of a certain color being randomly chosen, as well as her understanding of vocabulary such as "most likely" and "least likely."

Nine cubes. For the first task, I presented Amanda with a bucket containing four green, three red, and two yellow cubes. She used qualitative reasoning to determine that green was the most likely "because it has the most" and yellow was the least likely "because it's the least." When I asked her to compare the chance for picking a red cube to the chance for a green cube, she explained "they are one off, red is one off… when you mix it [the cubes] up, having the most blocks, more would probably be on top than the least." Although she never explicitly referenced the number of cubes of each color, her qualitative response indicates that she used part-part comparison to make her judgments. In addition, she also was considering the physical aspect of the cubes in the bucket in noting that green cubes would more likely be on top. To continue the task, I closed my eyes and randomly picked a cube from the bucket several times, each time asking her for a prediction. She always predicted either green or red but, even when asked, did not give a reason. After four tries, I had picked two green and two red cubes. Amanda then said "so it's equal, it's probably even then" but could not explain what she meant.

I continued the task by asking Amanda which color had the best chance after I removed a green cube (leaving three green, three red, and two yellow). She used a direct part-part comparison to say "you would probably pick out a green and red the same amount" but added that green was most likely "because I like the color green." After I removed another green cube from the bucket she stated that she was most likely to pick out a red cube “because it has the most blocks.” In addition she noted that green and yellow were tied for least likely because “they are both equal because there are just two greens and two yellows.” Although she expressed her personal bias towards green, she used part-part analysis and recognized that parts with the same number had equal chances.

Later in the interview, I brought this bucket back out and asked her to describe the chance for picking each color. She noted that for green “it’s pretty easy,” the yellow was “hard,” and that the red was “sometimes and sometimes not.” Her qualitative description was based on the amount of cubes of each color, but did not provide any indication of how “easy” or how “hard.” Her language was vague enough to capture the overall picture but did not give enough detail about the chances for someone who did not know the contents of the bag.

I continued the task by removing two green cubes and asked her if the chance of picking a green cube had changed from the first bucket (4 green, three red, two yellow) to the contents after I removed two green cubes (2 green, three red, two yellow).

> A: No. Yes.
> T: Yes. All right. Why has it changed?
> A: There are two yellows which we said was harder to pick. Now there are only two greens. So the yellows and the green are buds [buddies].
> T: They are buds! [chuckle] So before when these were in here, you said the chance of picking a green …
> A: The chance of picking a green is higher, no easier.
> T: Easier. So now the chance of picking a green is....
> A: Harder.
> T: Is harder. Has the chance of picking a red one changed now that I have taken these two out?
> A: No.
> T: Any why hasn't it changed?
> A: Yes.
> T: Oh it has changed. How?

> A: When you took the green was the easiest to pick and red was just one down from that. And so now you just took two out, so that makes red number one.
> T: And so the chance of picking a red has changed. Has it improved or gone down?
> A: Improved.
> T: How has the chance of picking a yellow changed? Or has it changed?
> A: No it hasn’t.
> T: It hasn't. And why hasn't it?
> A: Because it's always been third and we didn't really do anything to make it better. Just green and red were the only ones to be changed.

She used part-part reasoning in this task and was able to reason about chances for the green and red changing based on their change in relative position on a most-middle-least scale. Since the yellow cubes were initially “third” and were tied for least after the green cubes were removed, she did not think the chance of picking a yellow had changed. This demonstrates that she was not considering the total number of cubes in the bucket at all in her analysis.

Four cubes. For another bucket task we used three green cubes and one red cube. She described the chance of picking a green as “way easier” because “there is only one red and when you mix them up green is always going to be on top.” She described the chance for red as “you probably won’t even pick it.” Having only one red cube in the bucket, although there were only four cubes, seemed to skew her perception about the relative chances for picking each color. She does not seem to be considering the total number of cubes in making her judgment.

I asked Amanda if she could use numbers to describe the chance for picking each of the colors. She struggled and then said “I don’t know how to talk in number talk.” She then noted that for the chance of picking a red, “you are not going to pick a red, the most that you are going to pick a red is only once.” When I asked her how long she would have to pick out cubes before she got the one red cube, she said "probably for about a whole day" but then readjusted and said "no about maybe three minutes." She does not seem to have quantitative language to use to in her description but does describe the chance for a red in terms of a hypothetical experiment where it would take many draws to choose a red cube. Although she initially said "a whole day" she did reassess her judgment and drastically lower her prediction to three minutes. She could not explain why she predicted three minutes

After I removed one green cube from this bucket, she said the green was "still easier" and for the red, "it's probably going to get it maybe about three times." Although she did not say this, she may have thought the chance for red had improved because there were less total cubes or just because there was only a difference of one between two green and one red cube. When I removed the red cube (leaving two green cubes in the bucket), I asked her to describe the chance for picking a green.

> A: You may as well just take them out and get another color.
> T: Why?
> A: Because you are going to pick green the whole time.
> T: Why am I going to get green the whole time?
> A: Because there's no other color.
> T: Can you use any kind of numbers to describe for me, or any words, that I'm going to pick out the green?
> A: You are going to pick out, you can try all day and you'll just pick out green.
> T: You'll just pick out green?
> A: Unless you paint the green a different color.

Although she did not use the words "sure" or "certain" or a numerical descriptor like 100%, her description of experimental results that are always green is a strong indicator that she recognized the only possibility was green.

<u>Coin Tosses</u>

Two types of tasks were used involving coin tosses. The first task was designed to assess the concepts of equiprobable and sample space while the second task assessed the concept of independence.

<u>1, 2, and 3 coin toss.</u> I gave Amanda a penny and asked her what different ways it could land if I flipped it in the air. She noted that it could be heads or tails and “if it is heads when you are going to flip it over it will probably land on heads” but did respond “no” when I asked her if that was always the case because “there is one on each side.” She described the chance of getting heads as compared to tails as “they are equal … the heads is just on this side and the tails is just on that side.” Although she had initially described a hypothesis that you could affect the outcome by what started facing up when you flipped the coin, her latter response indicates she believed heads and tails were equiprobable.

For the next task in the interview, I asked her what the possibilities would be if I flipped two coins at the same time. She used the coins to model the first two possibilities and listed the results TT, TH, HT, HH. When I asked her if TH was different or the same as HT she described how she could tell the coins apart (color, size) and then modeled the nickel as heads and penny as tails then flipped each coin over to show the opposite side. I then asked her if any of the four possible ways were more or less likely to occur. She said the “matching” ones (HH and TT) were more likely because of her flipping hypothesis that “to me when I flip it when it’s on tails it mostly lands on tails and the same with the heads.” It seems that she thinks she can control the outcome by how she decides to flip the coin and that this would make it easier for her to control for HH and

TT. However, this reasoning is not consistent, because she could supposedly use the same strategy for trying to get TH or HT. It is interesting that she chose TT and HH over one of the “mixed” possibilities.

When listing the possibilities for flipping three coins (penny, quarter, nickel) she initially wrote TTT, HHH, THT, TTH, HHT, but then added HT to her third item to make THTHT. She then asked if she could start over because she “messed up.” On her second try (new piece of paper) she wrote TTT, HHH, HTH, THH, HHT and said she thought she had them all. She modeled her list using the coins and then found TTH. She could not verbalize a strategy and it is difficult to tell if she used a strategy because she wrote her list quickly and was leaning over the paper while she wrote. Nevertheless, she was not able to exhaust all possible outcomes.

In response to my question about whether any of the possible outcomes were more or less likely, she focused on the “mixed up” possibilities and did not refer to her flipping control hypothesis.

> T: Now all of these that you have here, do you think any of them are more likely to happen than the others?
> A: I think the mixed up ones are more likely to happen.
> T: And by mixed up you mean?
> A: Like tails, heads, heads. Heads, heads, tails … Not like all three of the same. I think the mixed up ones.
> T: And by mixed up you mean?
> A: Like tails, heads, heads. Heads, heads, tails.
> T: So they have some heads and some tails.
> A: Not like all three of the same.
> T: So why don't you think that all three of the same would be?
> A: Because the more coins you have, the least chance you are to have all of the matches.

She did not seem to be differentiating the “mixed up” ones and seemed to be basing her response on “mixed up” versus “matching” rather than the six ordered possibilities she had listed. If she was thinking about the unordered lists, then she is certainly correct that the "mixed up" ones are more likely to occur. In addition, she seems to have an intuition that the chance of getting "matching" results decreases as the number of coins flipped increases. When I asked her to explain why, she merely said with more coins "the harder it is to have matches … because, [pause] I don't know." Although she could not verbalize her reasoning, she displayed evidence of a primary intuition about the effect of the number of trials on the chance of "matching" results occurring."

Flipping a coin six times. Later in the interview I asked Amanda to predict what I would get if I flipped a penny six times. She noted that I should get TTHTTH because she thought she saw that I was holding the penny with the "tails up" and that she thought her list made "a nice pattern." I then flipped the coin and got THTTTH and asked her if either her sequence or my sequence was more likely to occur. She said that we both got four tails and then said "I don't know." When I rephrased the question as "we both in our list have four tails and two heads, but they are in different orders. Do you think that in the different orders that one of the orders is more likely to happen than the other?" she quickly said "no." Her response indicates that she may think that as long as both sequences have the same distribution (four tails, 2 heads), that the order would not matter and that both possible sequences would be equally likely.

For the next series of questions, I showed her four possible results from tossing a coin six times and asked her if any of the possible results of six were more likely to occur. For the first set (HHHHTT, THHTHT, THTTTH, HTHTHT), she thought HHHHTT would be more likely because "it's not as mixed up as these [other possibilities]" and then added "remember what I said about matches? This one has matches." She seemed to base her reasoning on "matches" being more likely to happen, which is the exact opposite reasoning that she used earlier. She also used the "matches" are better strategy for the next set of possibilities (HHHTTT, HHHHHH, THTHTH, HTHTHT). She thought HHHHHH was more likely to happen because "they are matched up." She then wanted to do an experiment with the penny to see what she got. She flipped the penny five times (not purposely starting with either heads or tails before she flipped) and got TTTTT. I asked her what she thought she would get next. She said "tails because I'm on a roll!" She flipped and got a heads and then reconfirmed her previous choice that HHHHHH was most likely to happen, but provided no further justification. Again, her "matched up" hypothesis was not consistent with her prior thoughts that the chance of getting all heads decreased as you did more trials. In addition, her favoritism towards the "matched up" possibilities does not reflect the findings from previous research about representativeness. However, she did employ the positive recency effect when predicting another tails after five had already occurred. It is important to note that the results from her experiment probably only reinforced her notion that matched up results were more likely to occur.

Sampling

To assess Amanda's sampling strategies, I gave her a black bag and told her that it contained 10 tiles of three different colors (5 blue, three red and two yellow). Her task was to make a reasonable guess at what was in the bag using a with-replacement sampling method. She chose to draw from the bag 10 times because "there are 10 blocks." After six draws she had gotten RBRRBB and said "there's only red and blue."

When I asked her is she was sure, she replied “yes because that’s all I’m getting.” She then drew out another blue. In her next pick she accidentally pulled out three tiles (one blue and 2 yellow) and said “I know there are two yellows” since two of the tiles she saw where yellow. She then predicted four blue, four red, and two yellow in the bag. When I asked her if she was confident she knew what was in the bag, she said “no” and said she would feel better if she picked out three more times. She picked out one additional yellow, blue, and red tile and then said “can I look?” Before I could answer, she tilted the bag and looked inside. When comparing her results with the actual contents, she commented that she knew she would “get about one or two wrong.”

Her sampling strategy seemed reasonable, but her accidental draw with three tiles gave her too much information to really assess her reasoning in making a prediction based on her sample. She did have an intuition that she would like to do more samples and that her prediction from the data she gathered would probably not be exactly right.

100 Gumballs

This task was used to assess her ability to use proportional reasoning and theoretical probability to make a prediction for a sample when the population is known. Given a gumball machine with 50 red, 30 blue, and 20 yellow gumballs, she predicted “5 or 6” red gumballs because “like I said the more the better chance that you are going to get the most.” She predicted three blue and two yellow and then five reds so it would equal 10. She was not able to give any justification beyond “red’s the most” for her predication. Although her prediction is in perfect proportion to the contents of the gumball machine, there is no verbal evidence that she used proportional reasoning in this task. However, her prediction at least is consistent with the most-middle-least rankings she used with the bucket of cubes earlier in the interview.

Marble-Bag Comparisons

For the first two pictures of bag of marbles, I only asked Amanda to describe the chance of picking a black marble. Bag #1 contained two black and two clear (2B2C) marbles. She said “you would probably pick out the same” because “they are equal.” She seems to be indicating that in an experiment, you would probably pick out the same number of black and clear marbles since there was an equal amount of each color in the bag. This is the third time that she has used a hypothetical experiment to describe the chance of something occurring. However, she was able to justify her reasoning based on the distribution of marbles in the bag.

In her assessment of bag #2 (5B3C), she described the chance of picking a black as “the most because it just has more” and compared black with white as “black is more likely to be picked than the white.” She did not use any reference to the quantity of black or clear marbles, although her judgment of “most” is obviously based on some type of numerical or visual comparison of the quantities.

For the remaining pairs of bags (#3 & #4, #5 & #6, #7 & #8), I asked her to choose which bag she would prefer to pick from if the goal is to try to pick a black marble. With each pair of bags, I reiterated a question such as “would you like to pick from bag #3, bag #4, or does it matter which bag you choose from?” For the first pair of bags, the distributions were proportional with 3B1C in bag #3 and 6B2C in bag #4. Amanda briefly looked at both pictures and said she would prefer to pick from bag #3.

A: I would rather pick from bag #3.
T: And why would you rather pick from bag #3?
A: Because there's more blacks than whites. Plus there's only one white so [pause]
T: So in bag #3 there are more blacks than there are whites?
A: Plus [long pause as she stares at bag #3]
T: And you said something about the one white. What's so special about having the one white there?
A: I'll probably touch it but not pick it up. There's only one white.
T: What's over here in bag #4?
A: There's two whites which is still less than the six blacks. But you have a better chance of picking a white in here [bag #4].

Amanda used a part-part comparison between the bags and chose the bag that had the least number of undesired events (clear) and also justified her reasoning based on the physical aspect of the real bags and that she might “touch it but not pick it up.” She reasoned that a lower chance of picking out a white would mean a higher chance for choosing a black marble.

She used similar reasoning in comparing Bag #5 (1B4C) and bag #6 (2B8C). She picks bag #5 but remarks “but you are not going to pick it” because there are too many white marbles in the bag. She reasons that although there are two black marbles in the other bag, there are “going to be a lot of whites on the top.” Again, her reasoning is based on minimizing the number of undesired events as well as the chance of the color having the most marbles being more likely to be at the top of the bag.

The last two bags were not proportional with 2B2C and 2B3C in bag #7 and #8, respectively. She chose bag #7 because “they are equal” and “there is going to be one black and one white most likely in the top.” She also noted that you could still “probably pick a black” in bag #8 but that she “would rather not risk it” since it is more likely to pick a white marble in that bag. She again referred to position of the marbles but used an appropriate judgment based on equal chances in bag #7 and a higher chance for white in bag #8.

Spinner Game

The next task consisted of the penny game (as described in Chapter 5) played with a spinner containing three unequal sectors (1/2 red, 1/3 blue, 1/6 yellow) and eight pennies. Amanda chose the red sector because "it covers up half of it, it's more likely to land on red." I chose the blue sector and asked her if this was a fair game. She replied "yes, I guess … we have the same amount of pennies and were not arguing [pause] I don't know the rest." But then she spontaneously references the spinner in her thinking.

> A: I wouldn't say it's not really fair on this [points to the spinner], because well I mean. Yes I think it's fair on here, because if it was all of the same [points at yellow and blue areas] it would probably, it wouldn't really matter what color you pick because it would be half and half.
> T: And so what do you mean by half and half, what is half and half?
> A: If this was blue right here and this was red like it is. It wouldn't really matter what color you pick because it would be boring. So you would have to
> T: So what does it do to it that the yellow is in here and none of us are yellow?
> A: Make it more harder to know which color it is going to land on.

It appears that Amanda recognized that the areas on the current spinner were not fair. She then used appropriate reasoning as she imagined the yellow and blue areas being the same color and then having the spinner half red and half another color. But, when I asked her about what it meant that the yellow sector was in there, she only noted that it made it harder to know which color the spinner arrow would land on and did not reiterate her notion about unfairness. Therefore, it seems that she thinks this spinner is unfair based on areas but fair because it seems worthwhile to play.

After we played the game and she won, I said “I don’t want to play this game again because I don’t think it’s fair.” I asked her to convince me whether the game was fair or unfair.

> A: It's fair because if the colors were the same amount, then like it would be 50 - 25, then if it was the exact same amount then it wouldn't really be fun because you would know that it would probably blue, red, blue, red the whole time. And then it would be oh [faking a yawn] this game is boring.
> T: Well what if I said I don't think I have as much as a chance of winning as you do because my area is not as big as yours?
> A: That's the point.
> T: That's the point. And why is that the point?
> A: Because if it was the same like I said it would be boring. Because it would be like a pattern.
> T: Would you mind if you were blue and I was red?
> A: No. Because I like the color blue.

This exchange illustrates that Amanda’s conception of fair is not necessarily connected with a notion of equal. She thinks that having the areas unequal make it more exciting so you can’t just predict a pattern of back and forth results. A fair game to her seems to include an excitement factor based on unpredictability.

For the final task in the interview, I asked Amanda to compare two spinners and decide which one she would rather use to play the penny game if she still won when the arrow landed on a red sector. Spinner A was the same one used in the previous task. Spinner B contained 12 equal sectors, six red, four blue, and two yellow spaced in the following pattern (r, b, r, y, r, b, r, b, r, y, r, b). She chose spinner A because it was more “solid” and had a better chance for landing on red. When I asked her if she could use numbers to describe the chance for landing on red in spinner A, she said 50% because “half of 100 is 50%…50 plus 50 is 100. So like they say in math, if this [half of spinner] was red and this [other half of spinner] was red, it would be 100%. If it was just half red it would be 50%." Her use of percent language was entirely appropriate and demonstrates that the circle was a familiar context for her as a representation of one-half and 50%.

Amanda then pointed to the red sectors in spinner B and thought they would add to 50% if she moved them. She modeled a pretend "trading places" technique but then concluded that the red areas would not quite be 50% and were probably 45%. She did not employ any numerical reasoning with the number of sectors in spinner B. Although her trading places strategy was appropriate, I think she had a difficult time visualizing the shifts on the spinner and misjudged the red areas to be slightly smaller than 50%.

Although she did not use proportional reasoning to compare the spinners, her references to 100% and 50% as well as the relationship of "half" was the most quantitative analysis she did in the entire interview. The circle definitely seemed to be a familiar context and give her a numerical and visual reference to half and 50%.

Strengths and Weaknesses from Pre-Interview

Amanda had a difficult time verbalizing her reasoning on many of the tasks. She had several primary intuitions about probabilistic concepts. Her sense of "fair" was based more on game-playing actions (e.g., taking turns) rather than the structure of the game (e.g., equal chance of winning). She associated fair in a sharing context with equal parts but did not apply this association in evaluating whether a probabilistic situation was fair based on equal chances. With the spinner tasks, she was able to recognize that "it's more likely to land on red," but wavered back and forth in her reasoning about whether the unequal areas on the spinner made the game fair. Her notion of a fair game in the spinner context seemed to include an excitement factor about not knowing where the arrow was going to land. She showed evidence of somewhat understanding the uncertain nature of random situations. However, she also expressed realistic ideas that the outcome of a coin toss could be determined based on how it was flipped and that the position of a cube in a bucket made it more or less likely to get picked. She could use part-part comparisons to reason quantitatively using most-middle-least benchmarks to estimate probabilities.

Amanda did not employ consistent reasoning with the coin toss tasks. She could reason about the difference between HT and TH but noted that the "matching" ones (HH and TT) were more likely to occur based on her controlling the flipping process. With the 3-coin task, she could only list six possible arrangements but noted that the "mixed up" ones (e.g., TTH, THT, HHT) were more likely to occur and that "the more coins you have, the least chance you are to have all of the matches." She has already switched strategies, but the inconsistency occurs yet again with the 6-string questions. Given both sets of four possible results, she chose the results (HHHHTT and HHHHHH) that were "not as mixed up" to be the most likely to occur. She also did an impromptu experiment, got five tails in a row, and predicted another tail since it "was on a roll." Her reasoning about coin tosses is not at all consistent. At times she thinks "mixed" results were more likely, and other times she thought "matching" results were more likely. Her conception of independence is certainly weak and she seems to have conflicting ideas about the effect of the number of trials on getting a sequence of the same result (e.g., all heads).

Amanda was able to easily list the four elements in the sample space for the 2-coin toss and differentiate between TH and HT. With the 3-coin toss, she did not appear to use any strategy and was only able to list six possible outcomes. However, she reasoned with the unordered combinations when asked if any arrangement was more or less likely or if they were all equally likely. Her intuition about the “mixed ones” having a higher chance was appropriate since she was considering the unordered arrangements.

In comparing the bags of marbles, Amanda used part-part reasoning and consistently used a strategy that the bag with the least number of undesired events (clear marbles) made the desired event (black marbles) most likely. In comparing the chance of landing on the red sector in both spinner A and B, she made reference to 50% and imagined shifting the red sectors in spinner B to make a solid 50% like in spinner A. However, she thought it would be easier to land on red in spinner A since it was more “solid.”

Overall, Amanda’s responses during the pre-interview demonstrate that she entered the research study with conflicting and unstable intuitions about chance. She relied on part-part reasoning and did not display any instances of part-whole reasoning. The only hint that she considered the whole is with her reference to the whole pie in the spinner as representing 100%. The area model and reference to 50% and “half” seemed to be a familiar context for her.

Amanda’s Meaning-Making Activity with the Microworld

Amanda participated in approximately eight hours of small group teaching sessions and two hours of individual sessions. As mentioned earlier, Amanda was not as verbal as the other two children and was easily frustrated with challenging tasks and her own inability to verbalize her thoughts. There were many times during the group teaching sessions that, although she may have been using the microworld tools to investigate something, she did not contribute anything verbally unless directly asked by myself or the other teacher/researcher. Amanda worked much better during sessions when she was in a child-researcher pair (third and fifth sessions) and in her individual sessions with me. During those sessions, she was able to verbalize her thoughts better since the teacher/researcher could help her work through the elaboration of her statements and give her plenty of time and encouragement to think through her reasoning.

The analysis of the teaching sessions with respect to Amanda brought forth four evidentiary themes in her development of probabilistic reasoning: 1) her interpretation and use of theoretical probability; 2) her intuitions about expected results; 3) her interpretation and use of the pie graph; and 4) her use of additive and multiplicative reasoning. Amanda's thinking with theoretical probability was tightly intertwined with her intuitions about expected results. Therefore, I have collapsed those themes into a single discussion. What follows are my observations and analyses of Amanda's meaning-making activities, mathematical ideas, intuitions, and conceptions under three themes. I also highlight how she used and interpreted the microworld tools. For cross-case comparison purposes, I have included a thick description of Amanda's work on the "twice as likely" task.

Theoretical Probability and Expected Results

As the teaching experiment progressed, Amanda developed her ability to interpret theoretical probability. She also developed her ability to predict and interpret experimental results based on the theoretical probability. At the beginning of the teaching experiment, Amanda only used qualitative descriptions of probability. For example, in the first teaching session, when asked to describe the chance for picking a black marble from a 2B2C bag, she claimed it was "both easy and hard." However, when Carmella described the chance as "50-50," Amanda interpreted that to mean "they are both even." In addition, for 100 trials, she thought it would be 50 blacks and 50 whites because "100 is an even number so half will be black and half will be white." Although she was not initially able to quantify the chance of picking a black marble, she was able to reason about the results based on the contents of the bag. Throughout the first teaching session, she always predicted "around even" results when experimenting with coin tosses. The other girls almost always predicted exactly "even" results and Amanda would predict results that deviated by one from "even" (e.g., for 20 coin tosses, Amanda predicted nine and 11). Although she did not give a reason for her prediction, she may have had an intuition that it is more likely to get either nine heads and 11 tails or 11 heads and nine tails than it is for five heads and five tails.

At the beginning of the second teaching session, I asked the girls if they thought any of the numbers on a die were more or less likely to occur than the others. Jasmine and Carmella both said "no" but Amanda said "I think [pause] the six is a little bit harder [pause] no actually I think it's even." Amanda's change in response may have been influenced by the other girls' responses. However, her initial instinct of six being harder to get is consistent with other research results (Green, 1985) that showed some young students feel that six is more difficult to occur. Green hypothesized that this intuition may be a result of game playing experiences where children are trying to roll a six and cognitively compare the chance of getting a six with all five other possible results. From this perspective, the number six is certainly more difficult to roll. Amanda followed up with “I think I changed my mind because [pause] I didn’t know why.” She may not have been able to justify her intuition of “six is harder” and decided to just agree with the other girls.

During experimentation with simulating a die toss, Amanda did not contribute much to the group discussion. When discussing the results from 40 trials, she noted that “mostly, a few [stacking columns] would be blank and sometimes they would all be close.” Amanda and Jasmine had gotten several results where one number on the die never occurred. She recognized the wide variability and did not think it was particularly unlikely to not get a number when rolling 40 times, although she did think getting all sixes was very unlikely. In addition, when doing a large number of trials, she thought “the low ones will have a comeback.” Amanda did not ever support her observations and intuitions about the expected results with reference to the theoretical probability.

During the third teaching session, Amanda described what might happen when choosing marbles from a 2B2W bag by saying “ it’s an even chance of being black or white ... because there’s two blacks and two whites ... none of them are more than the other, so there’s not a better chance of picking one or the other.” This was one of the first times Amanda shared an explanation with the group. She was able to reason directly with a part-part comparison of the marbles in the bag to support her judgment. Given the task to predict how many black and white marbles would be picked if we ran the experiment (with replacement) 10 times, Amanda guessed six white and four black, again slightly off from the theoretical expectation, because “it probably won’t be exactly even.” It seems that she knows that the best guess based on the theoretical probability is five and five, but that, perhaps from experiences with the microworld simulations thus far, exactly “even” results do not occur often.

Amanda also recognized this same relationship with a 5B5W bag of marbles and noted that results would be “the same as before” with the 2B2W bag because both bags have the “same number of blacks and whites.” Joe asked Amanda what she thought the results would be if there were not the same number of black and white marbles in the bag

> T2: What if we didn't have the same number?
> A: You would probably get one more than the other.
> T2: And which one would you get more of?
> A: It depends on what color has more in the bag.

Amanda’s response demonstrates that she saw a relationship between the distribution of colors in the bag and corresponding experimental data. This response also indicates that she was making the connection between physical objects and the digital representation, as well as understanding the simulation process. Perhaps the act of creating the bag of marbles herself made the abstractness of the random simulation process more concrete for her. Amanda later recognized the equivalence of the chances for getting the result of six black-4 white and six white-4 black because “they are exactly the same only switched around.” She also further explained why the chance for picking out a black marble was the same for the 2B2W and 5B5W bags.

> A: They are both even, they are just like this bag, it's just like that bag [2B2W bag] only this [5B5W bag] has more, that has more marbles than the other. That has five blacks and the white has five whites. But they are both equal and an even number. They are just littler numbers.

Although she did not use any part-whole statements of chance, Amanda was able to reason appropriately that the bags had the same chance of picking a black marble because both bags had an “equal” amount of black and white marbles, although those amounts were different in the two bags. Part-part reasoning was sufficient for her to compare the chances in these bags.

Amanda had some initial difficulty in analyzing an unequiprobable situation. After she designed a 3B1W bag in the marble environment, Joe asked her to make a prediction for picking out 10 marbles. The following dialogue, although lengthy, illustrates her conflicting intuitions about the relationship between the contents of the bag and experimental results.

> T2: You are going to pick out 10 marbles. What do you think you'll get with black and white?
> A: I think I'll get [pause] five blacks and five whites.
> T2: Why?
> A: Because five and five equals 10 and 10 is an even number [she hits run 10 times and gets all black marbles] Oh my gosh! I got all black! [she stacks them]
> T2: Let's see. Does that look like five and 5?
> A: No!
> T2: No. So what happened here?
> A: I got all 10 blacks.
> T2: Now how do you think that happened? Remember we tried it before and we never got all of one? How come we got all of one now on our first try?
> A: [pause] Because when we picked the marbles somewhere up here [pointing to the buttons along the top of the microworld], there are more blacks than whites and so I think we picked all black because there is only one white.
> T2: So having three blacks and one white what does that do?
> A: It makes it easier to pick out a black than a white. Because if you pick out a white you know you are just going to get another black. And you get a black with it. A lot of the times you are going....
> T2: Oh you mean if you pick one and one?
> A: If you pick at least 10 times you'll get a white, but you'll probably get only two because you don't have much amount of whites in there.
> T2: And in this case what did we get?
> A: Just black.
> T2: Let's try it again. [she clears trials, changes the number of trials to 10 and hits run, and then stacks them] Wow, tell me about this one now.
> A: We picked four whites and six blacks.

T2: Now would you have guessed that?
A: No.
T2: Why not?
A: Because to me I think it's unusual to get that much white. Well maybe not, if it went any higher it would probably get more uneven.
T2: All right. But the time before you said you would get five whites. Remember, just the time before we clicked, right before this one. The first time we guessed after putting three black and one white you said we would get five black and five white. And now you think getting four is unusual. Why is that?
A: Um, I just think there's only one white that it's, [pause] there is one white and to me it's unusual to get that many.
T2: So you think you could get five and 5?
A: Um ... Yeah.
T2: And you think it's unusual to get four whites?
A: No [tilting head]
T2: Okay. So what do you think?
A: I think that [pause] it's unusual to get 10 whites.
T2: Yes, it would be. Absolutely. Okay, let's try another. [She clears trials, runs another set of 10 trials, and stacks them.] What do you think about that one?
A: Nine black and one white.
T2: So do you think that's unusual?
A: No.
T2: How come?
A: Because we just got it and because there are more blacks in the pile.

It seemed that her initial guess of five black and five white was based on 10 being an even number and that perhaps she thought that since there where two choices, the results would be equally distributed. However, the experimental results (all black) seemed to cause a perturbation that made her rethink her prediction based on how she had designed the bag of marbles. She also continued to say that six black and four white was not very likely. She tried to support her judgment when reminded of her previous guess of five and five but did not seem sure of her reasoning. Her judgment of “unusual” seems to be linked with her actual experiences with the microworld. I think she noted that 10 whites was unusual because there was only one white in the bag, all 10 white marbles had not happened, and that perhaps since six black and four white had occurred, it's degree of "unusual" was small as compared to 10 whites. However, it is important that Amanda was able to link the occurrence of more black marbles to the contents of the bag.

During the fourth teaching session, the students were playing a coin tossing game on the computer and I had secretly changed the theoretical probability to be 5/6 heads and 1/6 tails. Up to this point, the students had not used the weight tool and did not know that capability existed. After many sets of 10 trials, as well as several sets of 100 trials, the continual occurrence of many more heads than tails seemed "unusual" to Amanda and she conjectured that the computer was "eating the tails for supper." When Carmella discovered the weight tool, Amanda interpreted the weights of five and one as "there's only a 1% chance to get tails," without being able to explain her interpretation. When asked how to make heads and tails more even, Amanda suggested "1 and 1" and then later said "how about 89 and 90." When I asked if the weights were even, she recognized the inequivalence and thought that tails had a 1% better chance. Although her use of percents is inappropriate, her attempt to quantify the theoretical probabilities demonstrates an appropriate recognition of more or less likely based on the weights. Since most of her judgments about theoretical probability up to this point have been on a most-middle-least scale, her interpretation was consistent with this scale. However, it was obvious that the numerical representations in the weight tool were too abstract for her.

During the fifth teaching session, the abstract nature of the weight tool became even more apparent. Carmella had suggested that they use weights of 2000 and zero in a coin experiment. Amanda thought that tails could still occur "after a long time." She wanted to run the experiment more than 2000 times, and she fully expected a tail to occur shortly after the 2000th trial. When this did not occur, she was not sure how to explain the results. When I asked her what would happen if we used one and zero as the weights, she thought there would still be more heads, but that a “few tails” could occur. She ran a simulation with these weights and after about 400 trials decided that tails would not occur because “it’s like there are none in the bucket.” It was actually this comment, and Amanda’s struggle to interpret the representations in the weight tool, that inspired me to create a dynamic link between the weight tool and the bag of marbles in the microworld. This change in the design of the microworld occurred between the fifth and sixth teaching session. Amanda’s interpretation of a weight of “0” demonstrates that the number zero itself was too abstract for her to interpret in a meaningful way. However, the ability and speed of the microworld to simulate an experiment with different weights allowed Amanda to test her conjecture with experimental data and to readjust her initial interpretation based on the data.

Also noteworthy during the fifth session was Amanda’s exploration with the secret weights of 2-2-1 for basketball, baseball, and soccer. When told she could use the microworld tools however she wanted to find out the secret weights, she immediately ran 500 trials and opened the data table during the simulation. She eventually opened the bar graph during her analysis and initially guessed that basketball had the “most,” baseball had “a little less” and soccer was the “least,” but would not quantify her hypothesis. I will discuss her analysis further as part of the section on her additive and multiplicative reasoning. However, for this theme, it is important to note that she used a large number of trials to investigate the secret weights. She was the only student who initially ran 500 trials. She may have had an intuition that the results from a large number of trials could help her guess the secret weights. Of course, since she was not able to give a reason to support her action, she may have merely ran 500 trials just because she could and that, perhaps, she found the simulation process engaging to watch. Her lack of verbalization made it difficult to interpret her actions and intentions.

To help Amanda connect the representation in the weight tool with a physical situation, we spent the majority of the sixth teaching session designing a variety of experiments with the marbles and comparing the contents of the bag with the display in the weight tool. She first filled up the bag of marbles with some of each of the six available colors until she had the maximum number of marbles allowed (24). She used a part-part comparison to identify the most likely color (red) and to identify the colors that had the same chance (yellow and blue). I then asked her to interpret the fractions displayed in the weight tool (see Figure 7.1)

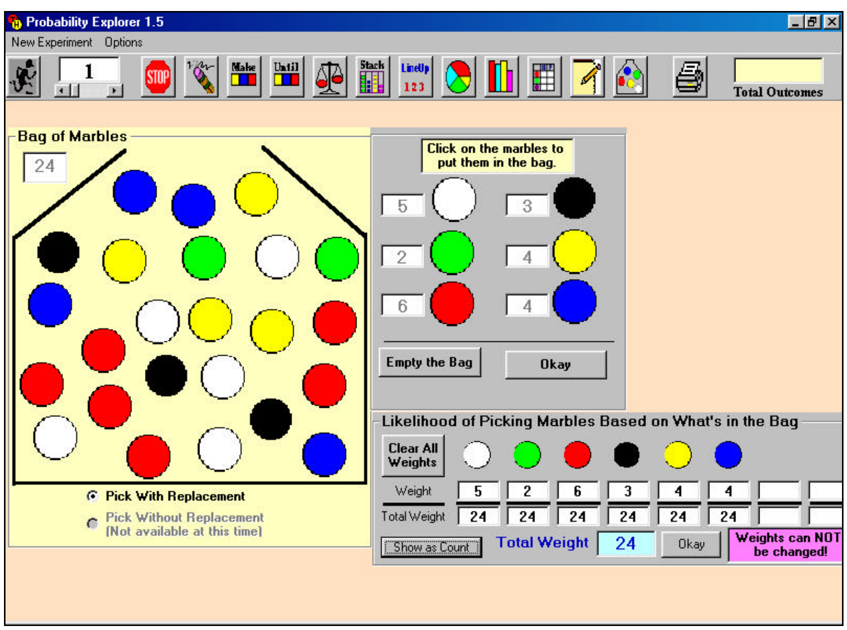


Figure 7.1. Amanda’s bag of marbles and associated display in weight tool.

T: Can you tell me what these fractions mean?
A: No.
T: Where am I getting this number 24 from that is on the bottom?
A: There are 24 marbles in the bag.
T: You are right. There are 24 marbles in the bag. So what does five over 24 mean with the white marble?
A: That there are five whites.
T: Five whites, right.
A: five out of 24.
T: Five of them out of 24 are white. That's right. So before you said that the yellow and the blue had the same chance of happening, that they were both four. And they both now are four out of 24. So do they still have the same chance?
A: Yes.
T: They do?
A: See, 24 and 24 and four and four. [pointing to each whole and each part in the weight tool]

Amanda needed focused questions for her to connect the numbers in the weight tool to marbles in the bag. However, she was eventually able to make the connection, including the equivalence of yellow and blue based on the same whole (24) and same part (4).

We continued to work with the bag of marbles as I asked her to empty the bag and put in two colors so that they had the same chance of being picked. She again filled up the bag using an alternating strategy of clicking on red, blue, red, blue, etc. until the bag had 12 red and 12 blue marbles in it. She justified that red and blue had the same chance because "I switched one [points to red], one [points to blue], one [points to red]." She interpreted the numbers in the weight tool as "there's 12 out of 24 which means each color, these two have the same amount of being picked and they are both the same number." When I asked her if she could design the bag any other way so that red and blue had the same chance, she immediately said "you could click here all six [red] and here all six [blue], and half of 12 is 6." She recognized that equiprobable chances could be designed with different total amounts of marbles as long as each color had the same amount. When I changed the bag to have one red and one blue marble, she commented "they are different numbers [than the 6 and 6] but they [red and blue] still have the same chance." In addition, with all these combinations, she reasoned that for 10 trials, all three bags of marbles would probably result in "5 and 5." It seems that working with equiprobable outcomes was a contextual way for Amanda to connect the theoretical probability in the weight tool and bag of marbles, and make estimates of the expected results based on the relationship of "half" in the theoretical probability.

After running several sets of 10 trials, Amanda got many 7-3 results, a few 6-4 and only one 5-5. When I asked her to predict results for 100 trials, she said "we got three and seven a lot … We wouldn't get five and five … I don't know what the number would be, but like the three and seven, I think that would mostly happen. But in a different number [pause] 30 and 70. Yeah, 30 and 70." Instead of reasoning from the contents of the bag, Amanda extrapolated the most common distribution from the experimental data to her prediction for 100 trials. Her extrapolation demonstrates her belief about the usualness or unusualness of a distribution based on whether that distribution occurs during her experimentation. Since 7-3 occurred more often than 5-5, she used that result with a small number of trials to predict for a larger number of trials. In this regard, it seems that she conjectured expected results based on the theoretical probability when she did not have any experimental data. However, given experimental data, she based her expectations on the experimental evidence rather than the theoretical probability. I conjecture that her understanding of theoretical probability and expected results is based more on subjective notions than objective analysis of information. If she did not know the contents of the bag (theoretical probability) and used experimental data to predict further experimental results, then her judgments would have been objective. However, she is either not considering the contents of the bag (perhaps she forgot what she put in the bag), or she is basing her judgments on subjective intuitions from her immediate past experiences with experimentation.

After running 100 trials several times, the experimental evidence again produced a perturbation that led to an accommodation in her judgment of expected results. However, she needed to be focused on the contents of the bag of marbles before she made the connection.

> A: It's usually in the 50's and 40's.
> T: Why do you think that's usually in the 40's and 50's?
> A: Because 50 is half of 100.
> T: But why is being half of a 100 so special?
> A: I don't know … I think it has to do with halves.
> T: Why would this have to do with halves?
> A: See it keeps going in the 50's and it keeps going in the 40's.
> T: Uh huh. It does. And you said half was 50, half of a 100 is 50. And it seems to be around 50. Why do you think we are getting around 50?
> A: I don't know. The only guess I can come up with is because 50 is half of a 100.
> T: But why is being half of a 100 important?
> A: I don't know. [pause]
> T: All right. Open up your bag of marbles. So what do you have in your bag of marbles here?
> A: 24 … oh, 12 is half of 24!
> T: It is half of 24. So half of the marbles in that bag are red. And half of them are blue. Do you think that has anything to do with maybe why we are looking at halves?
> A: Yeah. Because 12 is half of 24. And so 12 plus 12 is 24 [and] 50 plus 50 is a 100.

Although Amanda thought half of 100 was important, she did not make an explicit connection to the bag of marbles until I focused her on the contents. She was able to base her expected results for 500 trials on the theoretical probability. She thought the results would be "somewhere near half … because there is a half to 500 and since we put half in there [points to bag of marbles], then that probably means that we are going to get around half." She ran the 500 trials and noted how close the results were to "half." I then changed the bag contents to six red and six blue and asked her what she thought would happen this time if ran 500 trials. She quickly said "same thing" because "6 is half of 12 and 12 is half of 24" and that the pie graph "would look pretty close" to the one that she had gotten in the last 500 trials. Although she was referring to the previous experimental results, her justification came from the relationship of "half" in the bag of marbles.

During the seventh teaching session, Amanda and I did some more work with the weight tool. Some of those investigations will be discussed as part of the next two themes. However, at one point, Amanda was describing her experimental results in terms of how she had designed the weights and I asked her to explain how the weight tool affects the experiment.

> T: Go ahead and open up the weight tool. Let's take a look at this. So based on how we weighted this, how does that tell the computer what to do because it seems like we keep getting more soccer balls and smiley faces? You said that you thought it had something to do with the fact that we gave them both fours.
> A: It's like first you put all of the information in here [weight tool], then it goes, I guess, into each and every one of these [points to objects in the weight tool], so then it goes into this running thing and then it comes out.
> T: What do you think … what's it telling the running thing to do?
> A: It tells what the numbers are and then he gives them … he takes them in his bag and then just throws out what it should be.
> T: What do you mean by throws it in a bag – you mentioned something about a bag.
> A: Yeah, he takes it out of a bag and then he throws it into the screen.
> T: So what does he put in the bag – what's in the bag that he's picking from?
> A: This [points to the weight tool]
> T: Oh, and this is referring to what – what do you mean?
> A: This refers to this, this, this, and that. [she points to all the numbers listed below all four objects in the weight tool.]

It seems that the link between the weight tool and bag of marbles was effective in helping Amanda make the connection between the numerical representation of the theoretical probability and it's influence on experimental results. However, during the seventh session, Amanda used past experimental results to help inform her predictions for future experimental results. There were two related instances during this session that demonstrate her reasoning.

First, she had designed an experiment with four objects and had weighted them as in Figure 7.2. She recorded the weights on a piece of paper that she kept in front of her. She then ran several sets of 10 trials and we discussed the variability in results. Several times she got the most soccer balls, once the snowflake and smiley face were tied for the most, and once she got the most blue marbles (5 out of 10). When I asked her to predict what she thought would happen for 40 trials, she said she would get "more soccer balls … because we've been getting more soccer balls." The results from the 40 trials also varied with all four objects having the most occurrences at some point (note: the last trial of 40 resulted in the most blue marbles). Amanda's reasoning about predicting soccer balls shows how she used past experimental evidence in her judgments. Given the data and how closely the objects were weighted, her prediction was not unreasonable. Her reasoning demonstrates that experimental data was highly influential on her future expected results.

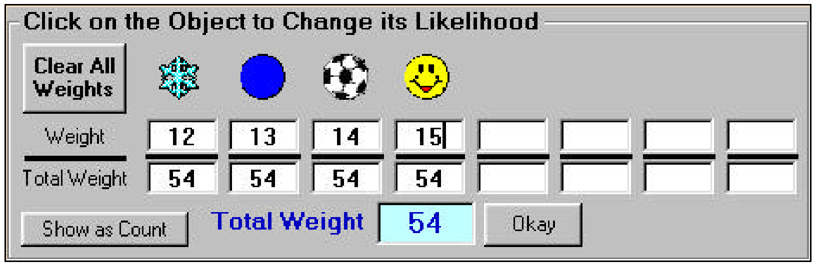


Figure 7.2. Amanda's weights for a four-outcome experiment.

To get Amanda to make some predictions and analysis with theoretical probabilities that should have different expected values, I changed the weights for those four objects to be two, two, four, and four (see Figure 7.3). She was able to reason that the snow and blue marble were equally likely, as well as the soccer ball and smiley face. In addition, she recognized that the soccer and smiley face were the most likely to occur. With the weight tool open, I asked her to make a prediction for 100 trials.

Figure 7.3. New weights for Amanda's four-outcome experiment.

T: So what do you think would happen if I would run this.
A: I don't know, you'd get some of each.
T: Some of each? Okay. Let's think about running this 100 times. What do you think is going to happen if I run this 100 times?
A: I don't know.
T: You don't know? Do you think…
A: Now it's a little bit harder to guess. [pause]
T: What do you think we will get? [pause] Instead of maybe just talking about exact numbers, which do you think we will get the most of?
A: [pause] The blue marble. Just a wild guess.

> T: The blue marble? Okay and why would you make a wild guess like that?
> A: It's not a WILD guess. I just think I won't pick soccer just to see what would happen.

Amanda was not able to initially reason from the theoretical probability to make a statement about the expected results. However, she eventually made a "wild guess" that blue would occur the most but justified her guess in a playful manner to "see what would happen." It seems that she thought soccer balls might occur more often, either based on the weights or past experiential evidence from the prior weights (see Figure 7.3), but wanted to pick blue marbles. Perhaps her choice of blue marbles was purely playful, or perhaps based on the last trial of 40 that she did with the previous weights when blue marbles occurred the most. Either way, Amanda may not always pay attention to, or remember, the weights, and appears to base her judgments of expected results on a variety of subjective and objective factors.

Overall, Amanda's interpretation and use of theoretical probability improved during the teaching experiment. Her predictions of expected results were based on both theoretical probability and her experiences with simulations in the microworld. She used part-part and some part-whole reasoning and only occasionally analyzed her experimental results with respect to theoretical probability. Her use of the marble environment and weight tool to model various situations demonstrates how the software facilitated her slightly improved conceptual understanding and helped her comprehend the electronic simulation process. However, Amanda seems to make predictions based on past experimental results more than she did with reasoning from the theoretical probabilities. It seems that experimental data was highly suggestive to her. She tended to take a more subjective approach, based on experiential data, instead of an objective analysis based on known theoretical probabilities.

Interpreting and Using a Pie Graph

A pie graph representation seemed initially unfamiliar to Amanda. In the earlier teaching sessions she had difficulty interpreting the pie graph display and was not able to make meaningful connections between the numerical data, bar graph display, and the size of the sectors in the pie graph. Whereas Carmella and Jasmine seemed to reason better from the pie graph than the bar graph, Amanda preferred the bar graph.

During the first teaching session, the students were each working on their computer simulating sets of 20 trials for a coin toss experiment. At one point, Joe asked if anyone could get a 20-0 result. Shortly thereafter, Amanda got a 17-3 and Jasmine commented that she could tell if she got a 20-0 by only looking at the pie graph because "it would be all blue or all gray." Amanda then embarked on her own experiment by continuously pressing the Run button to do many trials of 20 to see if she could get an all blue or all gray pie graph. Since she did not use the Clear button to erase the previous set of 20 trials, the number of trials was cumulative (i.e., 20, 40, 60,…200). She had divorced the pie graph display from the simulation process and the data that it represented. Even though the pie was part blue and part gray (at varying degrees early in her experiment), she still thought it was possible to fill the circle with only one color. Although her goal was inappropriate, I took the opportunity to have Jasmine and Carmella watch the pie graph as Amanda continued to run a large number of trials. Recall that this visualization was the birth of the EOP, as discussed in Carmella's and

Jasmine's case. Amanda recognized that "it's staying in the same place pretty much." But could not explain her observation. She also noted that she now realized she could never get the pie graph to be all blue or all gray when the pie already displayed both colors. However, the visualization of the dynamic pie graph during the simulation did not seem to affect Amanda's understanding of the effect of a large number of trials on experimental results as much as it did for Carmella and Jasmine. I conjecture the difference in effects was due to the students' differing levels of understanding about what a pie graph represents. Since Amanda did not understand what the pie graph represented, she could not make a conceptual connection between the simulation process and the changes in the pie graph display. Although she continued to have the pie graph open during further experimentation in this session, she did not verbalize any observations that would indicate that the pie graph was a meaning-making agent for her. In fact, I believe she was fascinated with the motion of the pie graph and did not make connections about why and how that motion was occurring.

During the second teaching session, the students did many experiments with simulating a die toss and analyzing the data with the stacking columns, data table, bar graph, and pie graph. At one point, Jasmine and Amanda got five 1's, two 3's, and three 4's. Using only the stacking columns and data table to guide them, I asked them to draw a picture of what they thought the bar graph and pie graph would look like for this data. Amanda was able to draw an appropriate representation of the bar graph, although she only accounted for four of the 1's in her graph. She had difficulty reasoning about the size of the slices for the pie graph. During her work with drawing a pie graph, Amanda referenced her bar graph that displayed four (instead of five) 1's, two 3's and three 4's, rather than the actual data shown on the computer screen.

> T: How many slices do you think are going to be in the pie?
> A: Three.
> T: Three. Okay. And why is it only going to be three?
> A: Because there are only three [kinds of] dice.
> T: What do you think about the sizes of the slices? [she shrugs her shoulders] You don't know. Can you tell which of the slices are going to be the smallest?
> A: Yeah, maybe the two [3's].
> T: Two. Okay. Which slice would be the biggest?
> A: The four of the dice.
> T: Okay. And which dice did you get four of?
> A: 1.
> T: All right, so this one would be the biggest slice? And this would be and you said this one [the 3's] would be the smallest slice. Okay. And compared to these two slices, how big would this one [the 4's] be?
> A: Um ….Between these two [1's and 3's].
> T: Okay. Why don't you just estimate the size of those slices? [long pause as she stares at her paper with an empty circle drawn on it] What do you think about the size of these slices? About how big? If you were going to draw a slice for the 1, how big do you think it would be?
> A: I have no idea.
> T: Do you think it would take up the whole pie?
> A: No.
> T: Do you think it would take up half of the pie? [pause]
> A: Um…Yes.
> T: All right. And why do you say it would take up half?
> A: because [pause] I know what I can do [she draws in vertical lines cutting across the pie graph. See Figure 7.4]
> T: Okay.
> A: Um no, that is not quite half [referring to the four vertical slices she just drew in the circle]
> T: Okay. [pause]
> A: I don't know what I'm doing.

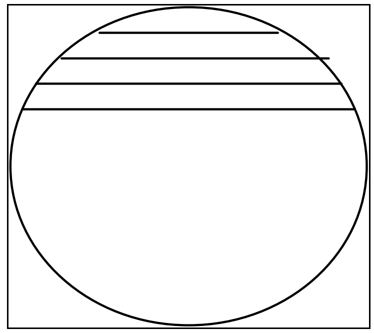


Figure 7.4. Re-creation of Amanda's drawing of four slices in a pie graph.

Amanda was able, with appropriate guiding questions, to think about some of the basic features of the pie graph (e.g., three slices, relative size of slices, four slices that constitute the slice representing the 1's). However, her sketch of the four vertical lines indicates that her understanding of how to create a pie graph seems to include making slices equidistant from each other, but that the area created by the slices is not necessarily equal. Although she spent a lot of time watching the pie graph during simulations, it seems that she had not internalized the pivotal movement around the center of the circle as the size of different slices change during the simulation process. Again, her focus on the pie graph may have been as a visual stimulant and not at all focused on the mathematical representation and interpretation of data.

During the third teaching session, Joe and Amanda did many experiments with bags containing 2B2W, 5B5W, and 3B1W marbles. Joe had Amanda interpret both the data table and pie graph simultaneously and asked her many questions to help her make connections between the numerical and graphical representations. With the 2B2W and 5B5W bags, she noted that she could tell if the experimental results were "even" by looking at the pie graph. She said that with "uneven" results, the pie graph had "bumps" in it but when the pie graph had no "bumps" on the line separating the black and white sectors, the results were "50-50" (see examples in Figure 7.5).

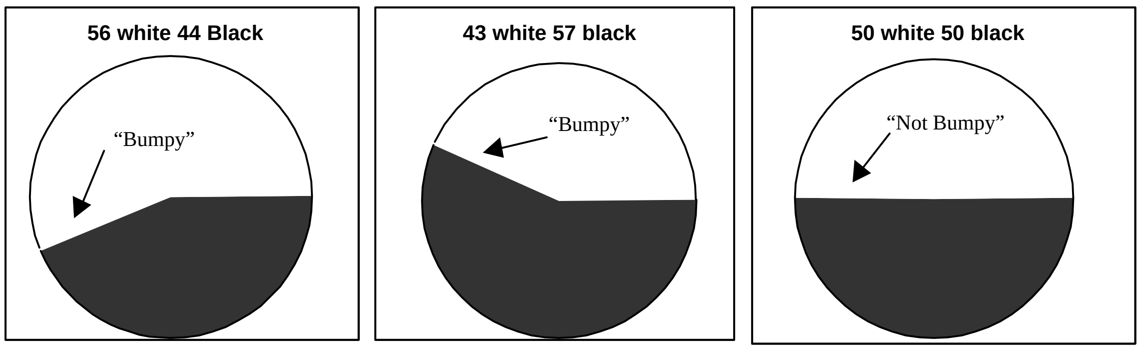


Figure 7.5. Amanda's analysis of the "bumps" in the pie graph.

Although her analysis of the "bumps" was focused on the staircase look of the line that separated the white and black sectors rather than the unequal size of the sectors themselves, she was at least making a connection between "50-50" and "half" of a circle "with no bumps." Amanda extended her analysis of the "bumps" in the pie graph with 100 trials with the 3B1W bag. When she finally got 25 white and 75 black marbles, she commented that there were no bumps in the graph and started to make a connection between the numerical result of 25-75 with the pie graph [see Figure 7.6].

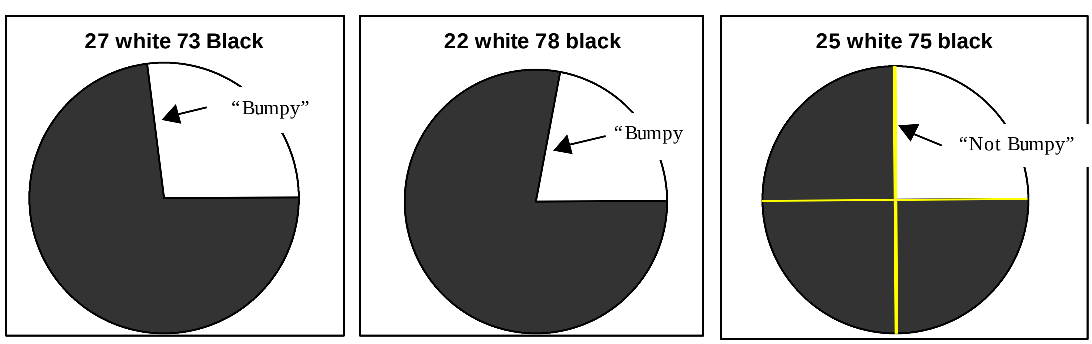


Figure 7.6. Amanda's "bumpy" analysis and connection with "quarters."

T2: Okay. Take a look [at the pie graph display of 25 white and 75 black] Now there's something else I would like for you to explain to me.
A: It's not bumpy.
T2: How do you know I was going to ask you that? That's what I was going to ask. It's not bumpy. And is it even?
A: No, but it's, there's four lines it cuts the pieces like a pie. There's one line right here [draws imaginary vertical line through middle with mouse]
T2: Okay.
A: And there's one line right here [draws imaginary horizontal line through middle with mouse]. If it's exactly on the line then it's not bumpy.
T2: Good. And why is it exactly on the line here?
A: Because there's a line right there and there's a line right there that cuts in to these little pieces right here.
T2: Now why do you think the colors match right on those lines?
A: Because 25 and 25 right here ... [pointing to top two imaginary quarter sectors]
T2: Yeah.
A: 25 and 25 equal a half, I mean yeah, a whole half. Then ... [long pause]

Although Amanda could visualize the four equal sectors in the pie graph and made the connection that each sector represented 25, she never made an explicit connection back to the 3B1W bag of marbles to explain why the results were close to the 25% mark on the pie graph. She did say that she thought black was more likely to occur and that with 500 trials she expected the pie to look similar, but she did not support her reasoning on the theoretical probability.

The visualization of four equal sectors in a pie graph must have made an impression on Amanda. After Carmella weighted a coin three to one in favor of tails in the fourth session, I asked both girls to draw a prediction of what they thought the results of 100 trials might look like as a pie graph. Carmella reasoned that if she thought about money then it would be 25 and 75. After hearing that, Amanda drew the pie graph shown in Figure 7.7 and explained that the “heads is 25.” It seems that reference to 25 and 75 may have helped her estimate the size of slices based on her memory from the quarter sectors she drew in the third teaching session.

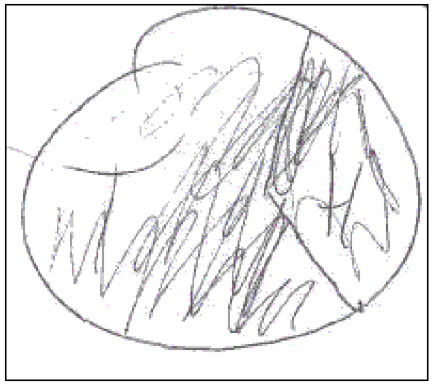


Figure 7.7. Amanda's pie graph prediction for 100 trials with a 3-to-1 weighted coin.

The quarter and half markings on a pie graph seemed to be a good reference tool for Amanda. In the sixth teaching session, Amanda was doing several sets of 10 trials from a bag with 12 blue and 12 red marbles. She had gotten several seven and three results, one 4-6 result and one 5-5 result. On the next set of 10 trials, I had her close the data table and only leave the pie graph open. She got the pie graph displayed in Figure 7.8. I asked her to predict what the numerical results were based on the pie graph. She was able to reason that it was not 5-5 since the red and blue sectors were not equal. She initially guessed seven reds and three blues based on secretly trying to count the marbles on the screen (one blue marble was hidden underneath the pie graph and was not visible).

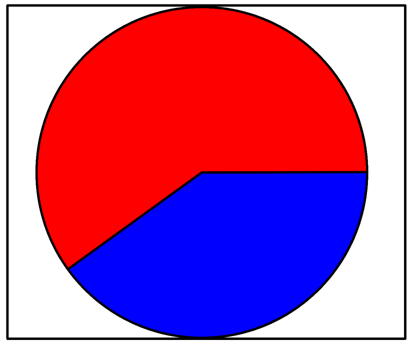


Figure 7.8. Pie graph display for six red and four blue marbles.

> A: [softly trying to count the marbles on the screen] one, two, three, four, five, six, [pause] seven reds and three blues.
> T: Seven reds and three blues? Okay.
> A: No. [tilts her head]
> T: No. Why don't you think so?
> A: Because last time it was like this [points to about the 70% mark on pie graph]
> T: You are right, the line was down there. So there's not going to be seven and three? [she shakes her head "no"] What else could it be?
> A: Six reds and four blues.

She knew that there were more red marbles and was able to use her memory of previous results of seven red and three blue to eliminate that possibility, even though she thought she had counted that many marbles on the screen.

Also during the sixth teaching session, Amanda was able to reason about the size of a sector representing "one" blue marble relative to the size of the other sector representing "two" black marbles. She did similar reasoning for results of 5 black and 1blue marble. Her complete reasoning on this task will be addressed in the discussion of her investigation of "twice as likely."

Still during the sixth teaching session, Amanda was able to extend her understanding of the size of the sectors on the pie graph to a situation with three colors, two of which had about equal numerical results. With results of 82 white, 83 red, and 35 blue, she reasoned appropriately about what the pie graph would look like.

> A: This one will be the same size.
> T: Which one?
> A: The red, I mean yeah, red and white.
> T: The red and white one will be about the same. And what about the blue one? Will it be bigger or smaller?
> A: Smaller.
> T: Smaller. Why is it going to be smaller?
> A: Because 35 is way off from 83.

She was able to use the numerical results to estimate the relationships between the sizes of the sectors. She knew that red and white would be about the same and that the blue had to be much smaller since it was "way off from 83." Amanda also made some connections between the number of trials and the number of imaginary slices in the pie graph during this session. I discuss those connections in the description of her work on the "twice as likely" task.

Another apparent reference to the half and quarter markings by Amanda occurred in the last teaching session. She was experimenting with a 3B1W bag of marbles and I asked her to predict the results for 100 trials. At first she reasoned that black would appear the most because "the three black marbles are fighting against this one white one." Since she was not able to give a numerical guess, I asked her if she could visualize what the pie graph would look like after 100 trials. At first she said the black would have "just a little" more than the white. I then drew a circle, shaded the bottom half and said "so if this equal, how much more?" She then took the pen, drew a vertical reference marking for quarters, shaded in the left-top quarter and then "a little more than that." Our shared drawing is shown in Figure 7.9. She could not give any reason for her drawing and shading besides “that’s what I think” and “it just helps me.” Although she seemed to be using four sectors as a guide, she did not verbalize any connection between the contents of the bag (four marbles) and her predicted drawing.

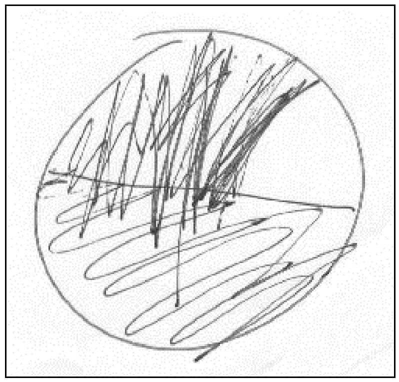


Figure 7.9. My and Amanda’s drawing for an estimated pie graph from 3B1W bag.

Several times throughout many of the teaching sessions, Amanda made reference to the pie graph “staying in the same place” during a simulation with a large number of trials. However, she was never able to explain why the pie graph was staying relatively the same beyond comments such as “the white can’t fight back any longer [in reference to her notion that with a large number of trials the white slice would fight back against the black slice].” She made only occasional references to the actual simulation process or the theoretical probability in her analysis of the motion of a pie graph during a simulation.

Amanda made some progress in understanding how a pie graph is constructed and how it corresponds to the numerical data. However, her ability to use the pie graph to reason probabilistically about chance situations and to analyze experimental data in terms of probability did not sufficiently develop. She made very little connection with the theoretical probability and resulting experimental data and graph. She also only occasionally verbalized any notions that a pie graph and data for experimentation was somewhat proportional to the theoretical probability. In this regard, Amanda's development with understanding the pie graph was purely from a numerical-to-graphical perspective and did not include ideas that helped her further develop conceptual understanding of probability.

Use of Multiplicative Reasoning

There were several tasks in which Amanda used multiplicative reasoning during the study. Recall that Amanda's early quantitative reasoning in the first three teaching sessions was primarily based on a most-middle-least scale. It wasn't until the end of the third teaching session that she started using multiplicative reasoning in her analysis of theoretical probability, expected results, and actual experimental results.

At the end of the third teaching session, Joe engaged all the girls in an experiment at the computer station where he and Amanda had been working. Amanda filled up a bag of marbles in the microworld with 23 blue and one red marble. Joe asked them each to predict how many red marbles would be picked if they ran 500 trials. Carmella predicted "not many, 50" and Jasmine predicted "52." Amanda said "no, I think you are going to get maybe 10." Carmella pressed the run button and the simulation started. Joe predicted 400 reds and the girls all laughed at him. Amanda commented "you would never get that!" Once the simulation stopped, there were 23 reds. Joe asked Amanda why she thought there would be so few red marbles. She said "there's only one red. So that means there's not going to be many reds. Not 400!" Carmella pressed the run button again and all the girls guessed around 50 red while Joe predicted 900. They all laughed again as the simulation continued and 45 reds (out of 1000) occurred. Before they had time to really discuss the results, Carmella clicked on the run button twice. Amanda started thinking aloud “let’s see, what’s 45…I think we are going to get 80.” She explained her reasoning “because 45 and 45 is about 80.” Amanda then asked Carmella if she clicked once or twice. When Carmella confirmed that she clicked twice, Amanda said, “then 90.” The girls then started talking about their brains being infected with “bugs” and how they were ready to go (this session was already 7-8 minutes over time). Joe and I never got to ask Amanda to explain her reasoning for adding 45 and 45. However, since she was concerned whether Carmella had clicked the run button once or twice, it seems that she had associated two clicks (500 trials per click) with about 45 red marbles. She then may have figured that two more clicks would add about 45 more red marbles to the data. Her reasoning indicates she may have inferred a ratio of 45 red for two “clicks” and used that ratio to predict the number of reds for an additional two “clicks” (1000 trials). Her reasoning occurred during a playful, impromptu task, and illustrates that in a non-threatening environment (she obviously felt at ease in the group exploration), she was able to construct a 45:2 ratio and employ an appropriate additive process to double the ratio in her guess of 90 reds for four clicks.

During the fourth teaching session, Carmella and Amanda had designed an experiment with two outcomes and were playfully changing the weight tool, making predictions, and running simulations. During this time, they spent some time using a variety of weights that were equiprobable and we had discussed that there were many ways to use the weight tool to model "equally likely." I then asked them if they could use the weights to make one outcome twice as likely to occur as the other.

> T: Let's do something. How would you weight it so that the soccer ball was twice as likely to occur as the shape?
> A: What do you mean by twice as likely?
> T: That if we do it I would get twice as many soccer balls as the shape. [Carmella changes the weights to two and 1] So you think two and one?
> C: Yes.
> T: Is that the only way to do it?
> C: No.
> A: You could do four and two … Eight and four … Sixteen and eight. [As she stated these weights, she changed the display in the weight tool accordingly.]
> T: Sixteen and eight. Would they all get the same results?
> A: No.
> C: Close.
> A: But they would all be twice.

As soon as Carmella established the ratio of two to 1, Amanda was able to generate three other examples of "twice." It seemed that she did not initially understand how to set up a "twice as likely" relationship, but once established, she knew to keep the relationship in her other examples. It is interesting to note that each time she made another example, it appears that she doubled the largest weight and used that answer as the largest weight in the next example.

During the fifth teaching session, Amanda employed multiplicative reasoning in her analysis of experimental data displayed in a bar graph. I had secretly weighted three outcomes as two, two, and one, and she had run 1000 trials in her attempt to gather data to figure out the weights. The results were 399, 397, and 204. She spent some time looking at the data displayed in the bar graph (Figure 7.10) and noted that she thought "these two are the same and this one is less." She then said that the green bar was "half way up the graph … see, 200 is half of 400." Based on her analysis, she predicted that the weights were 55, 53, and 32. She explained that “55 and 53 are about the same” and “32 is way less.” Although she was able to recognize the half relationship in the bar graph, she did not necessarily apply that relationship directly in her prediction of the weights. She may not have applied the relationship directly because she chose a difficult number to take half of (55) for the first weight. If that was the case, then 32 was a good estimate of half without having to actually divide 55 by two. Amanda may have also not have been thinking about the half relationship at all and merely chose 32 since it was considerably less than the weights of 55 and 53.

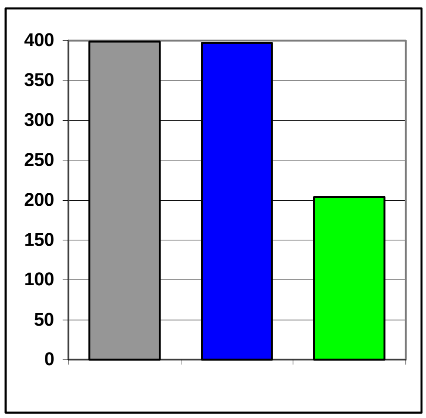


Figure 7.10. Bar graph display from experiment with secret weights of 2-2-1.

During the sixth teaching session, Amanda used multiplicative reasoning in the “twice as likely task,” which will be discussed in the next section. She also used multiplicative reasoning when I asked her to use three different colors of marbles in the bag so two of the colors were equally likely and those colors were twice as likely to get picked as the third color. She designed the bag of marbles in Figure 7.11 and explained that white and red were equally likely because “the numbers are four and 4” and they were twice as likely as blue since there were two blue marbles and “2 is half of 4.” She was easily able to make this bag of marbles and use appropriate multiplicative reasoning between each of the colors.

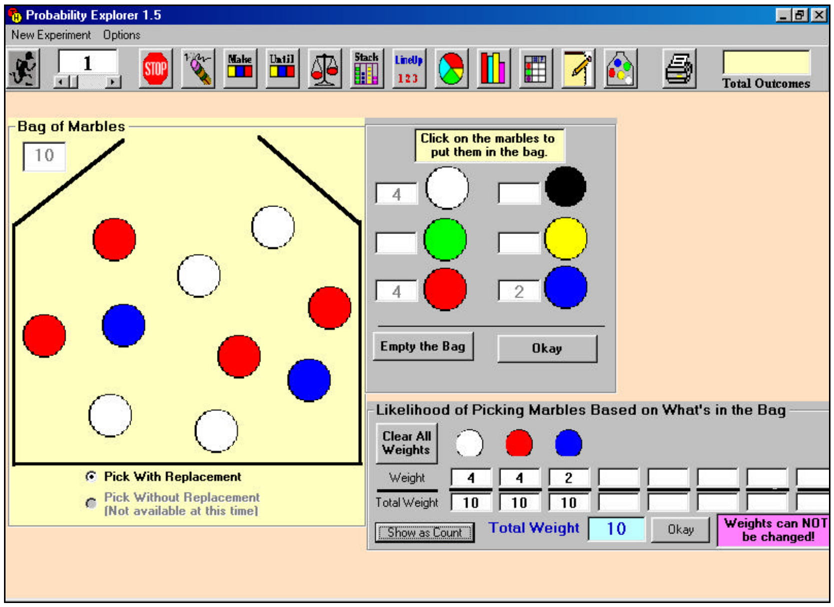


Figure 7.11. Amanda's design of two colors equally likely and twice as likely as the third.

With the bag of marbles and the weight tool showing (see Figure 7.11), I asked Amanda to describe how the chance of picking a blue marble compared with the chance for a white marble.

> T: How does the chance of picking out a blue marble compare with the chance of picking out a white marble?
> A: This will be easier to say for me. Let's say this was 40 [pointing to the weights of 4], that was 20 [the weight of 2]. There's going to be 20 more for the 40, this was Lydia [white], this was Carmella [red], and this was me [blue]. No, Jasmine [white], Carmella [red], and me [blue]. Jasmine gets 20 more pieces of candy. And Carmella gets 20 more pieces of candy. And I only get 10 more pieces of candy.
> T: Oh. So you said 20, 20, 10. Why did you say 20, 20, 10 when this said 4, 4, 2.
> A: I kind of messed up.
> T: Well, I'm just kind of curious.

> A: I don't know. I was saying 40 then 20 and I didn't want to go backwards and do it over again. So I just knew 10 was half of 20 so I just said that.
> T: So why was it important that 10 was half of 20?
> A: Because if you said something else besides 10 it wouldn't be half of 20.
> T: But why do I need half of 20 for the blue here? Because you said 20, 20, 10 [pointing to the white, red, and blue marbles in the weight tool]. Why do I need half here with the blue?
> A: Because two is half of 4.

She used a playful explanation to compare the chances, but the numbers she used in her explanation maintained the equally likely and twice as likely relationships. Not only could she set up the desired ratio, but she was able to maintain the ratio in her example and justify the use of 10 and 20 based on the original relationship in 4-4-2.

Since she seemed to be able to reason well with a twice as likely relationship, I posed another task for her containing a twice and "three times" relationship. She designed a bag of marbles with three blue, six green, and nine yellow marbles and explained how she made the bag.

> A: The yellow is twice as likely as the green. And the green is twice as likely as the blue … Because, I was adding by three. Adding up by three. Three equals three is six. And three plus three, no six plus three is nine.

Her use of adding three more to each weight was a good strategy. However, she reasoned that yellow was twice as likely as green and did not recognize that yellow was three times as likely as blue. However, when I asked her why yellow was twice as likely as green, she quickly cleared the bag of marbles and put in four white, eight green, and 12 yellow, and noted "that's just the same as three, six, and nine, but just different numbers." She went on to explain that green was twice as likely as white but then wanted to change the number of yellow marbles to 16.

A: Oh, that has to be 16.
T: Well why do you want it 16.
A: Because you have to double four plus four. Okay you have eight and you have to double it once more so it will be the same as [points to the weights of four and 8].
T: Oh, so if this was 16, then how …
A: Each of them would have, I don't really know how to explain it…
T: I think you are doing a good job. So with the 16, why were you thinking 16 here?
A: Because four and four is eight. And eight and eight is 16. So I have to double that.
T: Oh, so if you double that, how would the chance of green compare with the chance of yellow?
A: The same as white. I mean, the same as white and green.

Although Amanda had a "three times" relationship in both her examples, it seemed my question as to why she thought yellow was twice as likely as green changed the focus of the task and her goal became to establish that twice as likely relationship. It is important that she recognized the similar ratios in the weights of 3-6-9 and 4-8-12. She also recognized that making the yellow have a weight of 16 would establish a relationship the "same as white and green." Although her goal was different than my intended task, she reasoned well with maintaining the twice relationship. She may have recognized that her original iterative strategy created a three times as likely relationship between three and nine as well as four and 12, but she never verbally made that connection.

Amanda's multiplicative reasoning has all been done with relationships involving "half" and "twice," depending on how she thought of the task. She was not able to use multiplicative reasoning in the last teaching session when I had her compare a 3B1W bag to 6B2W bag of marbles (these are the same pictures of bags used in the pre-interview). Her task was to choose which bag of marbles she would rather design in the microworld if her goal was to pick out a black marble. Even though 6-2 was double 3-1, the ratio of

3-1 is not in the familiar “twice” relationship. She chose the 6B2W bag because “it has more blacks.” Even when I had her make each bag and look at the weight tool, she still maintained her choice of 6B2W because “more blacks give it a better chance,” which is contrary to the reasoning she used in the pre-interview about minimizing the undesired events. Any multiplicative reasoning that she used with the marbles and weight tool in the microworld did not transfer to this comparison.

The evidence suggests that Amanda was able to employ multiplicative reasoning with relationships of “twice and “half,” but used additive reasoning for other relationships. Her multiplicative structures were based on an additive process of “doubling.” She used both additive and multiplicative reasoning in her work with the “twice as likely” task. The description of her reasoning with this task provides further evidence of her multiplicative reasoning and how it relates to her development of probabilistic reasoning.

Amanda’s Investigation of “Twice as Likely”

Amanda’s work with this task occurred during the sixth teaching session. Since we were working with the marble bag and weight tool simultaneously, I decided to pose the “twice as likely” task to her in reference to marbles in a bag rather than in terms of the numerical weights as I did with the other girls.

> T: We are going to come back to the marbles. Open. All right. And I want you to put two colors in here. And this time I want it so that one of the colors is twice as likely to get pulled out than the other color.
> A: Okay. [She puts in 22 black and two blue marbles] Nope. [She clears the bag then puts in 10 black and five blue marbles] [see Figure 7.12.]

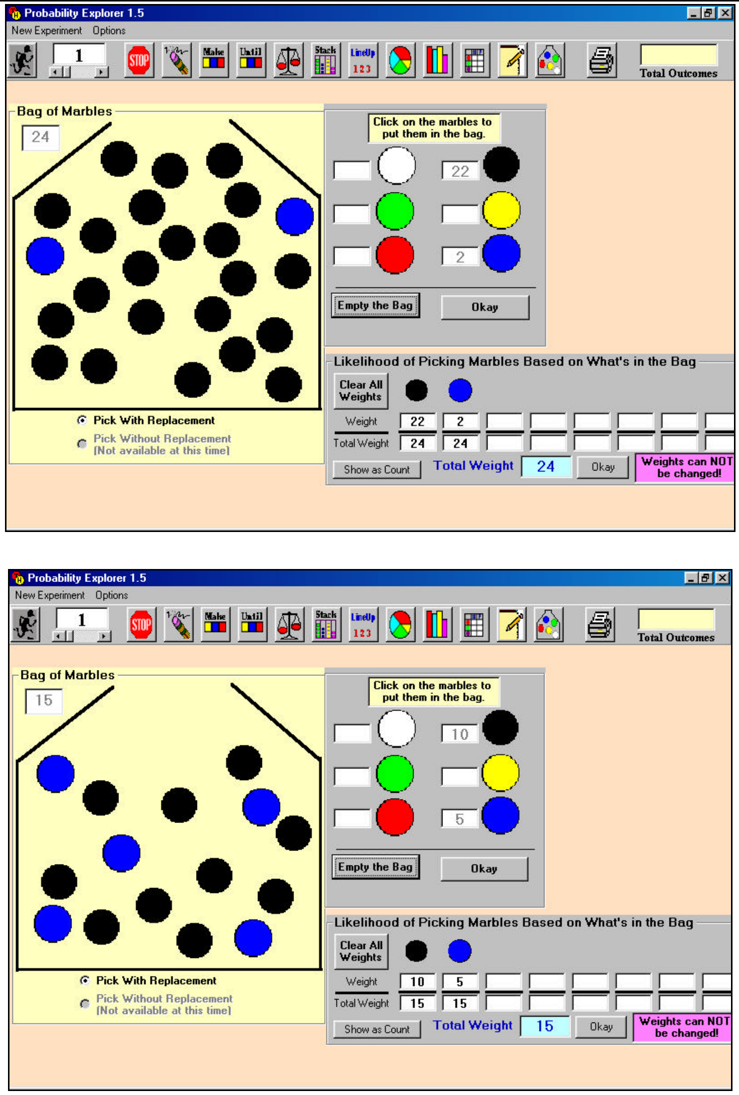


Figure 7.12. Amanda's additive and multiplicative interpretations of "twice as likely."

T: You've got 10 black marbles.
A: Is that right?
T: Ten and 5. So which color is twice as likely to happen?
A: The 10.
T: The 10. So the black is twice as likely to happen than the blue?
A: No.
T: No. Why not?
A: That's just half. No, I think you ARE asking for half of the 10.

Amanda seemed to use an additive approach when she designed the bag with 22 black and two blue marbles, but she quickly realized her error and used multiplicative reasoning to construct the bag with 10 black and five blue marbles. Although she initially was not sure that half of 10 fulfilled the “twice as likely” relationship, she was able to reconcile her doubt. I then engaged her in a simulation with the bag of marbles and a subsequent analysis of the results. She chose to run three trials and predicted “two blacks … No all blacks.” The experiment resulted in two black and one blue marbles. After she stacked the three marbles, I asked her to predict what the pie graph would look like.

> T: What do you think our pie graph is going to look like?
> A: There’s going to be black covered up and then there’s going to be this little tiny space.
> T: What kind of space? About how tiny? [She uses her fingers to estimate about a 15% slice.] That tiny? [I imitate her finger positions.]
> A: About that tiny.
> T: That tiny. All right. And why is it going to be that tiny?
> A: I guess that stands for one.

Her estimation of the size of the pie slice indicates that she was not considering the relative size of a “one” with only three trials. However, once she saw the pie graph, she was able to make the connection between the whole and each of the slices representing “one.”

> T: Okay. [She opens up the pie graph and looks surprised.] Oh, it doesn’t look so tiny does it?
> A: No.
> T: Why is it that [blue slice] so big? We only got one blue?
> A: Because if you cut this in half [She draws an imaginary line that cuts the black slice into two equal slices. See Figure 7.13]
> T: Yeah.
> A: That’s [first imaginary black slice] as big as that [second imaginary black slice], and that’s as big as that [blue slice]. And that’s two [points to the whole black slice]. And so that’s [blue] one. I guess they cut it up to be, they cut this, I guess it depends on how, on what the number is. So if it’s [the whole pie] 12, then this [blue slice] would be smaller to fit in more space for the other 11.

T: You are exactly right. It depends on that number that we are doing right there with the three … So if we were drawing that line in there, how many slices of pie would there be?
A: Three.
T: And how many of them would be black?
A: Two. And one of them would be blue.

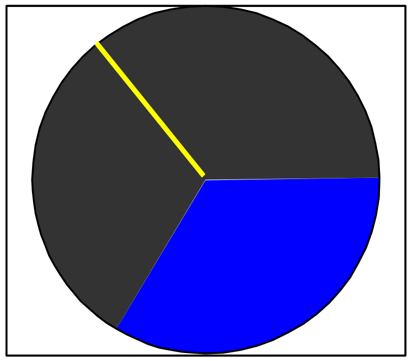


Figure 7.13. Amanda's construction of thirds in a pie graph.

Amanda's construction of the three equal slices demonstrates how the pie graph helped her make a connection between the parts and the whole. She extended this reasoning when I asked her to think about what would happen with six trials.

T: So if you do six, then how many slices would this pie graph be split into?
A: It was like that first. [using her fingers to create an estimated 1/3 slice.] So it's going to be half of that [She makes her finger slice smaller by about half.] Because three is half of six. And we chose six now.
T: So are the slices going to be smaller or bigger?
A: Smaller.
T: Very good. So let's Run six of them. Five and one. So here is one slice.
A: And half of that [draws imaginary line where the blue slice ended was with three trials] would be right there, which was half of it last time. So I guess. So that [blue] was half.
T: Oh, so this piece right here [blue] is half of what we had before?
A: Uh huh.
T: So how many slices do we have over here in the black area? Can we imagine?
A: [She draws in the imaginary lines as in Figure 7.14.] Four. Five.
T: Five. Yeah. Because I think we had five blacks and one blue.

A: Yeah. Five.

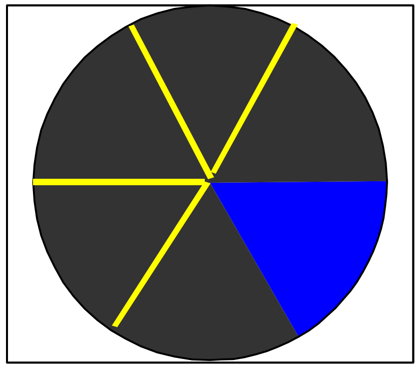


Figure 7.14. Pie graph of five black and one blue marble with Amanda's imaginary lines.

Amanda was able to use the pie graph as a tool for making a connection between the relative size of one slice compared to the total number of trials. The imaginary lines helped her reason about the size of the slices from a part-whole perspective. She also used multiplicative reasoning in estimating the size of a 1/6 slice by making it half the size of a 1/3 slice.

At this point she did not make any connection between the "twice as likely" relationship and the experimental results. She fully expected to pick out more black marbles and was surprised when once all blue marbles occurred with three trials. I asked her to predict the results for 100 trials.

T: So we have 10 blacks and five blues. So what's the chance of picking out a black?
A: More likely to pick out a black. But the last time we got three blues.
T: We did get three blues. Do you think we could get a 100 blues?
A: No.
T: No. Why not? Is that impossible?
A: Yes.
T: It is impossible. Is it impossible to get all blacks?
A: Yeah.
T: Why do you think it's impossible?
A: Because there are blues in there and you are going to pick out it a 100 times. So you are going to have to pick out a blue and a black.

T: Oh, okay.
A: So you are probably going to pick out blacks and blues.
T: Probably pick out blacks and blues. Okay. Let’s hit Okay. And let’s actually Run this a 100 times.
A: Okay.
T: If we run this a 100 times do you think we’ll get 50-50, like we were getting before?
A: Maybe once.
T: What do you think we’ll get? Why don’t you think we’ll get 50-50? You said maybe once.
A: Because 5, there’s not 10 and 10.
T: Oh, it has to be 10 and 10 in order for us to get …
A: 50-50.

Notice that Amanda’s judgments about chances and expected results are based on “more blacks” but also on experiential data (e.g., three blues). I think her judgment of “impossible” for all blue or all black marbles is also based on her experience with simulations. I think she does realize that it could happen, but to her, the probability is so small that it might as well be considered impossible. She also was able to reason appropriately from the contents of the bag that she did not expect “50-50” since it was not “10 and 10.” However, she was not able to ever give a numerical prediction of what she expected for the 100 trials. After 100 trials, by only looking at the pie graph, Amanda guessed that the blue slice was “30 pieces” and the black was “70 pieces.” The actual results were 68 and 32. We did several sets of 100 trials, but Amanda never made any connection between the experimental results and the “twice as likely” relationship between the marbles.

Although she was able to create a numerical “twice as likely” relationship, she did not use this relationship to explicitly make predictions for experimental data. The pie graph was a useful tool to help her construct a part-whole relationship to analyze the size of one slice. She was able to reason that as the number in the whole increased, the size of one slice decreased. The experimentation with the bag of marbles did not prompt Amanda to make any connections beyond “there are more blacks” to the theoretical probability.

Summary of Meaning-Making Activity in the Microworld

Amanda enjoyed using the tools in the microworld, but seemed more motivated to use the tools for playful activities and visual stimulation. She needed focused questions and specific tasks, mainly posed by either teacher/researcher or the other children, to engage her in meaning-making activity during the teaching sessions. There were only a few occasions when Amanda initiated an experiment or a “what if” exploration. She had initial difficulty learning how to use the various tools in the microworld (e.g., stacking columns, graphs, data table, weight tool). Amanda needed many experiences with the tools to make connections between them and to understand how they were related to the simulation process. Recall that she also had trouble verbalizing her reasoning and was quick to respond “I don’t know” in situations when she was frustrated.

Despite her initial difficulties, Amanda did make some progress during the teaching sessions and was able to use the microworld tools in meaning-making activity, although her engagement was almost always prompted by external questions. The vignettes described within the themes – theoretical probability, expected results, use of a pie graph, and multiplicative reasoning—capture the essence of her meaning-making activity, and suggest that the tools in the microworld facilitated her mathematical thinking when guided by a teacher/researcher, but resulted in very little development of probabilistic reasoning.

The connection between the weight tool and bag of marbles helped Amanda develop a somewhat better understanding of theoretical probability. The actual simulation process in the microworld facilitated Amanda's subjective reasoning about expected results based on experiential data. At times this reasoning was consistent with the theoretical probability, at other times her reasoning seemed to be wholly based on the experimental data. In this regard, she rarely made a connection between experimental results, the variability or stability of those results with a small or large number of trials, and the theoretical probability.

Her experiences with the multiple representations (e.g., iconic data on screen, data stacked in columns, pie and bar graph, data table, bag of marbles, weight tool) helped her better understand the pie graph representation of numerical data. Although Amanda made progress understanding the pie graph representation, she did not use the pie graph to analyze experimental data in terms of theoretical probability. Only on occasion did she indicate that a pie graph and data for experimentation was somewhat proportional to the theoretical probability. Her development with understanding the pie graph did not necessarily help her further develop conceptual understanding of probabilistic ideas.

Amanda's use of multiplicative reasoning only occurred in situations that involved "half" or "twice" relationships. She used multiplicative reasoning in several types of situations: 1) predicting results of an experiment with two equiprobable outcomes (five and five since five is half of 10); 2) comparing sizes of the bars (2-2-1 "secret weights" task) and slices ("twice as likely" task) in the bar and pie graph; 3) maintaining a ratio for predicting future results based on past results (e.g., predicted 30 and 70 based on prior results of three and 7); and 4) maintaining a ratio in establishing several weights that have the same chance (e.g., 3-6-9 and 4-8-12). Every time Amanda employed multiplicative reasoning, she was engaged in a task where she was relying on more than one tool in the microworld. Multiple representations, particularly numerical and pictorial, seemed to help her construct and maintain simple multiplicative relationships.

As noted earlier, when Amanda was not guided by directed tasks, focused questions, or experiments initiated by other students, her use of the microworld tools appeared to be more pleasurable actions. The majority of her meaning-making actions occurred during sustained interactions with one of the researchers/teachers. However, during these sustained interactions, the microworld tools helped her make sense of situations and somewhat further develop her intuitions, based both on subjective and objective reasoning, about probabilistic ideas.

Post-Interview Analysis

Amanda's post-interview was held on October 27, 1999, three 1/2 weeks after her last individual session (see Appendix G for post-interview protocol). Her work during the interview and my analysis of her responses are organized by the different tasks.

Cubes in a Bucket

Similar to the pre-interview, I asked Amanda a series of questions using a bucket containing six green, four red, and two yellow cubes. She used strict part-part reasoning to determine that green was most likely "because it has the most amount of cubes" and yellow was least likely “because it only has two cubes.” After I randomly chose a green cube and then replaced the cube in the bucket, I asked Amanda if I was more or less likely to pick another green next time. She said, “no, it has the same amount, but if you take it out, you’ll have a different amount and it sort of changes it but not much.” She did not let a previous event influence her analysis of the chance for picking a green and noted that the contents of the bag had not changed. In addition, she reasoned that the chance would change “sort of” if I would not replace the green cube. When I asked her to use numbers to describe the chances, she used part-part statements like “6 out of 6” for green because “there’s six greens and there are two yellows and four reds, so that makes it 6.” I then asked her to interpret her statement.

> T: So if I say you have a six out of six chance of getting the green, what does that mean to you? If I were to have just told you – I wouldn’t have showed you the bucket – and I said you have a six out of six chance of picking out the green, what does that mean?
> A: You have an even amount of getting both if there were two colors, but there’s three colors, so you’re most likely to pick out green because it’s most.

She also used a part-part statement to describe the chance for yellow as “two out of six” because “there are six greens and two yellows mostly there’s green because it’s still obviously more.” In this description, she did not account for the four red cubes in the bucket. It is as if she interpreted my question as a comparison task between green and yellow rather than a statement of chance.

<u>Coin Tosses</u>

Two types of tasks involving coin tosses were used in the post-interview. The first task assessed concepts of independence while the second task assessed concepts of fairness and the law of large numbers from a frequentist perspective. Since she had only used part-part statements to describe the chances in the last task, I first asked her to describe the chances for heads and tails on a coin.

> A: You would have one out of one … It means there's only one sign on one side and one sign on the other.
> T: Okay, so the chance of picking a head is what?
> A: No, one out of 2.
> T: Why is it two?
> A: Because there are two sides.

Although she initially used a part-part statement, she was able to self-correct her reasoning based on the whole of "two sides."

<u>Tossing a coin six times.</u> I showed Amanda four possible sequences of results from flipping a coin six times (HHHHTT, THHTHT, THTTTH, HTHTHT) and asked whether any of the results are more likely to happen than the others. Amanda first asked me "you really did this?" and I replied "yes." She then said that THHTHT was more likely to occur because it was "more mixed." When I asked her why it was good to be "mixed," she replied "I don't know, because it's normal … but anything could happen because you obviously got them." It was important for Amanda to know whether the sequences in the task were actual data, and that obviously informed her judgment of "anything could happen." This again demonstrates how she judges "usualness" on whether something has occurred or not (recall her judgments made during the teaching sessions). She also believed that "mixed up" results were better and "normal" and therefore more likely to occur. Although I emphasized the exact order in these sequences, Amanda may have been overgeneralizing to all combinations that are "mixed" are more likely as a group.

When shown the next set of results (HHHTTT, HHHHHH, THTHTH, HTHTHT), she said HHHTTT was more likely but could not give a reason for her judgment. I then asked her to compare HTHTHT and HHHHHH. She said "it (HHHHHH) could happen but I think that one (HTHTHT), but it wouldn't happen as much as that one (HHHTTT)." She seems to favor more "mixed" results, and although she did not say this, she may think HTHTHT is too regular. Nevertheless, she did not express any ideas about independence in her analysis of either of these tasks.

For the final question about independence, I asked Amanda if, after flipping a coin and getting the results HTHHHH, I was more likely to get a heads or tails on the next flip. She promptly stated "heads" because "they are on a roll." Although she seems to be ignoring the independence of events, her answer is consistent with the reasoning she has done in the past about making predictions based on past experimental data.

<u>Is this coin fair?</u> The intent of this task was to assess whether Amanda could reason from a frequentist perspective about the fairness (i.e., equiprobability) of a coin when given experimental results. To begin the task, I asked her to describe what it means for a coin to be fair.

> T: If I want to know if a coin is fair, what does that mean?
> A: I don't know.
> T: Okay. What does it mean for anything to be fair?
> A: It means to… if that person gets that chance, then you should get that chance unless it's that – you can get that chance.
> T: So what would it mean if we were going to flip this coin, how could be describe whether or not this coin was fair?
> A: It has two sides and they're both different unless you get a trick one.
> T: Okay, so they're both different so one side is a head and the other side is a tail. And how would we know if this coin is fair or not.
> A: By looking on the sides.

Amanda’s assumption is that all coins with a different symbol on each side are considered fair. Her understanding of a fair coin is that each side has an equiprobable chance of landing face up and as long as there are two different possible outcomes, the coin must be fair. Her definition of a fair coin is important in the analysis of her responses to the next series of questions.

In the first situation, I told her that I flipped a coin 10 times and got eight heads and two tails. When I asked her if she could tell if the coin was fair from the data, she immediately said “no … actually, you can tell it is fair because you can see there’s a tails and it has a head, so know that it’s not a trick coin.” Her reasoning was consistent with the definition she gave of a fair coin.

In the second situation, I told her that I flipped a coin 100 times and got 41 heads and 59 tails. She initially said “yes… they’re pretty close, the numbers are pretty close. The coin’s not cheating.” But when I asked her why it was important that the numbers were close, she could only reply “I don’t know.” It seems she had an intuition that a fair coin would result in about an equal amount of heads and tails, but was not able to verbalize her reasoning.

I then asked her what she thought about getting 175 heads and 325 tails with 500 trials. She seems to have an intuition that the “closeness” of the results can tell her something, but she is unsure of how to explain it.

> A: I can’t really tell how the coin’s going to be fair by the numbers.
> T: Okay, why not?
> A: Well, the coin itself, yeah the numbers can.
> T: And what about the numbers?
> A: If they’re close, far apart, in the middle.
> T: So could you tell me if the numbers were close? Would you consider these numbers close or far apart?
> A: Far apart.

> T: What do you think if we got results that are far apart, does that tell us anything about the coin?
> A: Would you say that again because I lost it?
> T: That's okay. If we flip a coin and we get these results – we do it 500 times and these are the results that we get – based on this, can we tell anything about the coin that we started with whether it was fair or not?
> A: I don't know.

She continued to say that it would be really unlikely to get all heads, especially "the higher the number, the lower the chance you're going to get all one." She seems to have an understanding about the effect of the number of trials, but is not able to make a direct connection between the number of trials, the "closeness" of the results, and a judgment of fair. This is probably due to her original definition of a fair coin as well as inexperience with real coins that are unfair. Nevertheless, she is expressing a notion that she thinks the data can tell her something, she is just not certain of exactly how to interpret the data. This is consistent with her experiences in the microworld. She would make statements about experimental data but not be able to connect the results back with theoretical probability.

Marbles in a Bag

As in the pre-interview, I presented Amanda with pictures of four pairs of bags containing black and clear marbles. Two of the pairs were in proportion to each other while two pairs were not proportional. When presented with each pair, I asked Amanda to determine which bag she would prefer to pick from, or if it mattered which bag, if she wanted to choose a black marble.

When presented with Bag A (3B3C) and Bag B (1B1C), she said "it doesn't matter which bag" and continued to explain her reasoning.

> A: Because there's one black here and one white there, so you have the same amount of chance of getting a black or white. There's three blacks here and three whites here, so you can't tell if there's going to be a black if you're going to pull out a black or a white one … Like Carmella says, 50-50.

She was easily able to recognize the equiprobability of picking out a black marble from each of these bags and even labeled the chance as "50-50" and referenced Carmella as originally stating equal chances with that language.

When shown Bag C (3B1C) and Bag D (5B2C), she chose bag C because "there's not a lot of blacks but there's only one white, so that puts the black in a higher spot because there's more black but not more black in bag D, but there's more whites [in bag D] and there's less whites here [in bag C], so I'd choose bag C." Her focus was on the amount of white marbles in the bag. Her comparison was done strictly as part-part and she focused on the undesired part as a deciding factor. In addition, I asked her to describe the chance of picking out a black marble in each bag. She used correct part-whole language and stated "3 out of four because there's three blacks but if you count the white, there's four marbles in the whole bag" and in bag D she noted the chance as "5 out of 7."

The next two bags presented, bag E (2B1C) and bag F (4B2C), were in proportion to each other. Although these bags each had a familiar "twice as likely" relationship, Amanda picked bag F and justified her reasoning based on a reference to 50-50.

> T: Bag F? Why do you want bag F?
> A: Because there are more blacks and less whites, in here [bag E] it's almost 50/50, so…
> T: So why is it almost 50/50 over here [in bag E]?
> A: Because two is the closest to one and three and there's three in the bag and two blacks so I'd choose this one [bag F] because the numbers aren't as close as this one [in bag E].

She did not employ any type of multiplicative reasoning in her comparison and instead based her judgment on an additive comparison to “50-50” and the additive difference between the parts in each bag. Although using “50-50” as a comparison base is an appropriate strategy, she needed to consider the relative difference from 50-50 rather than an absolute difference to make the strategy appropriate.

For the last pair of bags, bag G (2B3C) and bag H (5B6C), she again based her judgment on a comparison to “50-50.”

> A: [she quietly counts the marbles in each bag] Certainly not choose that one [bag H].
> T: Certainly not that?
> A: [She recounts marbles in bag G.] And not that one. [bag G]
> T: And not that one? Hmm…
> A: [She stares at the bags for about 20 seconds] This one [bag H] has more blacks than that one [bag G] and in this case I like it 50/50.
> T: You like it 50/50. Is this one 50/50?
> A: No.
> T: No? So how would you describe the chance of picking a black one over here [bag H] if it’s not 50/50?
> A: Five out of eleven.
> T: Five out of eleven, okay and what about over here [bag G]?
> A: Two out of five.
> T: Two out of five, okay. And so you want bag H? And tell me again why you wanted bag H.
> A: Because there’s more black and it’s closer to 50/50.

Although she is correct that bag H is closer to “50-50,” her reasoning is based on the number of blacks and she did not verbalize any reference to the total number of marbles. Altogether, she used strict part-part and additive reasoning to answer these tasks. She did not use any multiplicative reasoning nor did she explicitly consider the total number of marbles in each bag.

Constructing Sample Space and Theoretical Probability

The only sample space question used in the post-interview was for a three-event experiment. The context for this task was a family with three children, ages 9, 5, and 3. I asked Amanda to list all possible arrangements of boys and girls with respect to their ages. At first, she misunderstood the task, but once I focused her on the arrangements of the three kids from oldest to youngest, she said “oh, there’s nine different ways” and started listing possible combinations. The first six arrangements that she listed, in order, were GGG, BBB, GBB, BGB, BBG, and GGB. When I asked her to prove that she had all the possible arrangements, she systematically restated her list in the order that she made it. I began to ask her a question about the chance of a specific family when she realized she could make GBG because it was “buddies” with BGB (see Figure 7.15). She continued to pair the arrangements up as “buddies” and realized that GGB needed a “buddy” and wrote BGG. She explained that some of her pairs were “buddies” because “they are reversed” (e.g., BBG and GBB, GGB and BGG) and others were “buddies” because “they are the same pattern” (e.g., BGB and GBG, GGG and BBB). She further reasoned that she had found them since they all “have to have a buddy.” Even though she had a reason for each of the buddy pairs, since the reason was not consistent, she may not be able to generalize her approach to more difficult situations. However, the approach was sufficient for this task and demonstrates that she was at least trying to use a systematic approach to justify that she had made all possible arrangements.

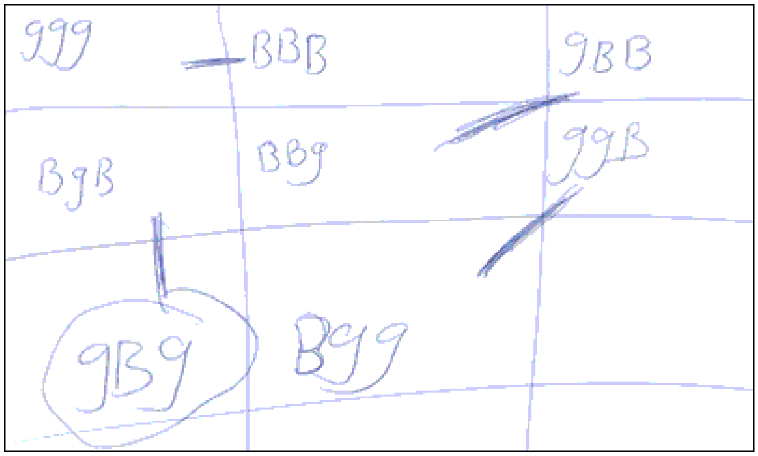


Figure 7.15. Amanda's list of all possible family arrangements.

To assess her ability to determine theoretical probabilities from a sample space, I asked her a series of questions regarding the chance of the actual family arrangement being certain arrangements or combinations of boys and girls. I first asked her to describe the chance that the family had all three girls. She replied "natural" because "I have a friend who has two other sisters, so they're all girls." When I asked if she could use numbers to describe the chance, she said "three-three, no three out of zero." When I asked what the zero stood for, she replied "boys." She only focused on the GGG arrangement and did not consider that arrangement as one out of the eight possible arrangements she had listed. Instead, she reasoned strictly by comparing the girl parts (3) and the boy parts (0). She used similar reasoning for the other questions I asked her.

T: What would be the chance that the family was girl-boy-girl?
A: Natural.
T: Okay and what about with numbers? What kind of number would describe those chances?
A: Two out of three.
T: Two out of three? Alright and where do the two and the three come from?
A: There are two girls and one boy but all together there's 3.
T: How would you describe the chance that the family had two boys and a girl in any order?

A: Two out of three.

Her sense of "natural" probably comes from her personal experience and that it is possible to have each of those families. Her description of the chances for the GBG family was based on that specific arrangements, but this time she considered the girl part (2) to the total number of kids (3). She did the same with two boys and a girl in any order. Amanda was not able to reason about each of the arrangements of three kids as being one of the eight possibilities. She could only reason from the individual arrangement and used both part-part and part-whole reasoning to state the chances. She did not seem to make any connection with the eight possible arrangements as constituting the whole. Thus, her reasoning from a sample space made up of compound events is weak. She reduced each of the arrangements to a sample space of three individual events.

Using Results to Design Experiment

The next two tasks were used to assess her ability to interpret and use information from both a pie graph and bar graph. I told Amanda that I had designed a bag of marbles in the microworld and ran an experiment. I showed her a graph of experimental results and asked what she could tell me about the bag of marbles.

Reasoning from a pie graph. For the first task, I showed Amanda the pie graph in Figure 7.16 (left-hand picture) and asked her if she could tell me how many times I ran the experiment. She noted "maybe 500, maybe 50." I then asked her what she could tell me about the bag of marbles that I designed.

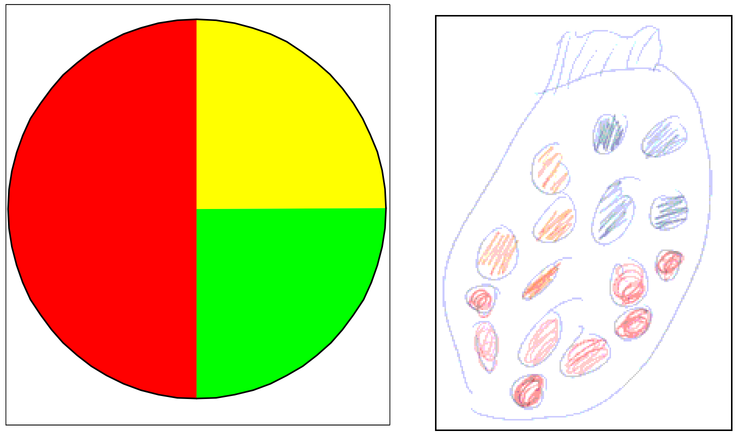


Figure 7.16. Given pie graph and Amanda's drawing of a possible bag of marbles.

A: There are more reds and an equal amount of chance for green and yellow.
T: So you said there are more reds and equal amount of chance for green and yellow?
A: Yeah.
T: What about the pie graph tells you that?
A: That this is half.
T: What is half?
A: This – this is half [traces the line down the middle].
T: Oh, this line right here?
A: And these two [points to green and yellow sectors] are half of the half.

She easily recognized the "half" relationships in the pie graph. She then used this relationship to make the bag of marbles on the right-hand side of Figure 7.16. She first drew 12 marbles in her bag and colored four of them blue, four of them orange, and four of them red.

T: Why do you think this bag would give me these kinds of results?
A: These two are half of the half that's 8.
T: Which two – the orange and the blue?
A: Yeah.
T: They are half of a half. Okay, so you said these are eight, so each of them are four. [She then draws in four more marbles in her bag and colors them red.] How come you're doing that?
A: Because it needs to be half of a half – I have to make these [red] 8.
T: Oh, okay. So you need eight reds?
A: Yeah, because that's [points to four orange marbles] half of the half.

Although she initially put an equal amount of each color in the bag, when we started discussing her reasoning, she noticed that she needed four more red marbles to establish the “half” and “half of a half” relationships. The pie graph and references to the familiar “half” seemed to facilitate her multiplicative reasoning with this task.

Reasoning from a bar graph. The same questions were posed when I showed Amanda the bar graph on the right-hand side in Figure 7.17. She thought that I had ran the experiment “about 500” since “they look pretty high.” She did not mention the scale on the graph at all. She also noted that the bag of marbles “has a lot of yellows … two or three blues … and maybe seven green.” She then drew a bag of marbles with two blue, four green, and six yellow marbles. I then asked her to compare how many marbles she drew of each color with the bar graph.

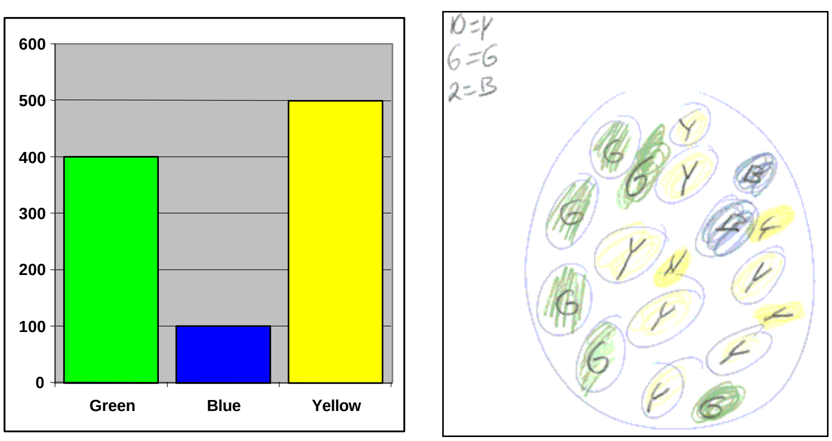


Figure 7.17. Given bar graph and Amanda’s drawing of a possible bag of “marbles.”

> T: So tell me why you did it the way you did.
> A: Because there are more yellows than greens, so I did green four, and I did the rest yellow but left two for the blue.
> T: Over here, how would you compare the green and the yellow from the graph?
> A: Pretty close.
> T: They’re pretty close? Are they pretty close in your bag of marbles?
> A: Yeah.

T: Alright and what about the blue compared to the yellow?
A: There's a lot more.
T: A lot more, okay and is that true in your bag?
A: Yeah.
T: And what about the blue to the green?
A: Close, no, not close.
T: Close, not close, what do you mean?
A: In the middle.
T: In the middle? What's in the middle?
A: There is more green compared to the [pause] blues have a little chance.
T: The blue has a little chance compared to the green? Okay, and is that true in your bag?
A: No.
T: No, why not?
A: Because it [green] only has two more than the blue. [She then adds two more green marbles and four more yellow marbles to her bag and records the number of each color. See Figure 7.17]

Once I focused her on comparing each of colors to the pie graph display, she realized that having the green and blue marble only two apart did not make sense if the green and yellow were only two apart. She then adjusted her bag of marbles to have green be more in the "middle." She justified her final bag of marbles by saying that "six is close to 10" and that the in the bar graph green and yellow "are pretty close" while blue "is not very much." Her reasoning was based more on a most-middle-least scale with a bit of additive reasoning when she adjusted the part-part relationships. Overall, she used a very informal quantitative approach to this task and did not try to establish any multiplicative relationships between the colors like she did with the pie graph. Granted, the relationships between 500, 400, and 100 are more difficult than the visual representations of half and "half of half."

<u>100 Gumballs</u>

The gumball task used in the post-interview was similar to the one used in the pre-interview but contained 30 yellow, 60 blue, and 10 red gumballs. I asked Amanda how many gumballs of each color she would predict if someone picked out 10 gumballs. She guessed six blue, three yellow and one red and based her reasoning on a multiplicative relationship between 60 blue and 30 yellow.

> A: Because I thought about 60 and 30 and I thought about 10, so 30 is half of 60 but since 60 is higher, 10 is half of 10, five is half of 10, so I just raised it up one more … So I told myself five is half of 10 but I raised it up one more and then I realized if I put one red, I would have three yellows and that would be half of 6.

It seems that she recognized the "half" relationship between yellow and blue in the gumball machine, but used that relationship inappropriately to think about the number of blues as half of 10 and then "one more." But then she realized by predicting one red she would have to predict three yellows and that did maintain the half relationship between 60 and 30 and six and three. Although her guess is in exact proportion to the gumball machine, her original guess of six blues was based on the relationship between 60 and 30 rather than 60 and 100. This shows that she can use multiplicative reasoning, but inappropriately applied it to this situation.

<u>Spinner Game</u>

The last task in the post-interview was similar to the spinner game used in the pre-interview; however, different spinners were used. With the first spinner (1/4 blue and ¾ red), Amanda picked the red area "because it has more of chance." I then chose the blue sector and asked her if this game was fair. She replied "yeah," and justified her reasoning based the excitement of the game.

> T: Why is it fair?
> A: Because if it was straight, this [half] would be blue, that [half] would be red then you'd be like "Oh it's blue" [yawning] "now it's red" [yawning] "ooh, this is boring."
> T: Oh, it would be very boring.
> A: Yeah, because we'd probably end up with the same amount [of pennies].
> T: So for a game to be fair, it has to be not boring – it has to be exciting?
> A: Yeah, like… well, it would not be fair like this but it would make it more exciting.
> T: So what do you mean it would not be fair like this?
> A: Well, there's less of the blue and then everybody would start arguing that if there are kids it would be a big fuss.
> T: So you think the way the spinner is now that it's not very fair or it is fair?
> A: In a way fair, in a way not fair. But you want it to be exciting so it needs to be like that.

At first she based her judgment of fair on the excitement factor of the game but later reasoned that the spinner was not fair the way it was. However, she still preferred to have the spinner in unequal parts since it would make the game exciting.

I showed her spinner B with eight equal sectors, two blue and six were red, and asked her if it mattered whether she used spinner A or B if she wanted the arrow to land on red. She said "it doesn't matter because I put those two [blue slices on spinner B] together … and then add these [red slices on spinner B] up to here and your luck with the reds is obviously the same size as that [red slice on spinner A] so that's the same size but they're in different places." She was able to reason about the equality of the size of the six slices based on visual appearance and a rearrangement of the sectors on spinner B. However, when I asked her if it would matter which spinner we used to play the game, she wanted to use spinner B because it would be more "exciting" and "the arrow will have a hard time finding the smaller pieces." She then used a metaphor of a family of ducks to explain her reasoning.

> A: Pretend this is a mother duck [the arrow] and here's her line of ducks [the eight sectors in spinner B] but she has two missing [the blue sectors], so she wants to find them but it's a big world out there so she keeps looking in different directions but the ducks are apart. It depends on what size they are, so if the ducks are real small it's real hard to find them.

Obviously, Amanda believes that red and blue areas on each of the spinners are equal, but thinks it would be better to play with spinner B since blue sectors are smaller and more spread out.

For the last part of this task, we revisited the question of whether the game, with spinner B, was fair. Amanda was able to base her description of fair on the equal size sectors in the spinner.

> A: Yeah, it's fair for the excitement. It's not fair for the amount.
> T: What do you mean by the amount?
> A: The amount of reds and the amount of blues.
> T: Oh, is there a way to make it fair as far as the amount?
> A: Draw on the computer.
> T: What would I have to do?
> A: Make a new one.
> T: Yeah, but what would that new one have to look like?
> A: Can I draw a picture? It could look like this [left-hand circle in Figure 7.18]. It could look like that [middle circle in Figure 7.18]. It could look like this [right-hand circle in Figure 7.18], anything like that.

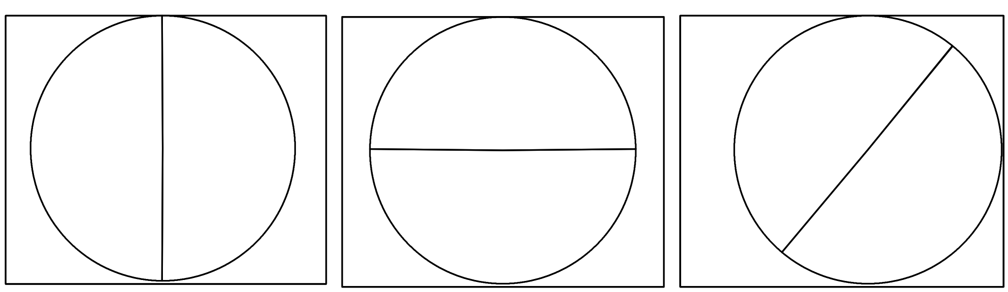


Figure 7.18. Amanda's drawing of three possible "fair" spinners.

> T: Could it look like… that? [I draw a picture with about a 20% slice]
> A: No, it could look like this too. [She draws a circle with quarter sectors and colors the sectors yellow, blue, red, and orange.]
> T: So, it would be four different colors?
> A: Yeah, blue-yellow-red- and orange.

All three of her examples of fair spinners were cut in half, but had different orientations. The other example she gave had quarter sectors. This is consistent with her references to “even” in the pie graph during the teaching sessions. It seems that half and quarter are two easily recognizable ways for her to make equal size slices. It is interesting, however, that she drew three examples in different orientations. I am not sure if she thought those spinners were different or was merely trying to illustrate that half of a circle is still is fair regardless of orientation to the person spinning the arrow on the spinner.

I also asked Amanda if she could think of way to make spinner B fair. She immediately said that she could color two more of the sectors blue so she would have “4 blues and four reds” and it would be fair since “they’re evened out.” Amanda’s sense of fair is certainly more grounded on a notion of equal parts, although she does still factor in the “excitement” of the game as also being a way to consider the fairness of a game.

<u>Strengths and Weaknesses in Post-Interview</u>

Amanda’s conception of fair, both in a coin toss and the spinner, was based on a notion of equal parts and equal chances. However, she still believed that “excitement” had to also be considered when establishing fairness in a game-playing context. Her initial language for describing theoretical probability was based on a part-part comparison, but she self-corrected her reasoning and used some part-whole language later in the interview. Although, with the help of her “buddy” system, she was able to eventually list all eight possible arrangements in the family task, she did not reason about the theoretical probabilities of each arrangement based on eight as the total number of choices. Instead, she used a part-part comparison within each arrangement of how many boys and how many girls there were.

Her understanding of independence of events was not stable. In the bucket situation, she recognized that as long as you replaced the green cube, it would still have the same chance of being chosen. Yet in a coin toss situation, she thought that flipping several heads in a row would increase the chance that a heads would occur because “it was on a roll.” She also thought with a coin toss “the higher the number, the lower the chance you’re going to get all one.” At a fundamental level, she seems to have a grasp on the effect of the number of trials on the likelihood of getting all heads. However, when faced with experimental data (e.g., four heads in a row), she relied on a subjective intuition based on the results rather than either independence or the number of trials.

With the coin tosses, she believed that “mixed” results were more likely to occur. In the task about determining if a coin was fair based on experimental data, she did express an intuition that results that were “closer” (and thus more mixed) would be more likely to occur and that would probably be an indicator of the fairness of the coin. However, recall that she was not able to explain her reasoning for this idea.

Amanda used 50% as a benchmark in her analyses in both the marble bag and spinner tasks. She could recognize the equivalence of the chance for picking a black marble for the 3B3C and 1B1C bags and the unequivalence, with reference to half, of the blue and red sectors on spinner A. She could also equate the areas on spinner A and spinner B as equally likely based on a visual rearranging on the sectors in spinner B to match spinner A. However, Amanda was not able to recognize the equivalence in the 2B1C and 4B2C bags of marbles. In her analysis of this pair of bags, as well as two other pairs, she used part-part reasoning and a comparison of closeness to 50-50 in determining which bag had the better chance for choosing a black marble.

Amanda used multiplicative reasoning in designing a bag of marbles based on the pie graph but did not employ such reasoning with the bar graph. Her reasoning with the bar graph task was based on a most-middle-least scale. She did employ multiplicative reasoning in the gumball task, albeit for an inappropriate reason.

Overall, Amanda showed evidence of developing some of her primary intuitions about chance and randomness. She was beginning to develop her probabilistic reasoning and could compare some probabilities based on a reference to 50%, while others she incorrectly compared using a part-part difference from 50%. Her reliance on part-part reasoning and additive relationships shows she has not developed a stable use of part-whole reasoning or establishing and maintaining multiplicative structures. Her multiplicative reasoning in the post-interview was always in reference to “half.”

Putting It All Together: Amanda’s Development of Probabilistic Reasoning

Amanda enjoyed the playful nature of the computer microworld environment and was fascinated by many of the visual stimulants that occurred during a simulation (e.g., motion of pie graph). However, directing her play into mathematical meaning-making activity usually required prompting from an external source, either tasks or focused questions posed by other students and teachers. Once engaged in an activity, she was able to use the tools to help her reason through an investigation and she occassionally made some connections between the multiple representations in the microworld (e.g., weight tool and bag of marbles, pie graph and numerical data).

At the beginning of the teaching experiment, Amanda had several primary intuitions about chance occurrences and quantified her belief in the chance of an event occurring with verbal descriptions such as “easy,” “sometimes and sometimes not,” and “hard,” corresponding to a most-middle-least scale. She did not consistently use any reasoning based on independence or the effect of number of trials. In addition, during her pre-interview, she did not explicitly connect the concept of “fair” in a game situation with equal chances. This was probably due to her game playing experiences and her interpretation of a fair game as a worthwhile game.

Amanda made some progress in her understanding of theoretical probability. She could sometimes use part-whole statements to describe the chance of an event occurring. Although the weight tool was initially too abstract for her, the connection between the weight tool and bag of marbles helped her interpret the weights in terms of a concrete model of a collection of objects. The use of the weight tool and graphical displays seemed to only occasionally help her interpret theoretical probability and make judgments of the likelihood of certain events. It is also important that Amanda used experimental data to help her make judgments about the likelihood of something occurring. Even if a result seemed to deviate from her expectation, she based the “usualness” of that result on whether or not it actually occurred. Thus, a result that may be theoretically “unusual,” may be not-so-unusual by Amanda’s standards. She seemed to usually reason well with *a postiori* situations, both objectively and subjectively, but did not significantly develop her *a priori* reasoning.

Although her initial prediction of expected results did not always reflect the theoretical probability (e.g., predicting 30-70 from a 5B5W bag, predicting 5-5 from a 3B1W bag), she used the microworld to test her prediction and several of the tools (e.g., graphs, data table, stacking columns) to analyze the data. When the actual results did not match her expectation, she was only sometimes able to adjust her hypothesis in light of the evidence. Amanda was only sometimes able to link the experimental results to the theoretical probability as a factor influencing the resulting data. Other times, she was able to accept the data, but was not able to explain why the data did not match her initial expectation. During the post-interview, Amanda used this notion in her analysis about what experimental data could tell her about the fairness of a coin. She had a sense that she expected about an equal distribution of heads and tails. With the results of 175-325, she thought the results seemed too far apart but was not able to directly connect her reasoning with the fairness of the coin. However, I believe this was due to her definition of a fair coin having one head and one tail and that she probably could not conceptualize how a real coin could be unfair that had two different sides.

Amanda almost always used part-part reasoning and only employed part-whole reasoning in tasks while she was using multiplicative reasoning. Those tasks always involved relationships of “twice” and “half.” The exception was when Amanda made part-whole connections with the pie graph when she figured out why the blue slice representing one out of three, and then one out of six, was the appropriate size by creating the same number of equal-sized sectors as the number of trials. Although Amanda made progress in her interpretation and use of the pie graph, she rarely used the pie graph to help her reason probabilistically about expected results or theoretical probability. Her use of multiplicative reasoning also only minimally helped her develop probabilistic reasoning. She did not use multiplicative reasoning during the post-interview in her analysis of probability comparisons. She did, however, appropriately use multiplicative reasoning when she designed a bag of marbles based on the pie graph results. But again, this was in a familiar "half" situation.

Amanda's conception of independence barely improved during the teaching experiment. She was more consistent in the post-interview about believing that "mixed up" results were more likely to occur rather than her inconsistent reasoning based on both "mixed" and "matched" in the pre-interview. I believe her experiences with the microworld and running a large number of trials may have facilitated this more consistent reasoning. However, she still used reasoning based on the positive recency effect and discounted the number of trials and her "mixed is better" notion when she expressed that heads could be more likely "if they are on a roll." I conjecture that in this point in her development of probabilistic reasoning, her reasoning based on *a postiori* evidence supersedes any other developing ideas she has about the effect of the number of trials, "mixed" results being more likely, and independence of events.

Overall, Amanda made only some progress in her development of probabilistic reasoning throughout the teaching experiment. I believe the microworld tools helped her to make connections between numerical and graphical representations and that the combination of the weight tool and bag of marbles facilitated her slightly improved understanding of theoretical probability and ability to reason multiplicatively with simple tasks. Her primary intuitions about chance did develop but I believe that lack of self-confidence and frustration when she could not verbalize her thoughts hindered her engagement and risk-taking during the teaching sessions. She was much more willing to take risks when working in a researcher-child pair when the researcher would constantly praise and encourage her thinking. However, she still did not take advantage of the tools in the microworld to investigate experiments of interest to her. Almost all of the meaningful probabilistic reasoning was initiated and sustained through researcher-child interactions.

CHAPTER 8

A CROSS-CASE COMPARISON OF CHILDREN'S MICROWORLD INTERACTIONS AND DEVELOPMENT OF PROBABILISTIC REASONING

The previous three chapters contained detailed descriptions of each child's development of probabilistic reasoning throughout the entire research study.
The purposes of this chapter are to focus on the children's interactions with and use of the microworld tools, and their subsequent development of probabilistic reasoning. As evidenced in the descriptions of each child's meaning-making activity, the children's uses of the microworld tools and their understanding of probability concepts varied greatly. This is actually a strength of the study, as it illustrates how children use the tools for different purposes, and that a tool that acts as an agent for one child can serve as a deterrent for another, depending on individual intuitions and current schemes. Before detailing the particular interactions that affected each child's development, I first discuss the various ways other researchers have described interactions with a microworld for cognitive growth. These descriptions provide a framework for my characterization of each child's use of the *Probability Explorer* tools.

Characterizing Children's Interactions with Microworld Tools

Open-ended learning environments, such as microworlds, are purposefully designed to maximize the potential for meaningful interaction. The following are several of the inherent characteristics of a well-designed open-ended learning environment (OELE):

- OELEs enable learners to build and test their intuitive, and often misconceived, notions.
- OELEs support experiences wherein learners begin to explore, build upon, and make explicit their intuitive notions.
- OELEs assume that understanding is a continuous and dynamic process that evolves as a result of observation, reflection, and experimentation.
- OELEs support experiences for learners to identify, question, and test the limits of their intuitive beliefs. (Land & Hannafin, 1996, p. 38)

Children interacting in such open-ended microworld environments can learn through developing "theories-in-action" (Karmiloff-Smith & Inhelder, 1975, cited in Land & Hannafin, 1996) in which they generate intuitive-based theories and modify them as they reflect upon experiences that either confirm their intuitions or challenge the validity of their theory through perturbations. Children's intuitive notions may drive many of their initial actions in a microworld. Child-computer interactions and goal-oriented tasks are central in the development of intuitive notions into theories-in-action.

The children in my study exhibited many intuitive notions about probability. Each of them used the microworld tools in a variety of ways and developed different levels of understanding. My analysis of students' use of tools in *Probability Explorer* is based on the descriptions of types of interactions and levels of mathematical activity identified in the research of Tzur (1995) and Steffe and Wiegel (1994).

Tzur characterized interactions by focusing on the various ways children used the computer. In his research on children's use of a microworld to develop fraction knowledge, he identified four major types of child-computer interactions.

1. Children use the microworld as a *medium for actions* to solve a task.
2. Children use the microworld tools to *explain and justify* a solution to others.
3. Children's actions in the microworld *induce a perturbation*.
4. Children use features and tools in the microworld to *gain control* over a situation.

Each of the children in the current research study engaged in all four types of child-computer interactions. However, for each child, the four types of interactions occurred with different frequencies and to varied degrees.

Steffe's and Wiegel's (1994) characterization of children's activities with a microworld include the playful orientation of the children's actions, the mathematical purpose of the actions, and whether the actions were initiated and sustained by children or by a teacher. The following four characterizations of children's activities are slightly modified from Steffe's and Wiegel's model.

1. Cognitive play – playful actions in a microworld initiated by a child and not intentionally mathematically oriented.
2. Mathematical activity – goal-oriented actions in a microworld initiated by a teacher's instructional intervention and usually sustained by teacher guidance. (Cognitive play can be transformed into mathematical activity with an intervening question about an action or result during cognitive play.)

3. Independent mathematical activity – goal-oriented actions in a microworld initiated by a teacher but sustained by a child as he or she pursues the original goal with minimal teacher guidance, and possibly initiates additional mathematical investigations.
4. Mathematical play – independent mathematical activity that has a distinctive playful orientation that further engages the child in investigating mathematics.

These four ways in which children can be engaged in activities with a microworld are highly dependent on the learning environment, and on a child's current understandings, tendency to take risks in an exploratory manner, and successful connections between actions in the microworld. Carmella and Jasmine were, to various degrees, engaged in all four activities, while Amanda was engaged in the first two types of activities.

I will use Tzur's characterizations to highlight the direct child-computer interactions that occurred while a child was engaged in playful and mathematical activities as characterized by Steffe and Weigel. In this regard, I am combining both of their frameworks in hopes to provide a more in-depth characterization of children's computer interactions.

In the next section, I elaborate on each child's microworld interactions with respect to the frameworks, and how those interactions affected her development of probabilistic reasoning. After the discussion of each child, this section concludes with a cross-case comparison of the children's types of interactions (Tzur) and levels of mathematical activity (Steffe & Weigel).

Effect of Child-Microworld Interactions on Probabilistic Reasoning

The children's interactions with the *Probability Explorer* tools were a primary way for them to enact their intuitions about chance and develop further intuitions and theories-in-action. Since each child interacted differently with the microworld tools, I briefly discuss their individual interactions and how those interactions affected their probabilistic reasoning. This section concludes with a brief illustration of how the three children compared in the types and levels of microworld interactions.

Carmella's Advanced Interactivity and Development of Theories-in-Action

Carmella's advanced use of the microworld tools was facilitated by the schemes she had at the beginning of the teaching experiment, as well as her confidence, willingness to take risks in an investigative manner, and continual verbalizations as she talked aloud to make sense of her microworld experiences. She entered the teaching experiment with a strong understanding of several concepts. She associated "fair" directly with equal chances of winning, and she could explain theoretical probability using her hypothetical experiment strategy (HES). She used simple fractions and could use both part-part and part-whole reasoning, although in her part-whole reasoning, she used both inappropriate additive and appropriate multiplicative reasoning. Carmella also had an unusually strong understanding of independence and had intuitions that the number of trials affected the chance of something happening.

I suggest that Carmella's previous and developing probabilistic understandings facilitated her advanced level of interactivity with the microworld. She interacted with the microworld in all four ways as described by Tzur (1995). She playfully explored the microworld tools (cognitive play), and took advantage of the flexibility and multiple representations in the environment to solve a variety of tasks of posed and facilitated by a teacher/researcher (mathematical activity), as well as those initiated and sustained by herself or another child (independent mathematical activity and mathematical play).

The dynamic link between numerical and graphical results, especially visualization of the pie graph, during the simulation process facilitated Carmella's theory-in-action about the evening out phenomenon (EOP). After her initial experience with the EOP in the first teaching session, Carmella used the tools in the microworld to help develop and refine her understanding of the EOP. She systematically investigated the effect of increasing the number of trials on the results from a die toss by running several sets of 10, 20, 40, and 80 trials and using the stacking columns and data table to comment about the range between the lowest and highest result (medium for action). She used the tools to carefully control her investigation to test out her theory-in-action about the effect of increasing the number of trials (gain control). After several sets of 80 trials, she began running sets of 200 trials and opened the pie graph to observe the results visually "wiggle" and then "hardly move at all." Carmella was able to further refine her theory-in-action when I asked her to think about and then simulate an additional 200 trials added onto the current 200 trials. She initially expected the "wiggling" process to begin again and needed the simulation experience and visualization with the pie graph to induce a perturbation. The iconic stacked representation of the data and the numerical displays in the data table enabled her to pay attention to the absolute variability in the range.

However, the pie graph facilitated her recognition that the important variability in the range was relative to the total number of trials (e.g., “the rate stays the same”).

Carmella also used several tools throughout to enact and refine her theory-in-action about the effect of the total number of trials on the probability of getting the exact theoretical distribution and a relatively close distribution. The weight tool allowed her to transform her HES into her total weight approach (TWA) to experimentation. The connections she made between the weight tool, simulation process, pie and bar graphs and data table facilitated her experimentation, perturbations, and further refinement of her theory-in-action. Her use of these tools to investigate this theory-in-action was almost always during independent mathematical activity and mathematical play. Carmella used these tools to her advantage and engaged in a high level of interaction with the microworld to explore questions of interest to her.

Carmella occasionally used the tools to justify her prediction or solution to a task. For example, during the seventh teaching session, she created an elaborate “bubble map” on paper to predict the outcomes for 100 trials for both a 3B1W and 6B2W bag of marbles. After making a 75-25 prediction with multiplicative reasoning, she ran several simulations and used the pie graph and data table to confirm that the results were close to her prediction. Her work on that task demonstrates her independent mathematical activity in developing a strong understanding of the relationship between empirical and theoretical probability as the number of trials increases.

Carmella entered the teaching experiment with an intuition about the effect of the number of trials on the likelihood of an event occurring. Through interactions with the microworld tools, she developed an intuitive-based theory-in-action about the effect of the number of trials and the law of large numbers. Her advanced level of interactivity was influenced by her strong intuitions and risk-taking personality. Overall, Carmella's use of the microworld demonstrates that the tools were appropriately designed for her to use, effective in enacting connections between the representations, and flexible enough for her to formulate and investigate her theories-in-action.

<u>Jasmine's Use of Multiple Representations and Sustained Mathematical Play</u>

Jasmine was highly engaged in tasks with the microworld, enthusiastically shared her thinking with the group, and used playful metaphors to describe her actions and data analysis. At the beginning of the teaching experiment, Jasmine's intuitive ideas about chance included a strong preference for "mixed up" results as representative of more likely random events, and a recognition of likelihood that she could describe and compare with her non-normative, but consistent, use of percents. A circle representation seemed to be a familiar context for her as it helped her correct her non-normative use of percents. In addition, Jasmine only used additive and part-part reasoning in her pre-interview.

Jasmine's parents had a concern about her visual perception learning disability and her possible difficulty in interpreting the displays on the computer, especially displays that would be changing rapidly. Throughout the entire teaching experiment, there is no evidence to suggest that Jasmine had difficulty interpreting the simulations and resulting data in this highly visual microworld environment. In fact, the visual nature of the environment seemed to facilitate her probabilistic reasoning. She made explicit connections between the multiple representations and noted in the first teaching session "we have four ways to see everything," referring to the stacked iconic data, pie graph, bar graph, and data table. In addition, she was able to explain the changes in the pie graph during the simulation process by referring to the process of random data being generated and the distribution of results continually changing.

Although Jasmine had some intuitions about chance prior to using the microworld, her understanding of probability was not as well developed as Carmella's. In addition, she was shy and tentative during the pre-interview and beginning of the first teaching session, until she started to work with the microworld. Instead of being intimidated, she appeared to be empowered by the ability to manipulate the objects on the screen and analyze the results from an experiment in a variety of ways. She literally "came alive and took control of her learning," as the non-participant observer noted, when using the microworld to explore chance situations. Although Jasmine did not have as strong of an understanding for probabilistic ideas as Carmella did, and did not appear to be as verbal and as much as a risk taker, she quickly engaged herself with the tasks in *Probability Explorer* and used the tools to her advantage.

Jasmine also interacted with the microworld in the four ways described by Tzur (1995). She especially took advantage of the multiple representations in the environment to solve a variety of tasks posed and facilitated by a teacher/researcher (mathematical activity), as well those initiated and sustained by herself or another child (independent mathematical activity and mathematical play).

Jasmine continually used the microworld tools as a medium for action to solve a variety of tasks (e.g., figuring out the "secret weights" through experimentation and data analysis). She sometimes used the tools to justify a prediction or solution to a task. During the third teaching session when, with a 3B1W bag of marbles, she conjectured that she was more likely to get more white than black marbles if she did a small number of trials. She continually ran sets of three trials until she got two white and one black marble and noted that “if we were doing 50, we’d be here all day.” Her work on that task demonstrates one example of her independent mathematical activity that contributed to her developing an understanding of the effect of the number of trials.

The pie graph was a major cognitive prompt for Jasmine in a variety of contexts and induced perturbations that helped Jasmine develop her understanding of the EOP, part-whole relationships, and multiplicative reasoning. The visualization of the pie graph during simulations helped her develop an understanding of the EOP. She was able to transition from part-part to part-whole reasoning by thinking about the whole pie as a set number and then reasoning about the number associated with each slice relative to the whole. In this regard, the pie graph was also an aid in facilitating her multiplicative reasoning. The combination of the pie graph, weight tool and bag of marbles helped Jasmine make connections between empirical and theoretical probability.

Jasmine had a distinctive playful approach to using and interpreting the actions in the microworld (e.g., “gray guy eating the sky”). However, except on one occasion, her playful approach was still goal-oriented and focused on the mathematics (independent mathematical activity and mathematical play). Jasmine was able to use *Probability Explorer* to simulate experiments that were both playfully contrived (e.g., designing an experiment with a “girl, her house, and her dog” to see if “every girl has one house and one dog”), and reflective of her real world experiences (e.g., modeling the bingo barrel). Both cases are an example of her ability to initiate and sustain mathematical play.

Jasmine's probabilistic reasoning, especially her part-whole and multiplicative reasoning, with the microworld did not transfer to her reasoning with many of the tasks in the post-interview. The only post-interview instance of part-whole reasoning in a probabilistic situation was during her creation of a bag of marbles given a pie graph of experimental results. It seems that the multiple representations available to her in the microworld facilitated her reasoning, and that with only one representation available in the post-interview tasks (e.g., bag of marbles or spinner), she mostly did not employ appropriate part-whole or multiplicative reasoning in assessing the probabilistic situations. Perhaps Jasmine needed more experiences with the multiple representations in order to transfer her probabilistic reasoning to situations with only one representation, or perhaps the types of representations available in the microworld were not sufficient in helping her truly develop a mental scheme for probability concepts.

Amanda's Confirmations, Perturbations, and Playful Orientation

Amanda's interactions with the microworld had a playful orientation that often disengaged her from the context of a probability task. Many of the actions in the microworld served either as confirmations of her existing intuitions, or perturbations that were mainly outside of her "zone of potential construction" (Olive, 1994). Amanda had many primary intuitions about chance at the beginning of this study. She could only verbalize notions about theoretical probability and probability comparisons by using descriptors such as "easy" and "hard" based on a most-middle-least scale. She was not consistent with her ideas about whether strings of results that were "mixed" or "matched" were more likely. She used the positive recency effect in determining that a heads was more likely to occur after a sequence of several tails. She also had a tendency to believe that random results could be determined better if you attended to physical characteristics (e.g., marble at top of the bag) or how you used a physical object (e.g., heads are more likely if you start the flip with heads side up). Amanda had a very difficult time reasoning with most tasks and verbalizing any possible explanations.

Amanda's initial interactions with the microworld demonstrated that she did not comprehend what the various displays represented and what was happening during the simulation process. She seemed to be fascinated with the visualizations (e.g., changing pie graph, scattered icons on the screen) during experimentation and tended to run a large number of trials when given a choice. Her choice of a large number of trials does not appear to be a purposeful attempt on her part to explore the effect of a large number of trials, but to prolong the simulation process so she could "watch" the actions on the screen. I conjecture that since she did not understand the processes on the screen, this was a way to disengage herself. She rarely engaged in thoughtful reflection, even when asked directed questions, about the relationships between multiple representations as well as between empirical results and the theoretical probability. The visualization during the simulation process was more of a deterrent for her probabilistic reasoning rather than an agent in promoting appropriate experiences. Amanda never enacted a theory-in-action about the EOP like the other children.

Amanda's interactions with the microworld could technically fall into each of the four types described by Tzur (1995). However, her use of the tools to solve problems and to justify a prediction or solution was mostly limited to instances when the interaction was guided by teacher questions, and sometimes even explicitly suggested. There were only a few occasions when she used the tools to take control of a situation (e.g., changing the weights to 89 and 90 make a coin more even). Amanda's goals in most of her interactions did not seem to have a mathematical focus unless explicitly addressed by another child or teacher/researcher. Because of this, many of the results during simulations that seemed "unusual" to her may have caused a perturbation, but these perturbations rarely prompted her to accommodate her understanding of probability.

Land (1995) addressed this perturbation phenomenon in her dissertation research with students using a science microworld. She observed that sometimes students adapted contradictory data from their microworld experiences to fit into their existing theory-in-action rather than use the data as a perturbation that could result in a refinement of the theory. An example of this occurred when Amanda intuitively thought that weights of 2000 and zero for a coin toss experiment would result in "some" tails. After running about 100 trials, there was certainly no evidence that a tail had occurred. She subsequently used this contradictory evidence to fit into her theory-in-action that tails had a very small chance of occurring. She refined that theory by hypothesizing that the number of trials had to surpass 2000 before a tail would occur. She ran these trials and fully expected a tail to occur shortly after the 2000th trial. When this did not occur, she seemed to accept this as evidence that a tail would not occur. However, with weights of one and zero, she reverted back to her previous theory-in-action and thought that she would not have to run as many trials before a tail would occur since there was less weight for the head. Although she was eventually able to refine her intuitive-based theory to account for all the contradictory evidence, she needed many instances of the contradiction before a perturbation was actually within her zone of potential construction.

Amanda’s use of the microworld tools can only be characterized by the levels of cognitive play and mathematical activity. Because of her lack of understanding and frustration level, most of her interactions in these two categories were prompted by and sustained by other children and the teachers/researchers. An instance where she initiated cognitive play was her experiment in the first teaching session to see if she could “fill up” the pie graph with all blue or all gray by running a large number of trials. Although she initiated this cognitive play, her goal not mathematically appropriate.

Although her understanding of the structure of the pie graph improved during the teaching experiment, she only occasionally made connections that indicated she understood the inter-relatedness of the various tools (e.g., marble bag, weight tool, pie graph). There was also little evidence that she truly understood the effect of the weight tool and contents of the marble bag on the likelihood of an event and the empirical data. In effect, I think her lack of reflective meaning-making activity in the microworld is linked to the current tools available in the microworld, the types of tasks used in the teaching experiment, her intuitive-based misconceptions, and her level of frustration with reasoning and verbalization.

Comparison of Interactions and Levels of Mathematical Activity

The types of interactions and levels of mathematical activity of each of the three girls were different and affected their probabilistic reasoning in various ways. Carmella obviously had the highest level of mathematical activity and wide spread use of different types of interactions with the microworld. Jasmine’s playful and empowering orientation toward using the microworld tended to also engage her in higher levels of mathematical activity and various types of interactions. However, Amanda's playful orientation mostly distracted her from engaging in mathematical activity and limited her self-initiated interactions with the computer. Figure 8.1 illustrates a summary of the frequency of the types of child-computer interactions (Tzur, 1995) that each child was engaged in at the four levels of increased mathematical activity (Steffe & Weigel, 1994).

Levels of Mathematical Activity (Steffe & Weigel,

Levels of Mathematical Activity	Medium for Action	Justify a Solution	Induce a Perturbation[1]	Gain Control
Mathematical Play	C (o) J (o)	C (s) J (s)	C (s)	C (o) J (s)
Independent Mathematical Activity	C (o) J (o)	C (o) J (s)	C (o) J (s)	C (s) J (s)
Mathematical Activity	C (o) J (o) A (o)	C (o) J (s) A (s)	C (o) J (o) A (o)	C (o) J (o) A (o)
Cognitive Play	C (s) J (o) A (o)	Not Applicable	C (s) J (o) A (s)	J (s) A (o)

Types of Child-Computer Interations (Tzur, 1995)

Key:
C = Carmella J = Jasmine A = Amanda
(o) = occurred often **(s) = occurred** sometimes

[1] For Carmella, perturbations almost always led to increased mathematical activity and subsequent improvement in understanding. For Jasmine, the perturbations sometimes led to increased mathematical activity and subsequent improvement in understanding. However, for Amanda, the perturbations caused frustration and rarely led to improved understanding

Figure 8.1. Chart of children's interactions and levels of mathematical activity.

The chart in Figure 8.1 illustrates that Carmella and Jasmine were engaged in all four types of child-computer interactions and all four levels of mathematical activity. However, Amanda was only engaged in the two lowest levels of mathematical activity and all four levels of child-computer interactions, although she only sometimes interacted with the computer to justify a solution. It is also important to note that Carmella was only sometimes engaged in cognitive play and that she was often engaged in the three highest levels of mathematical activity. Jasmine was often engaged in cognitive play and mathematical activity. Although she was also often engaged in independent mathematical activity and mathematical play, she was only sometimes involved in interactions at those levels where she was justifying a solution, gaining control, or where her interaction induced a perturbation. It is also important to note that although each child was engaged in interactions with the microworld that induced a perturbation, the perturbations were most effective in altering Carmella's understanding and rarely affected Amanda's probability understanding.

The differences and frequency of the types of interactions, levels of mathematical activity, and effect of perturbations are important considerations when analyzing the development of probabilistic reasoning of the three children. The next section includes a summary and cross-case comparison of the children's development of probabilistic reasoning in each stage of the teaching experiment.

Children's Development of Probabilistic Reasoning

The purpose of this section is to provide a summary and cross-case comparison of the children's meaning-making activity during the teaching sessions and their overall development of probabilistic reasoning. Although the summaries are focused on the children's probabilistic reasoning, it is important to remember their interactions and levels of activity with the microworld when considering their reasoning during the teaching sessions.

Cross-Case Comparison of Meaning-Making Activity With Microworld

Recall that within each case study, the children's meaning-making activity was discussed through analyzing their development throughout the teaching sessions in several evidentiary themes. Each theme characterized the main focus of their mathematical development while problem solving with the computer microworld. Figure 8.2 summarizes the evidentiary themes for each child.

Carmella	Jasmine	Amanda
Total Weight Approach	**Theoretical Probability**	**Theoretical Probability**
"Evening Out" Phenomenon *	**"Evening Out" Phenomenon ***	**Expected Results**
Close vs. Exact	**Part-Whole Reasoning ***	**Interpreting Pie Graph**
Proportional Reasoning *	**Additive and Multiplicative Reasoning ***	**Multiplicative Reasoning ***

***** Indicates that the pie graph was a significant tool used in meaning-making activity in a particular theme.

Figure 8.2. Cross-case evidentiary themes of meaning-making activity during teaching sessions.

Although the three children developed different mathematical ideas in their personal meaning-making activity, there were some similarities. As shown in Figure 8.2 with the colored boxes, Jasmine and Amanda were both engaged in developing a better understanding of theoretical probability, albeit at different levels. Carmella and Jasmine spent a significant amount of time investigating and making sense of the "evening out" phenomenon. Lastly, all three children further developed their multiplicative reasoning in the context of probability tasks. However, Carmella's development of multiplicative reasoning facilitated her proportional reasoning in a variety of contexts (e.g., comparing bags of marbles, making proportional numerical predictions, comparing circle areas), while Jasmine and Amanda had more struggles between appropriate use of additive and multiplicative reasoning. In addition, Jasmine's and Amanda's multiplicative reasoning was almost always done in contexts with "double" or "half" relationships.

The chart in Figure 8.2 also indicates that the pie graph was a significant tool used in meaning-making activities for each of the children in many different themes. The pie graph served as a major cognitive prompt for both Carmella and Jasmine in themes involving dynamic contexts (e.g., EOP) as well as static contexts (e.g., proportional and multiplicative reasoning). Amanda's development under the "pie graph" theme was mainly trying to understand that representation. For the most part, the pie graph was a deterrent in her meaning-making and probabilistic reasoning.

Noting the similarities and differences in the children's meaning-making activities during the teaching sessions helps to frame the discussion of the children's zone of development of probabilistic reasoning. The next section contains a summary and cross-case comparison of the children's development of probabilistic reasoning.

Children's Zone of Development of Probabilistic Reasoning

To consider the children's development of probabilistic reasoning, before, during, and after their use of the microworld tools, it is helpful to build upon prior research on children's understandings of probability. Jones *et al* (1999b) developed a framework (see Figure 8.3) that characterized four levels of understanding of six probability concepts.

Construct	Level 1 Subjective	Level 2 Transitional	Level 3 Informal Quantitative	Level 4 Numerical
Sample Space	* Lists an incomplete set of outcomes for a one-stage experiment	* Lists a complete set of outcomes for a one-stage experiment and sometimes for a two-stage experiment	* Consistently lists the outcomes of a two-stage experiment using a partially generative strategy	* Adopts and applies a generative strategy to provide a complete list for two- and three-stage cases
Experimental Probability of an Event	* Regards data from random experiments as irrelevant and uses subjective judgments to determine the most or least likely event * Indicates little or no awareness of any relationship between experimental and theoretical probabilities	* Puts too much faith in small samples of experimental data when determining the most or least likely event; believes any sample should be representative of the parent population * May revert to subjective judgments when experimental data conflict with preconceived notions	* Begins to recognize that more extensive sampling is needed for determining the event that is most or least likely * Recognizes when a sample of trials produces an experimental probability markedly different from the theoretical probability	* Collects appropriate data to determine a numerical value for experimental probability * Recognizes that the experimental probability determined from a large sample of trials approximates the theoretical probability * Can identify situations in which the probability of an event can be determined only experimentally
Theoretical Probability of an Event	* Predicts most or least likely event on the basis of subjective judgments * Recognizes certain and impossible situations	* Predicts most or least likely event on the basis of quantitative judgments but may revert to subjective judgments	* Predicts most or least likely event on the basis of quantitative judgments * Uses numbers informally to compare probabilities	* Predicts most or least likely event for one- and simple two-stage experiments * Assigns a numerical probability to an event (a real probability or a form of odds)
Probability Comparisons	* Uses subjective judgments to compare the probabilities of an event in two different sample spaces * Cannot distinguish “fair” probability situations from “unfair” ones	* Makes probability judgments on the basis of quantitative judgments – not always correctly * Begins to distinguish “fair” probability situations from “unfair” ones	* Uses valid quantitative reasoning to explain comparisons and invents own way of expressing the probabilities * Uses quantitative reasoning to distinguish “fair” and “unfair” probability situations	* Assigns a numerical probability and makes a valid comparison
Conditional Probability	* Following one trial of a one-stage experiment does not always give a complete listing of possible outcomes for the second trial * Uses subjective reasoning in interpreting with and without replacement situations	* Recognizes that the probability of some events changes in a without replacement situation; however, recognition is incomplete and is usually restricted only to events that have previously occurred.	* Recognizes that the probability of all events changes in a without replacement situation * Can quantify changing probabilities in a without replacement situation	* Assigns numerical probabilities in with replacement and without replacement situations * Uses numerical reasoning to compare the probability of events before and after each trial in with replacement and without replacement situations
Independence	* Has a predisposition to consider that consecutive events are always related * Has a pervasive belief that one can control the outcome of an experiment	* Begins to recognize that consecutive events may be related or unrelated * Uses the distribution of outcomes from previous trials to predict the next outcome (representativeness)	* Can differentiate independent and dependent events in with and without replacement situations * May revert to strategies based on representativeness	* Uses numerical probabilities to distinguish independent and dependent events

Figure 8.3. Framework for describing probabilistic reasoning (Jones *et al.*, 1999b, p. 150)

Recall that the Jones *et al* research was with elementary-age children, and although they used an experimental approach to teaching probability, their teaching approach did not utilize technology tools. Before using this framework to summarize the children's development of probabilistic reasoning in this technology-rich study, there are several suggested amendments to this framework that I propose will make the framework more meaningful and applicable to the probabilistic reasoning of the three children in my study.

Amending the framework. In regards to approaches to probability, I agree with Hawkins and Kapadia (1984) that a subjective approach (i.e., learning from experience, availability of information, etc.) can be appropriate and useful. Having the lowest level of probabilistic understanding labeled as "subjective" insinuates that such an approach is not as viable as other objective approaches to probability tasks. Based on the descriptions in many of Jones *et al* articles (1997, 1999a, 1999b), I believe they use the term subjective to primarily refer to those judgments a child makes that reflect personal preferences and a person's actions that could cause an event to occur. Therefore, I suggest replacing the term subjective with "egocentric" throughout the framework. I do not believe replacing their use of the term subjective with egocentric changes their original intent, it merely removes the insinuation that a subjective approach to probability is at the lowest level of understanding.

Jones *et al*, argue that their framework is relatively consistent between constructs (i.e., a child at level three in sample space should be at level three in the other constructs as well). Although the mapping of development within each construct is meaningful, I do not believe that children's progress in developing understanding within each construct develop at the same rate. For example, I do not view the concept of independence as equally accessible to children as the ability to systematically generate a sample space for a 3-stage experiment. In addition, understanding independence of events, especially while developing understanding about the relationship between empirical and theoretical probability, is very difficult for children. Being able to appreciate that the distribution of results from random events tends toward a very regular pattern and understanding that future results are not dependent on past results is cognitively demanding for a child trying to understand random processes. Therefore, while Jones *et al* strive to be able to label a child as “level 3” across all constructs, I do not believe that type of “consistent” labeling is needed, nor constructive.

There is another change I suggest for the framework. Based on my work with three children in this study, two of those children (Carmella and Jasmine) reached a level of understanding while using the microworld that surpassed the numerical level. By this I mean, they not only could use reasoning in their problem-solving consistent with level 4, but they were using their level four understandings of several constructs in relation to each other. I call this fifth level “relational.” At the relational level, children can use their numerical understandings of a construct as it relates to another construct in a meta-numerical manner. The child understands the relationship between two or more constructs at an *analytical* level. For example, a child can list all possible outcomes for a three-stage experiment, use that sample space to determine the theoretical probability of an event, both ordered and unordered, and then compare the likelihood of several events. Another example of thinking at the relational level is when a child can maintain an understanding that every random occurrence is independent while developing an understanding for the law of large numbers and the effect of the number of trials on the theoretical probability of an event occurring.

Zone of development. I used the amended framework to create a mapping of each child's understanding of probability concepts before, during, and after their use of the microworld (see Figures 8.4, 8.5, and 8.6). These mappings are not a reflection of *measures* of their understanding, rather, I stretch the mapping across a range of understandings to reflect an *approximate zone of development*. The mapping is a purposeful attempt to highlight the children's varied and continuously developing understanding on a variety of tasks and within a variety of contexts. On some tasks, within some contexts, they may have demonstrated a higher level of understanding, while other tasks may have elicited reasoning at lower levels.

The mappings reflected in Figures 8.4, 8.5, and 8.6 were carefully derived from the evidence presented in each case study and the children's interactions with the microworld and their levels of mathematical activity. For each of the six concepts, I analyzed their reasoning on tasks that used those concepts and made a judgment, based on the evidence in their reasoning, use of the microworld tools, and responses to questions, to place the majority of their reasoning within a certain range. The lines extending from the boxes (in both directions) in Figures 8.4, 8.5, and 8.6 represent the contextual and tool-based extensions in their reasoning. There were many instances when the children used a higher or lower level of reasoning on a certain task, using a certain microworld tool, or in certain contexts. Although the solid boxes in Figures 8.4, 8.5, and 8.6 represent the approximate zone for the majority of their reasoning, it is necessary to consider the stretched zone of development that occurred. Thus, I am merely using these mappings as a visual summary of the children's zone of development of probabilistic reasoning that was detailed in each of the case studies.

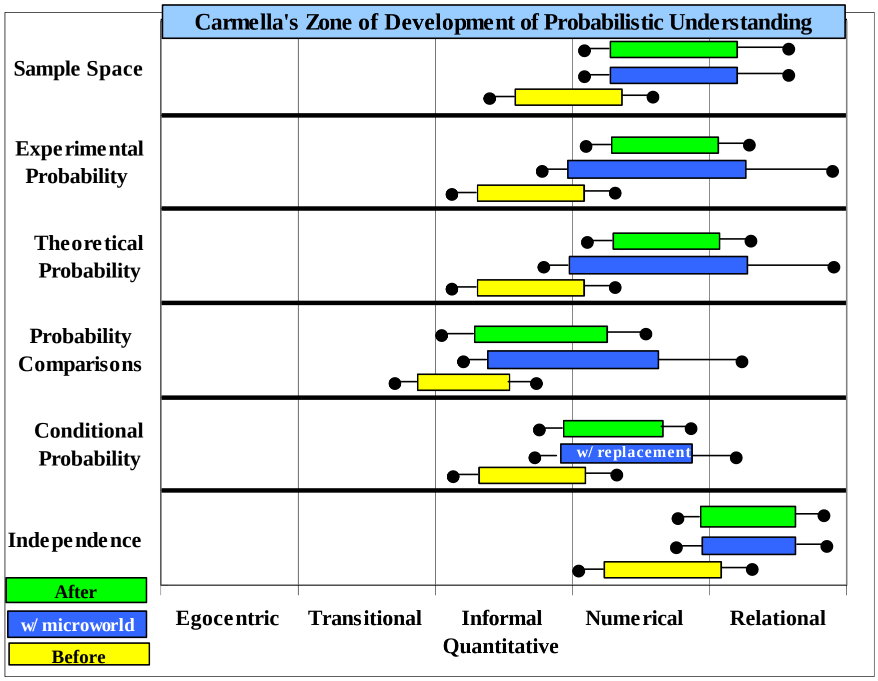


Figure 8.4. Mapping of Carmella's zone of development at each stage of the teaching experiment.

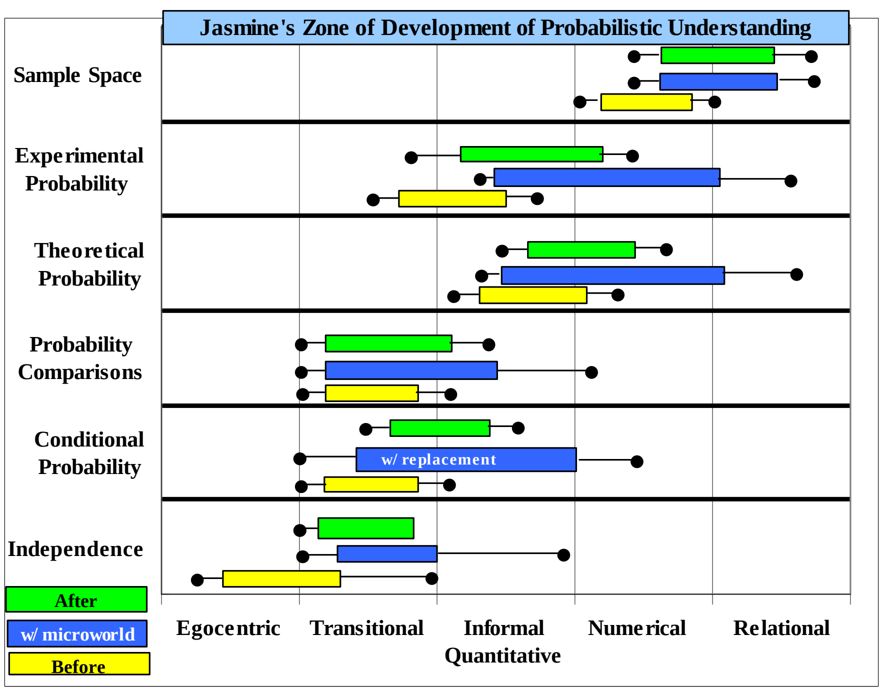


Figure 8.5. Mapping of Jasmine's zone of development at each stage of the teaching experiment.

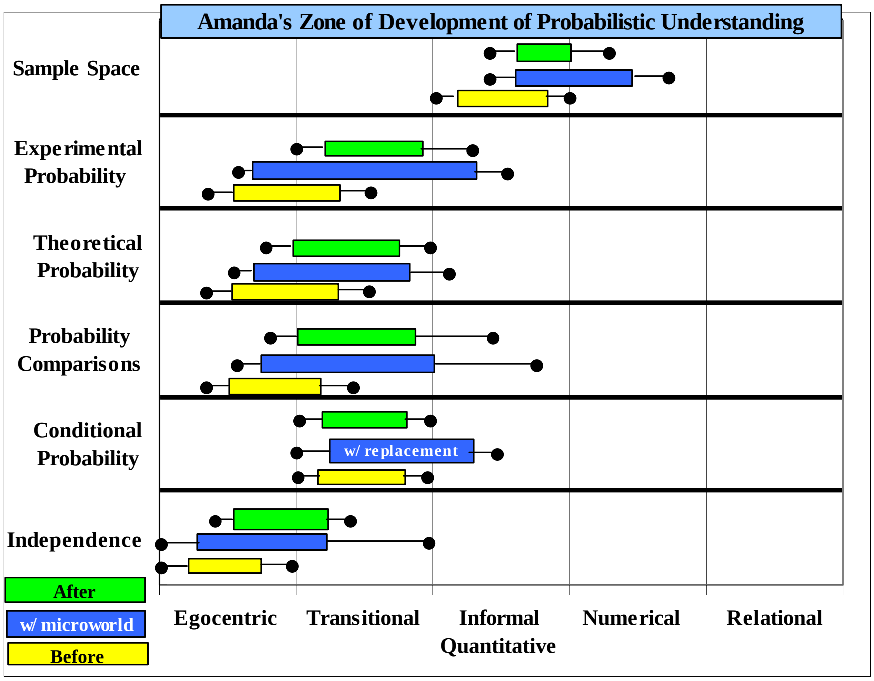


Figure 8.6. Mapping of Amanda's zone of development at each stage of the teaching experiment.

As an example of the construction of the mappings, consider Amanda's development with understanding theoretical probability at each stage of the teaching experiment. Amanda's reasoning with theoretical probability in the pre-interview was a combination between egocentric and transitional. The position of a cube in a bucket, how she held a coin before flipping it, and her inconsistent preferences for "mixed" or "matched" outcomes with two and three coin toss experiments are a few examples of her egocentric judgments that affected her reasoning about the theoretical probability of an event occurring. At times she did use part-part comparisons based on a most-middle-least scale to predict the likelihood of an event (e.g., comparing the number of white and black marbles in a bag) and thus also reasoned at a transitional level.

During the teaching sessions, Amanda's reasoning about theoretical probability was mainly in the transitional level as she sometimes used quantitative judgments to compare theoretical probabilities (e.g., 12/24 is the same as 6/12) and expect similar results "around half" with both theoretical probabilities.. However, she still sometimes made judgments based on personal preferences (e.g., "I like soccer") or prior experimental evidence. She had difficulty interpreting the weight tool and only occasionally made judgments based on informal quantitative reasoning (e.g., her work on tasks involving "twice as likely" and "equally likely").

Amanda's reasoning with theoretical probability in the post-interview was mainly in the transitional level. Recall her egocentric analysis with the family task in which she stated that that the chance of a family having all 3 girls was "natural" because she had a girl friend who had two sisters. Amanda also could use part-whole language to describe a theoretical probability, but there was no evidence that she considered the relationship between the part and the whole. Thus, her reasoning barely stretched up to the informal quantitative level.

All three children had certain tasks or contexts that tended to facilitate their reasoning at an extended level than the majority of their reasoning on other tasks. As an example of the task-based or context-based extensions in the children's reasoning, consider Amanda's extended reasoning with probability comparisons with tasks that included simple comparisons of equiprobability (e.g., 5B5W and 2B2W bag of marbles) and weights that were a special case of "twice" or "half" (e.g., 3-6 and 4-8). Also consider that Carmella's and Jasmine's relational analyses between experimental and theoretical probability almost always occurred when they were using the pie graph and analyzing experimental data during a simulation.

Cross-Case Comparison

At the beginning of the teaching experiment, Carmella's zone of development was primarily in an informal quantitative to numerical level (see Figure 8.4). Jasmine's zone of development in the six concepts ranged from egocentric through numerical (see Figure 8.5), while Amanda's zone was primarily in the egocentric to transitional level (see Figure 8.6). As evidenced by their responses in the pre-interview, different contexts seemed to elicit different levels of reasoning. For example, the children all reasoned better when comparing probabilities with the spinner task than they did with the marble bags. It seems that an area model facilitated their reasoning much better than a discrete model. In addition, Jasmine and Amanda were able to recognize the independence of events in choosing cubes out of a bucket with replacement, but they did not seem to consider independence in the coin toss questions.

During the teaching sessions, the children's use of the microworld tools extended their zones of development, at least somewhat, past their levels of understanding in the pre-interview. Many of those extensions were context-based and tool-based. However, it is evident that their experiences in the microworld did facilitate probabilistic reasoning at higher levels. It is important to recall that Carmella and Jasmine had higher levels of mathematical activity and interactions than Amanda exhibited. Thus, even the small amount of extension in Amanda's zone of development was almost always guided and sustained by a teacher/researcher. It is also important to note that only the children (Carmella and Jasmine) who entered the teaching experiment with a zone of development in the informal quantitative and numerical levels were able to self-sustain reasoning at the numerical and relational level.

The mappings of the children's zone of development in the post-interview demonstrate that in almost all cases, their reasoning was not as strong nor consistent as it was during their use of the microworld. Carmella had the most stability in her reasoning with the microworld to her reasoning with the tasks in the post-interview. I propose that, especially for Amanda and Jasmine, the lack of multiple representations available in the post-interview limited their ability to reason at the levels they exhibited during the teaching sessions. Amanda's improved reasoning during the teaching sessions was also due to sustained interactions with a teacher/researcher. Thus, since she did not make sense of many of the situations on her own, I suggest that she was not able to alter her schemes in a transformative way. However, I want to emphasize that Amanda's growth from mainly egocentric to the transitional and informal quantitative reasoning levels was a critical step towards a more analytical approach to probability.

Chapters 5, 6, and 7 detailed the probabilistic reasoning of the three children in this study while the current chapter focused on interactions with the microworld and the children's zone of development of probabilistic reasoning. In the final chapter, I briefly summarize the results from this research to answer the original research questions. I also discuss implications from this research in four areas: 1) iterative changes in the microworld based on research; 2) how microworld interactions could enhance implications from prior research; 3) contributions to understanding children's development of probabilistic reasoning; and 4) how children's use of microworld environments can facilitate probabilistic reasoning. Future research on children's probabilistic reasoning in a technological environment and iterative development of the *Probability Explorer* will also be discussed.

CHAPTER 9

CONCLUSIONS AND IMPLICATIONS

The goals of this research study were to understand children's probabilistic reasoning, to study how children used the microworld tools in *Probability Explorer* to investigate probability tasks, and to further refine the microworld based on children's experiences. The case study chapters (5, 6, and 7) included descriptions of the children's probabilistic reasoning at each stage in the teaching experiment (i.e., before, during, and after using the microworld). The previous chapter further described their interactions with the microworld and the children's development of probabilistic reasoning. The purpose of this chapter is to discuss conclusions from the research, implications, and future research and development.

Research Conclusions

Previous researchers have studied elementary children's understanding of probability, while other researchers have studied elementary children's interactions and meaning-making activities in computer microworlds in a variety of domains. However, this is the first research study to investigate elementary children's probabilistic reasoning with a computer microworld. In this section I will highlight the findings from this study that answer the original research questions for this study:

1. What are children's understandings of probabilistic concepts (e.g., fairness, equivalence, sample space, experimental probability, theoretical probability,

probability comparisons, and independence) and how do they develop appropriate probabilistic reasoning?

2. How are children's conceptions affected by their use of *Probability Explorer* as a problem-solving tool? What are the benefits and drawbacks of the instructional design and utility of tools in *Probability Explorer* for facilitating appropriate probabilistic reasoning in children?

I will concurrently answer the first question regarding the children's understanding and development of probabilistic reasoning and the first part of the second question with regard to how their conceptions were affected by their use of *Probability Explorer*.

In response to the first research question, the children in this study exhibited different levels of understanding in a variety of contexts. Although all three children were the same age and had attended the same elementary school, their understandings at the beginning of the teaching experiment were very different. There are many unknown factors that may affect children's probabilistic reasoning on particular tasks. For example, Jasmine referred to her experience watching "Bill Nye the science guy" in her prediction of the increased likelihood of flipping a tail after a sequence of heads had occurred. However, if she would not have shared that experience with me, I may have only attributed her reasoning on that task to an egocentric view of independence. Thus, children's understandings of probability concepts and their development of appropriate probabilistic reasoning is extremely sensitive to their everyday experiences that can either confirm or disconfirm their intuitions or help them build experience-based subjective views of probability that may or may not be aligned with normative probability theory.

In response to the first part of the second research question, one needs to consider the mappings in Figures 8.4, 8.5, and 8.6 that approximate the children's zone of development while they were using the microworld tools. With each child, and every probability concept, their zone of development of probabilistic reasoning was extended, at least somewhat, when using *Probability Explorer*. In some cases, that extension only occurred when they were investigating certain tasks and using certain tools in the microworld. Another important finding from this research is that only the children (Carmella and Jasmine) with an initial zone of development within levels three and four were able to stretch their developmental zone into a relational level where they considered relationships between probability concepts. Amanda's use of the tools only slightly extended her zone of development of probabilistic reasoning.

To answer the second part of the second research question, I want to first discuss the benefits and drawbacks in terms of developing appropriate probabilistic reasoning. The design of the microworld as it was described in Chapter 3 seemed to be beneficial to Carmella and Jasmine. However, Amanda needed an alteration in the environment to help her interpret the weight tool and to slightly improve her understanding of theoretical probability. Carmella's advanced level of interactivity with the microworld was probably due to many factors, including her already developed understanding of several probability concepts, her willingness to take risks and experiment in the microworld, and her ability to easily interpret the representations used in the microworld (e.g., pie graph and weight tool). The multiple representations and flexibility of the tools in the microworld seemed to also facilitate Jasmine's development of probabilistic reasoning while using the tools. However, her stretched zone of development during the teaching sessions resulted in only slightly improved understandings after her use of the microworld as compared to her prior understandings. The abstract nature of the simulation process and dynamic connection between representations seemed to distract Amanda from the underlying mathematics. This may have been caused by her egocentric view of randomness, her lack of previous experiences with probabilistic situations, her lack of confidence in risk-taking, her playful orientation, her misunderstanding of the representations (e.g., pie graph, weight tool), or a variety of unknown factors that could have contributed to her mainly subjective approach to probability tasks. Overall, the original tools designed in *Probability Explorer* served as a benefit and drawback to different students and in different situations.

The research questions have been answered within the bounds of the purposeful limited scope of the study. The answers to those questions may have very well been different with more children participating in the study, including children of different age groups with different educational and personal experiences. Nevertheless, I believe the results are sufficient to fulfill the purpose of such an exploratory study of a newly designed microworld environment. Based on these results there are several implications that will contribute to my personal continued research and development with *Probability Explorer*.

Implications from Research

There are several implications from this research. First, I will discuss the changes in the microworld environment based on my research results. The second implication is how children's use of the microworld can enhance implications from prior research. The third implication entails how my findings contribute to the body of research on children's understanding of probability and development of probabilistic reasoning. The final implication considered is how children's use of microworld tools can facilitate probabilistic reasoning.

Iterative Development of the Microworld

The intent of this study with respect to the microworld was not only to engage in concurrent research and development with *Probability Explorer*, but to identify strengths and weaknesses based on the children's work that would inform future iterations of the software. Roschelle and Jackiw (in press) argue that successful technology designs need to go through an iterative and transformative process concurrently with research on students' learning with the technology. Based on the children's interactions with the microworld and their subsequent development, or lack thereof, of probabilistic reasoning, I propose several changes in the environment that may better facilitate children's meaning-making activity.

"Line up" outcomes. Neither Jasmine nor Amanda made substantial progress in their understanding of independence and that both still thought that "mixed up" strings of results were more likely to occur than those in a pattern. The current representations in the microworld help children analyze data grouped by outcome, and not data in the order in which it was randomly generated. In my original design of *Probability Explorer*, I wanted to make the iconic representations of the data moveable so that the children could re-present the data as they saw fit. However in my attempt to give students more flexibility than existed in other software such as *Chance* (Jiang, 1992) and *Probability Constructor* (Logal, 1997), I removed an important representation. The only time the children ever attempted to line the data up in order was during simulations with 10 or less trials. The latest version of *Probability Explorer* includes a "Line Up" button on the toolbar that can be used to place the randomly generated icons in the order in which they occurred, from left to right on the screen in rows of 10 (Figure 9.1).

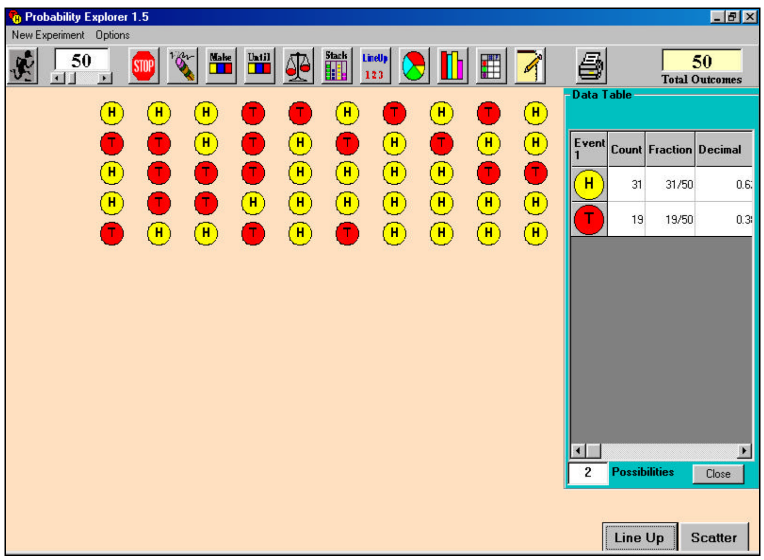


Figure 9.1. Results from a fair coin toss lined up in the order in which they occurred.

In the results displayed in Figure 9.1, notice that after seven consecutive heads (trial 34, 35, 36, 37, 38, 39, 40) only three tails occurred in the next ten outcomes. Students who may tend to employ the negative recency fallacy might benefit from observing such occurrences. However, those who employ the positive recency fallacy might view this occurrence as a confirmation to their intuitive theory. Students would need repeated experiences with analyzing the ordered data to observe instances that disconfirm both types of fallacies. In addition, the arrangement of 10 outcomes per row is a purposeful attempt to provide a visual arrangement that could facilitate discussions about comparing the likelihood of several sequences of 10 results. This could further help students overcome biases with a representativeness heuristic.

Because I want to maximize the possible representations available in the microworld that students can use simultaneously in their analysis, I will eventually create a permanent history box that will contain an ordered list of the icons that students can view by scrolling. The students will be able to display or hide this history box just as they can the graphs and data table. With the addition of the history box, students will be able to re-present the icons on the screen in either the stacking columns or their own arrangements but also be able to analyze the data in an ordered sequence.

Simulation speed. Amanda was fascinated with the “motion” of the pie graph during the simulation process and she seemed to divorce the pie graph from the data it represented. In addition, she never made explicit connections with the “motion” and the theoretical probability. To give students more time to analyze changes in the graphs and data table during a simulation process, and to hopefully help them connect these changes directly with the newest randomly generated outcome, I added a feature to allow the user to change the speed at which the simulation occurs (Figure 9.2). The “slow” setting causes a purposeful delay of 0.6 seconds between each random event. The “medium” setting causes a 0.3-second pause and the “fast” setting does not trigger any pause in the system and will occur as quickly as possible depending on a computer’s processing speed. Slowing down the simulation process will hopefully prompt students to attend to the changes in the graphs and data table and give them more time to reflect on the effect of the number of trials on the results, rather than just watching the “motion” of the scattering icons on the screen and the dynamic changes in the graph. The speed setting can even be changed during the simulation process if the user wants to use “slow” to carefully analyze the data during a particular part of the simulation and then wants to change back to “fast” to quickly finish a large number of trials.

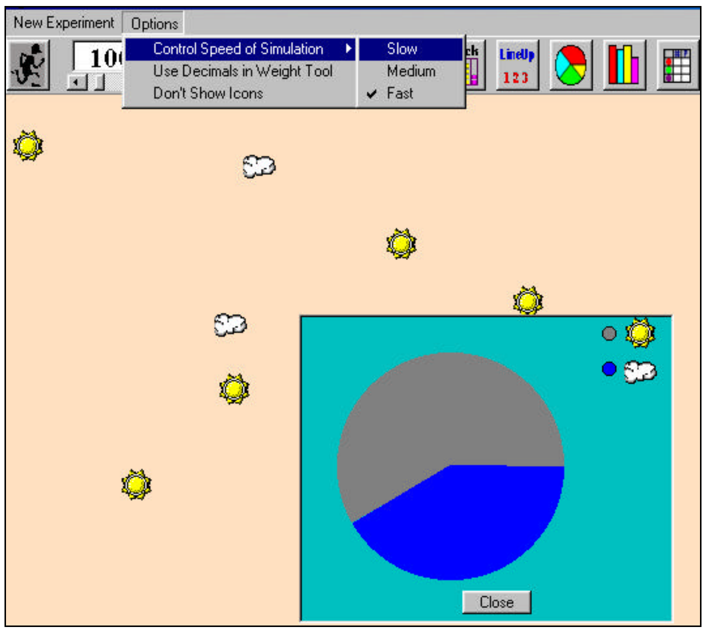


Figure 9.2. Feature to control the speed of the simulation added to Options menu.

Flexible data control. Several times throughout the teaching sessions, the children were conducting experiments by running several sets of 100 trials. They often recorded the numerical results on paper or in the Notebook, but needed to rely on their memory of the graphs and iconic representations. Thus, there needs to be some sort of “copy/paste/save” capability built into the software so a child could copy a picture of the current pie graph and also place that picture in the Notebook for later reference, or into a word processing or presentation software application. In order to encourage older students to do further analysis with the numerical data from each set of trials, there also needs to be a capability to create another data table that the students can record, or copy/paste/save, numerical results and then perform additional calculations and/or generate graphical displays. Students also need to be able to export data to a spreadsheet application for further data analysis. Until the technological design issues of these suggestions can be addressed, I have added a print feature to allow users to print a copy of the entire screen, including graphics and data table on a local printer. The print feature will at least provide some sort of visual record for students to reference during a meta-analysis of several sets of trials.

Additional improvements. Several other changes are planned to increase the flexibility and usability of the tool. These include: 1) making the data table flexible enough to be resized so that students can see all representations (count, fraction, decimal, percent) concurrently; 2) displaying results with compound event experiments in pie and bar graphs; 3) adding a spinner environment in which students can design their own spinners with various sized sectors and colors; 4) adding more icons to the “Design Your Own” options to allow students to model a wider variety of experiments; and 5) increase the total number of outcomes possible in an experiment to 10 (the current maximum is eight). Of course, minor “bugs” and nuances in the environment have been, and will be, continually corrected (e.g., miscellaneous lines that appear in the bar graph).

Microworld Interactions can Enhance Implications from Prior Research

All the research on elementary-aged (grades K-6) children’s understandings of probability done prior to this study was in non-technological environments. Now that a technological tool appropriate for children to explore probability is available, and this initial research on children's use of the tools is complete, it is important to revisit prior research. The advent of *Probability Explorer* and results from this exploratory research study can shed some light on prior research results and perhaps bring forth possible ways that the microworld can be used to address researcher's prior implications and suggestions.

Several researchers (e.g., Fischbein & Schnarch, 1997; Tversky & Kahneman, 1982) have studied students' intuitions about the effect of sample size on the likelihood of an event occurring. Many students put too much faith in experimental results from a small number of trials and also will consider proportionally equal results (e.g., two heads-1 tail and 200 heads-100 tails) as equally likely to occur. Jasmine's and Carmella's investigations with the EOP and their recognition that an exact distribution is more likely with a small number of trials demonstrates that, for the most part, they did not fall prey to this type of fallacious reasoning. Amanda showed a few instances where it appeared that she understood the effect of the number of trials, but for the most part, her reasoning in a variety of task situations demonstrated that she puts too much faith in results from a small number of trials in predicting future results. The use of the microworld to easily simulate any given number of trials can help students investigate questions about the effect of the number of trials on the likelihood of an event occurring. In fact, throughout the teaching sessions, the children almost always brought the issue of the number of trials into their investigations without prompting from either teacher/researcher. It seems that merely having the flexibility to run a variety of simulations with both small and large number of trials, as well as the tools available for the subsequent analysis of the data, brought forth children's intuitions about the effect of the number of trials.

The effect of the number of trials is the cognitive link between empirical and theoretical probability. Although researchers (e.g., Jones *et al*, 1999a) have attempted to identify children's understanding of both empirical and theoretical probability, I posit that most children can only develop a deep conceptual understanding of the relationship between both types of probability by repeated experimentation with an increasingly larger number of trials, and with many different types of experiments. Jones *et al* suggest a "pooling of data" (p.153) strategy where pairs of children conduct an experiment with physical objects 20 times, and the pair data is pooled to help students compare their individual results from 20 trials to the class data which may be in excess of 200 trials. They even suggest analyzing the data after each successive "pool" (40, 60, 80, … , 200). Although this strategy can give children a sense of the effect of the number of trials, due to enormous time constraints in creating and revising graphs, the analysis involved in this method is probably through a tally mark and counting system to determine frequencies.

With *Probability Explorer*, this investigation can take on a different type of "pooled data" strategy. Pairs of children working in the microworld can quickly run cumulative sets of 20 trials and analyze the results numerically, graphically, and for the first few sets of 20 trials, by direct manipulation of the iconic representations on the screen (see Figure 9.3). By observing the dynamic link between the numerical and graphical representations during the simulation, each pair of children can now experience the intent of the "pooled" data strategy by analyzing numerical and graphical *frequencies* and *relative frequencies* as the number of trials increases.

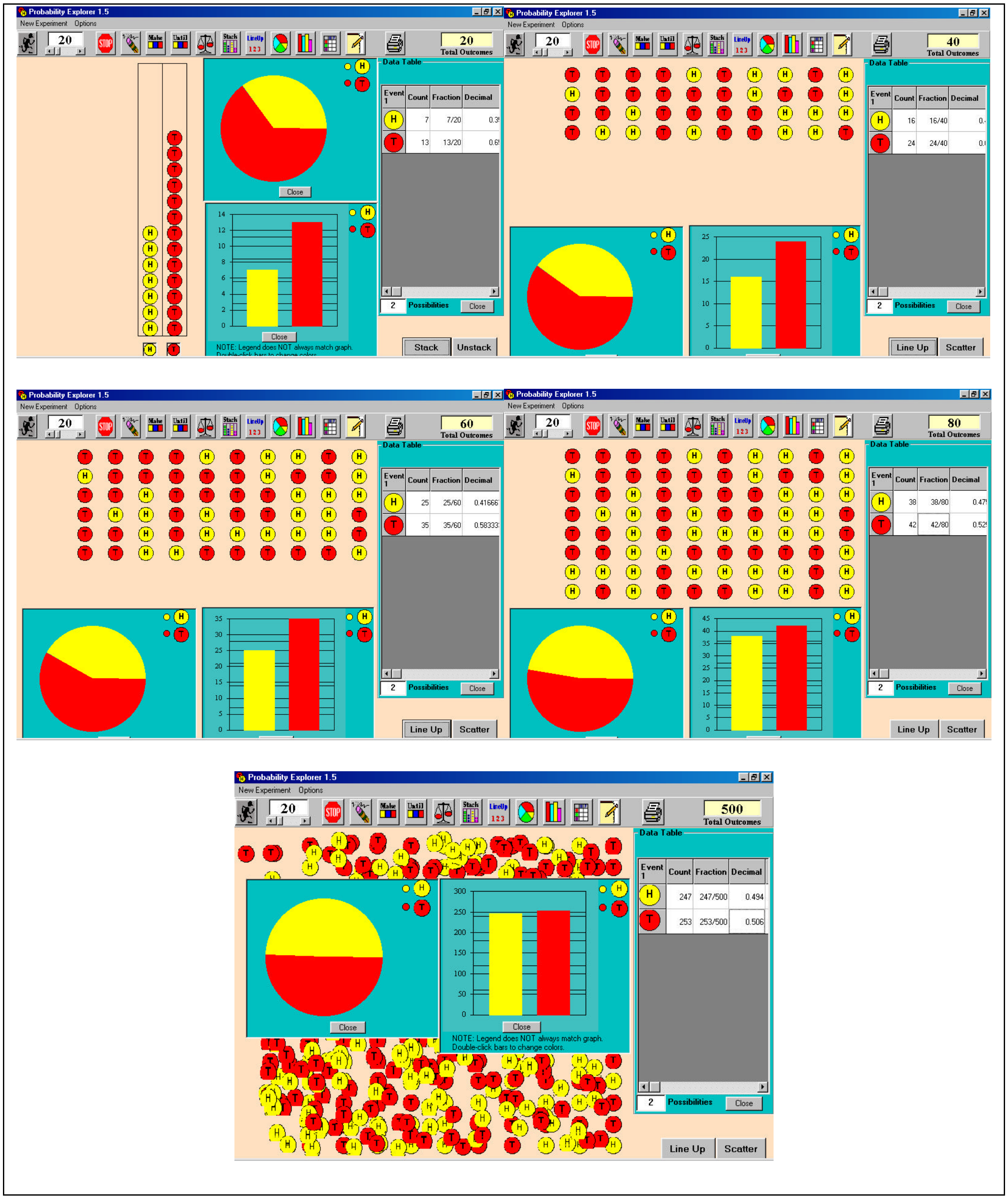


Figure 9.3. Example of "pooled data" (20, 40, 60, 80, … , 500) with *Probability Explorer*.

Notice that in the successive screenshots in Figure 9.3, the distribution of heads and tails did not seem to "even out" from 40 to 60 trials but were more equally distributed by 80 trials. After 80 trials, other children may be getting different results on their computer screen (e.g., more tails than heads). Of course, after the pairs of children run, say 500 trials, they can all glance at other computers and observe that the numerical and graphical results from different pairs of children looks incredibly similar. I conjecture that the suggestive power of this "pooled data" is apt to elicit conceptual understanding of the relationship between empirical and theoretical probability and to facilitate student's development of a theory-in-action.

Many researchers (e.g., Fischbein, 1975; Jones *et al*, 1997, 1999b; Piaget & Inhelder, 1975) have used tasks with bags of marbles to assess children's ability to compare probabilities. Many of the most common strategies reported in prior research were also used by children in my study (e.g., favoring the bag with the least number of undesired events or favoring the bag with most number of desired events). While using the bag of marbles in the microworld, Carmella and Jasmine were able to reason that bags with marbles in proportion to one another had the same chance for choosing a certain color. Amanda sometimes reasoned proportionally (e.g., 3-6-9 has the same chances as 4-8-12) and other times did not (e.g., 6B2W had a better chance for choosing black than 3B1W). Only Carmella used proportional reasoning on these tasks in the post-interview, while Amanda and Jasmine did not; although they did use proportional reasoning to design a bag of marbles given a pie graph of experimental results. It seems that the marble environment, the link to the weight tool, and the ability to quickly run many experiments with the bags, can help students analyze and compare the probability of an event occurring in two different bags and make subsequent comparisons. I suggest that the microworld tools can help children develop reasoning with fractions and proportional reasoning while they approach these tasks experimentally.

When dealing with compound events, many researchers (e.g., Fischbein, 1975; Jones, *et al*, 1997, 1999b; Piaget & Inhelder, 1975) found that children have difficulty generating all the possible outcomes in the sample space and do not always consider the importance of order in their analysis of probability. Jasmine and Amanda tended to not consider order during the pre-interview while Carmella did consider the order of events. Only one task was used during the teaching session to address constructing sample space and the importance of order. During this task, Carmella and Jasmine were able to use systematic strategies to construct all the possible outcomes using the Make It tool. They were also able to recognize and explain the difference between stacking the outcomes with order mattering and in unordered groups. Amanda initially created all possible outcomes without regard to order, but when I made an additional outcome by changing the order of one of her previously made outcomes, she noted "oh, well if you do it that way then…" and she continued to make the remaining possible outcomes in all possible ordered arrangements. Her intuition was to consider unordered arrangements while Carmella and Jasmine intuitively considered ordered ones. It seems like the ability to construct a sample space of compound events and to analyze the data as both ordered and unordered arrangements, has the potential to help children explicitly investigate the differences in probability of events as well as real world situations where order matters, or it does not matter.

The research on children's use of the representativeness heuristic indicates that there are many types of strategies and intuitions that children elicit when approaching probability tasks. As noted in the previous section regarding changes in *Probability Explorer*, the new “Line Up” feature can help children analyze ordered data in conjunction with the grouped data displayed in the graphs and data table. This feature was added to the microworld because the children in this current study, especially Amanda, did not significantly change any of their intuitive-based notions related to representativeness. Further research with this specific feature in the microworld is needed to assess whether the analysis of ordered data can help children reason appropriately with regard to representativeness.

Overall, there are many possible ways that *Probability Explorer* can help facilitate children’s probabilistic reasoning. The prior research with children showed that children have many intuitive-based conceptions of probability, some aligned with normative probability theory and most not aligned. Researchers such as Hawkins and Kapadia (1984) also posited that probability instruction needed to tap into those intuitions by allowing children to take a subjective approach to probability tasks as well as the traditional *a priori* and frequentist approaches. The microworld environment and the many tools available for designing experiments and analyzing data provide enough flexibility for children to approach tasks in all three ways.

Understanding Children’s Development of Probabilistic Reasoning

The framework, as I have amended it from Jones *et al* (1999b), is only the beginning of what is needed in understanding the complexities of children’s development of probabilistic reasoning. The mere use of my approximate zone of development implies that it is difficult to assess children’s probability understanding and rate of development of probabilistic reasoning. This also implies that some types of tasks, with or without technology, seem to place children within their own zone of potential construction (Olive, 1994), while other tasks can merely frustrate children and even elicit reasoning based on judgmental heuristics such as representativeness and availability. Based on the findings from this study, it seems that tasks that include “twice” and “half” relationships tend to elicit multiplicative and part-whole reasoning that facilitates a child’s probabilistic reasoning. These types of tasks may be able to give children appropriate experiences and serve as comparison experiences with tasks that do not involve such easily recognizable relationships. The findings from this study also indicate that tasks which utilize an area model, either through spinners or pie graph representations of data, can help facilitate children’s part-whole reasoning and their development of probabilistic reasoning. It seems that perhaps an area model more closely connects with children’s understanding of rational numbers than a discrete model.

The findings from this study support other research findings (e.g., Green, 1983; Jones *et al*, 1999a; Konold 1987, 1991; Shaughnessy,1977, 1992) that suggest the teaching of probability be grounded in an experimental approach that allows students to reason from an *a priori*, frequentist, and subjective approach. The children’s use of *Probability Explorer* suggests that an experimental, exploratory approach to teaching probability can enhance their probability understandings by using a variety of tasks and contexts, as well as making explicit connections between multiple representations of data.

Microworld Tools can Facilitate Probabilistic Reasoning

Although this research was only done with a limited number of children, the results *suggest* that some children's use of carefully designed microworld tools can facilitate their probabilistic reasoning while using the tools. However, the different ways that the children in this study used the tools and how those tools helped or hindered their probabilistic reasoning *implies* that many factors, including their current schemes and disposition towards open-ended learning, contribute to how interactions with a microworld can facilitate an individual's development. Land and Hannifin (1996) noted that "some learners meet the cognitive and metacognitive demands of open-ended learning; many others, however, do not. OELEs [Open-Ended Learning Environments] must intentionally facilitate on-going learner needs, reflection, and interpretation" (p. 44). With this in mind, it is imperative to continue to do research on children's use of *Probability Explorer* and their development of probabilistic reasoning. The original design of the environment as used during the teaching experiment did not facilitate Amanda's reflection and interpretation. Perhaps if she would have had different tasks or more experiences with the environment as it has been changed (e.g., slower simulation speed, ordered outcomes), she may have been more reflective and interpretative of the experimental process and results.

As shown with Carmella and Jasmine, dynamically linked multiple representations have the *potential* to facilitate children's probabilistic reasoning and encourage them to develop theories-in-action about the law of large numbers, especially when reflecting on changes in the multiple representations during a simulation. In addition, the research results from this study have shown that the pie graph seems to be a very important representation for developing understanding of the law of large numbers as well as facilitating multiplicative and part-whole reasoning. Children's multiplicative and part-whole reasoning seems more consistent with an area model (pie graph or spinner) rather than a discrete model (bag of marbles).

Using *Probability Explorer* can substantially extend typical experiences with physical objects and lead children to play, experiment, predict, and discover probabilistic ideas. Although I believe *Probability Explorer* can be useful in a classroom situation, there is no research evidence to support this belief. Nevertheless, I conjecture that teachers who carefully engage students in active experimentation, reflection, and classroom discourse can help students develop probabilistic reasoning. I suggest that they use a variety of tasks and contexts for their probability explorations and that they *explicitly* focus students on making connections between the multiple representations, both during and after a simulation.

Future Research and Development

Further research is needed on children's understanding of probability, the types of tasks, and technological and non-technological learning environments that truly facilitate the development of appropriate probabilistic reasoning. Research is also needed to connect students' understanding of probability with their understandings of other mathematics concepts and their everyday experiences. Both of these research suggestions need to be done with children in a variety of grade levels and a variety of cultural and socio-economic situations. Since children's primary intuitions are affected by their everyday experiences, it is important to study children who have a variety of real life and educational experiences. This type of research will help in further developing and refining a developmental model of probabilistic reasoning, and how children's real life experiences can be used to extend the types of tasks used in instruction beyond use of coins, dice, spinners.

Further research needs to be conducted with children using *Probability Explorer* in a variety of small group and classroom situations. This research should be conducted in situations where myself, as the designer of the microworld, serves as teacher/researcher, and where other researchers and regular classroom teachers serve as the primary teacher. In addition, I agree with Rochelle and Jackiw (in press) who call for concurrent innovation in curricula, technology and pedagogy and urge technology designers to work as part of a team in researching and developing these innovations. They further note that technology design projects that are developed in isolation of developments in curricula and pedagogy rarely make it into the mainstream classroom, and thus, have little potential of making an impact on educational practices. Because I believe that *Probability Explorer* has the potential to impact the teaching and learning of probability in elementary schools, it is imperative that the future developments in the microworld are aligned with developments in school curricula and pedagogy.

My intent for future research and development with *Probability Explorer* includes a concerted effort to involve other researchers, technology designers, mathematics educators, and curriculum specialists in a team project that addresses the above suggestions for further research. Such a project could transform the software through an iterative process and build a curriculum and support system to maximize the potential impact on children's probability understanding. This exploratory study has planted a seed and outlined a trajectory that will guide my research for many years to come.

REFERENCES

Battista, M. T. (1998). *Computer environments that engender students' construction of mathematical ideas and reasoning: A constructivist perspective.* Paper presented at the ENC Technology and NCTM Standards 2000 Conference. Arlington VA, June 5-6, 1998.

Behr, M. J., Harel, G., Post, T., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning*. (pp.296-333). New York: Macmillan.

Bezuk, N. S. & Bieck, M. (1993). Current research on rational numbers and common fractions: Summary and implications for teachers. In D. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp.118-136). New York: Macmillan.

Biddlecomb, B. D. (1994). Theory-based development of computer microworlds. *Journal of Research in Childhood Education, 8(2),* 87-98.

Biehler, R. (1991). Computers in probability education. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education,* (pp.169-211). Boston: Kluwer Academic Publishers.

Blumer, J. (1969). *Symbolic interactionism.* Englewood Cliffs, NJ: Prentice Hall.

Borovcnik, M. & Bentz, H.-J. (1991). Empirical research in understanding probability. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 27-71). Boston: Kluwer Academic Publishers.

Borovcnik, M., Bentz, H.-J., & Kapadia, R. (1991). A probabilistic perspective. In R. Kapadia & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 27-71). Boston: Kluwer Academic Publishers.

Bowers, J. (1995). Coordinating top-down and bottom up approaches: An alternative perspective for developing a mathematical microworld. In J. L. Schnase & E. L. Cunnius (Eds.), *Proceedings of the First International Conference on Computer Support for Collaborative Learning* (121-125). Bloomington, IN: Indiana University.

Bramald, R. (1994). Teaching Probability. *Teaching Statistics, 16*(3), 85-89.

Castro, C. S. (1998). Teaching probability for conceptual change. *Educational Studies in Mathematics 35(3),* 233-54

Clements, D. H. & Battista, M.T. (1994). Computer environments for learning geometry. *Journal of Educational Computing Research 10(2),* 173-97

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23(7),* 13-20.

Cramer, K., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp.159-178). New York: Macmillan.

Einhorn, (1982). Learning from experience and suboptimal rules in decision making. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 268-283). New York: Cambridge University Press.

Erickson, F. (1986). Qualitative methods. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd. ed. pp. 119-161). New York: Macmillan.

Estes, W. K. (1964). Probability learning. In A. D. Melton (Ed.), *Categories of human learning* (pp. 89-128). New York: Academic Press.

Falk, R. (1979). Revision of probabilities and the time axis. In *Proceedings of the Third International Conference for the Psychology of Mathematics Education* (pp. 64-66). Warwick, U. K.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Boston: D. Reidel Publishing Company.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96-105.

FolioViews. (1998). Software program (Version 4.2). Open Market, Inc.

Goldin, G. (1998). Observing mathematical problem solving through task-based interviews. In A. R. Teppo (Ed.), *Qualitative research methods in mathematics education*. (Journal for Research in Mathematics Education, Monograph No. 9, pp. 40-62). Reston, VA: National Council of Teachers of Mathematics.

Green, D. R. (1983). School pupils' probability concepts. *Teaching Statistics, 5(2),* 34-42.

Graue, M. E. & Walsh, D. J. (1998). *Studying children in context: Theories, methods, and ethics.* Thousand Oaks, CA: Sage Publications.

Hawkins, A. S., & Kapadia, R. (1984). Children's conception of probability – A psychological and pedagogical review. *Educational Studies in Mathematics, 15*, 349-377.

Jiang, Z. (1992). *Students' learning of introductory probability in a mathematical microworld.* Unpublished doctoral dissertation, University of Georgia, Athens, GA.

Jiang, Z. & Potter, W. D. (1994). A computer microworld to introduce students to probability. *Journal of Computers in Mathematics and Science Teaching 13*, 197-222.

Jones, G. A, Langrall, C. W., Thornton, C. A., & Mogill, T. A. (1997). A framework fro assesing and nurturing young children's thinking in probability. *Educational Studies in Mathematics 32*, 101-125.

Jones, G. A, Langrall, C. W., Thornton, C. A., & Mogill, T. A. (1999a). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30(5)*, 487-521.

Jones, G. A, Thornton, C. A, Langrall, C. W., & Tarr, J. E. (1999b). Understanding students' probabilistic reasoning. In L. V. Stiff (Ed.) *Developing mathematical reasoning in grades K-12* (1999 Yearbook, pp. 146- 155). Reston, VA: National Council of Teachers of Mathematics.

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.

Konold, C. (1987). *Informal concepts of probability* (Paper based on a Ph.D. Thesis.). Massachusetts: University of Massachusetts.

Konold, C. (1991). Understanding students' beliefs about probability. In E. von Glaserfeld (Ed.), *Radical Constructivism in Mathematics Education* (pp. 139-156). Holland: Kluwer.

Konold, C., Polletsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. *Journal of Research in Mathematics Education, 24*(5), 392-414.

Land, S. M. (1995). *The process of developing theories-in-action with open-ended learning environments: An exploratory study*. Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.

Land, S. M., & Hannafin, M. J. (1996). A conceptual framework for the development of theories-in-action with open-ended learning environments. *Educational Technology Research and Development 44(3)*, 37-53.

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

Lidster, S. T., Pereira-Mendoza, L., Watson, J. M., & Collis, K. F. (1995). *What's fair for grade 6?* Paper presented at the annual conference of the Australian Association for Research in Education, Hobart, November.

LOGAL. (1997). *Probability Constructor.* Cambridge, MA.

Mack, N. (1990). Learning fractions with understanding: Building on informal knowledge. *Journal for Research in Mathematics Education 21,* 16-32.

Mack, N. (1995). Confounding whole-number and fraction concepts when building on informal knowledge. *Journal for Research in Mathematics Education 26,* 422-441.

Madsen, R. W. (1995). Secondary students' concepts of probability. *Teaching Statistics*, *17*(3), 90-93.

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

National Center for Educational Statistics, (4/18/99). *1996 Assessment Mathematics–Public Release: Grade 4* (URL: http://nces.ed.gov/nationsreportcard/ sampleq/index.shtml).

National Council of Teachers of Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: author.

National Council of Teachers of Mathematics. (1991). *Professional Standards for Teaching Mathematics*. Reston, VA: author.

National Council of Teachers of Mathematics. (1994). *Professional Development for Teacher of Mathematics* (1994 Yearbook*)*. Reston, VA: author.

National Council of Teachers of Mathematics. (1998). *Principles and Standards of School Mathematics*: *Discussion Draft*. Reston, VA: author.

Olive, J. (1994). Building a new model of mathematics learning. *Journal of Research in Childhood Education 2(2)*, 162-173.

Olive, J. (1999). *Computer tools for interactive mathematical activity in the elementary school.* Paper presented at the annual meeting of the National Council of Teachers Mathematics, San Francisco, CA, April 22-24.

Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas.* New York: Basic Books.

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* Newbury Park, CA: Sage.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: Norton.

Pothier, Y. & Sawada, D. (1983). Partitioning: The emergence of rational number ideas in young children. *Journal for Research in Mathematics Education, 14*, 307-317.

Riedesel, C. A., Schwartz, J. E., & Clements, D. H. (1996). *Teaching Elementary School Mathematics*. Boston: Allyn and Bacon.

Rochelle, J. & Jackiw, N. (in press). Technology design as educational research: Interweaving imagination, inquiry, and impact. In A. Kelly & R. Lesh (Eds.) *Handbook of research design in mathematics and science education*. Hillsdale, NJ: Erlbaum.

Schlottman, A. & Anderson, N. A. (1994). Children's judgment of expected value. *Developmental Psychology, 30(1),* 56-66.

Schwartz, J. E. & Riedesel, C. A. (1994). *Essentials of classroom teaching: Elementary mathematics.* Needham Heights, MA: Allyn and Bacon.

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,* 285-316.

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

Shaughnessy, J. M. (1993). Connecting research to teaching. Probability and statistics. *Mathematics Teacher, 86*(3), 244-248.

Shaughnessy, J. M. & Bergman, B. (1993). Thinking about uncertainity: Probability and statistics. In S. Wagner (Ed.), *Research ideas for the classroom: High school mathematics* (pp. 177-197). New York: Macmillan.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1976). Cognitive processes and societal risk taking. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior*. Hillsdale, NJ: Erlbaum.

Steffe, L. P. (1988). Children's construction of number sequences and multiplying schemes. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades*, (pp.119-140). Reston, VA: National Council of Teachers of Mathematics.

Steffe, L. P., & Thompson, P. W. (in press). Teaching experiment methodology: Underlying principles and essential elements. In A. Kelly & R. Lesh (Eds.) *Handbook of research design in mathematics and science education.* Hillsdale, NJ: Erlbaum.

Steffe, L. P., & Wiegel, H. G. (1994). Cognitive play and mathematical learning in computer microworlds. *Educational Studies in Mathematics, 26*, 111-134.

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). Boston: Kluwer Academic Publishers.

Thompson, P. W. (1994). The development of the concept of speed and its relation to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics.* Albany, NY: SUNY Press.

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

Tversky, A., & Kahneman, D. (1982). Judgment under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 3-22). New York: Cambridge University Press.

Tzur, R. (1995). *Interaction and children's fraction learning.* Unpublished doctoral dissertation, University of Georgia, Athens, GA.

Vahey, P. (1997). *Toward an understanding of productive student conceptions of probability: The Probability Inquiry Environment.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL , March 24-28.

Vahey, P. (1998). *Promoting Student Understanding of Elementary Probability Using a Technology-Mediated Inquiry Environment.* Unpublished doctoral dissertation, University of California-Berkeley, Berkeley, CA.

Van de Walle, J. A. (1997). *Elementary and middle school mathematics: Teaching developmentally.* New York: Addison Wesley Longman, Inc.

Virginia Department of Education (1995). *Virginia Standards of Learning*. Richmond, VA: author.

Virginia Department of Education (1996). *Making Sense of Data*. Richmond, VA: author.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning.* Washington, DC: Falmer Press.

Watson, J. M. and Moritz, J. B. (1997). *The C&D PD CD: Professional development in chance and data in the technological age.* CD-ROM produced by the University of Tasmania, Tasmania, Australia.

Williams, J. S. (1995). *11-12 year old children's informal knowledge and its influence on their probabilistic reasoning.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, April 18-22.

Yackel, E., Cobb, P, & Wood, T. (1993). Developing a basis for mathematical communication within small groups. In T. Wood, P, Cobb, E. Yackel, & D. Dillon. (Eds.) *Rethinking elementary school mathematics: Insights and issues* (Journal for Research in Mathematics Education, Monograph No. 6, pp. 33-44). Reston, VA: National Council of Teachers of Mathematics.

Zawojewski, J. & Heckman, D. S. (1997). What do students know about data analysis, statistics, and probability? In P. A. Kenney & E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the National Assessment of Educational Progress* (pp. 195-223). Reston, VA: National Council of Teachers of Mathematics.

APPENDIX A

PARENTAL INFORMED CONSENT AGREEMENT

University of Virginia

CHILDREN'S PROBABILISTIC REASONING WITH A COMPUTER MICROWORLD

Please read this consent agreement carefully before you decide to let your child participate in this study.

Purpose of the research study: The purpose of this study is to investigate in what ways and to what extent the use of a computer-based probability curriculum may help students understand and apply probability concepts.

What your child will do in the study: As a participant in this study, your child will be asked to participate in activities using the *Probability Explorer* computer software package. These activities will take place during a free mathematics course held at the University of Virginia' s Curry Center for Technology and Teacher Education. Your child will also be interviewed twice at the Curry Center. During these interviews your child will be asked to complete a short series of probability tasks. Both the in-class activities and the interviews will be audio and videotaped (see attached consent forms).

Time required: Each interview will last approximately 45 minutes. The mathematics course will begin with three days of two hour classes. The remainder of the course will be held during the months of August and September for approximately one hour each week. Participants will also be asked to complete a weekly homework assignment for approximately 30 minutes.

Interviews:	2 @ .75 hours	1.5 hours
Course:	3 @ 2 hours	6 hours
	6 @ 1 hour	6 hours
Homework	6 @ .5 hours	3 hours
Total Time Commitment		**16.5 hours**

Risks: There are no risks associated with participation in this study.

Benefits: There are no direct benefits to your child for participating in this study.

Confidentiality: All information collected in this study will be confidential. Only aliases will be used in the final write-up. For purposes of data analysis, only myself, my peer debriefer, and members of my dissertation committee will view or hear the recordings. All video and audio tapes will remain in my possession at my home: 1052 Beaver Hill Drive, Charlottesville, VA 22901

Voluntary participation: Your child's participation in this study is completely voluntary.

Right to withdraw from this study: Your child will have the right to withdraw from this study at any time without penalty.

How to withdraw from the study: If your child wants to withdraw from this study at any time, please inform me, Hollylynne Drier of your child's decision. Your child may withdraw at any point during the course of the study.

Payment: Your child will receive no payment for participating in this study. Those interested will receive a complimentary copy of the *Probability Explorer* software. In addition, when completed you may request a summary of the findings.

Who to contact if you have questions about this study: If you have any questions feel free to contact me at any of the following:

Hollylynne Drier
1052 Beaver Hill Drive
Charlottesville, VA 22901
(804)-924-3399 (w)
(804)-823-4154 (h)
hollyd@virginia.edu

or my advisor:

Prof. Joe Garofalo
Ruffner Hall
University of Virginia
Charlottesville, VA 22903
(804) 924-0845 (w)

Who to contact if you have questions about your child's rights in this study: Dr. Luke Kelly, Chairman, Institutional Review Board for the Behavioral Sciences, Washington Hall, East Range, University of Virginia, Charlottesville, VA, 22903. Telephone: (804) 924-3606.

Agreement to participate: I agree to allow my child to participate in the research study described above.

Signature of parent ________________________________ Date ___________

Agreement to allow data to be used in publications, presentations, and conferences: I give informed consent to Hollylynne Drier to use any data my child provides during interviews and classroom activities for the purposes of publications, presentations, and conferences. All data will remain anonymous.

Signature of parent ________________________________ Date ___________

You will receive a copy of this form for your records.

APPENDIX B

STUDENT INFORMED CONSENT AGREEMENT

University of Virginia

CHILDREN'S PROBABILISTIC REASONING WITH A COMPUTER MICROWORLD

Please read this consent agreement carefully before you decide to participate in this study.

Purpose of the research study: The purpose of this study is to investigate in what ways and to what extent the use of a computer-based probability curriculum may help students understand and apply probability concepts.

What you will do in the study: As a participant in this study, you will be asked to participate in activities using the *Probability Explorer* computer software package. These activities will take place during a free two-week mathematics course held at the University of Virginia' s Curry Center for Technology and Teacher Education. You will also be interviewed twice at the Curry Center. During these interviews you will be asked to complete a short series of probability tasks. Both the in-class activities and the interviews will be either audio or videotaped (see attached audio and video consent forms).

Time required: Each interview will last approximately 45 minutes. The mathematics course will begin with three days of two hour classes. The remainder of the course will be held during the months of August and September for approximately one hour each week. You will also be asked to complete a weekly homework assignment for approximately 30 minutes.

Interviews:	2 @ .75 hours	1.5 hours
Course:	3 @ 2 hours	6 hours
	6 @ 1 hour	6 hours
Homework	6 @ .5 hours	3 hours
Total Time commitment		**16.5 hours**

Risks: There are no risks associated with participation in this study.

Benefits: There are no direct benefits to you for participating in this study.

Payment: You will receive no payment for participating in this study. Those interested will receive a complimentary copy of the *Probability Explorer* software.

Confidentiality: All information you provide in this study will be confidential. Your name will not be used in the final write-up. My faculty advisor may listen to or watch some portions of the tapes but your name will not be known to him. All video and audio tapes will remain in my possession at my home: 1052 Beaver Hill Drive, Charlottesville, VA 22901.

Voluntary participation: Your participation in this study is completely voluntary.

Right to withdraw from this study: You have the right to withdraw from this study at any time without penalty.

How to withdraw from the study: If you want to withdraw from this study at any time, please tell me, Hollylynne Drier, or have your parents contact me and let me know. You may withdraw at any point during the course of the study. My address is given below.

Who to contact if you have questions about this study: If you have any questions feel free to contact me at any of the following:

Hollylynne Drier
1052 Beaver Hill Drive
Charlottesville, VA 22901
(804)-924-3399 (w)
(804)-823-4154 (h)
hollyd@virginia.edu

or my advisor:

Prof. Joe Garofalo
Ruffner Hall
University of Virginia
Charlottesville, VA 22903
(804) 924-0845 (w)

Who to contact if you have questions about your rights in this study: Dr. Luke Kelly, Chairman, Institutional Review Board for the Behavioral Sciences, Washington Hall, East Range, University of Virginia, Charlottesville, VA, 22903.
Telephone: (804) 924-3606.

Agreement to participate: I agree to participate in the research study described above.

Signature of student______________________________ Date __________

Agreement to allow data to be used in publications, presentations, and conferences: I give my permission to Hollylynne Drier to use any data I provide during interviews and classroom activities for the purposes of publications, presentations, and conferences. All data will remain anonymous.

Signature of student______________________________ Date __________

You will receive a copy of this form for your records.

APPENDIX C

AUDIO RELEASE FORM

During this research study, all interviews and classroom sessions will be audiotaped. These audiotapes will serve as part of the data record of this study. They will be reviewed and analyzed by me, Hollylynne Drier, and possibly my advisor, Prof. Joe Garofalo. Portions of these audiotapes may be shown to other members of the research committee or to members of the education community at large. However, you will remain anonymous.

___ I give permission for my interviews and participation in classroom activities to be audiotaped and used as described above.

___ I do NOT give my permission for my interviews and participation in classroom activities to be audiotaped.

Signature of student: ___________________________ Date:**_________**

___ I give permission for my child's interviews and participation in classroom activities to be audiotaped and used as described above.

___ I do NOT give my permission for my child's interviews and participation in classroom activities to audiotaped.

Signature of parent: ___________________________ Date:**_________**

You will receive a copy of this form for your records.

APPENDIX D

VIDEO RELEASE FORM

During this research study, all interviews and classroom sessions will be videotaped. These videotapes will serve as part of the data record of this study. They will be reviewed and analyzed by me, Hollylynne Drier, and possibly my advisor, Prof. Joe Garofalo. Portions of these videotapes may be shown to other members of the research committee or to members of the education community at large. However, you will remain anonymous.

___ I give permission for my interviews and participation in classroom activities to be videotaped and used as described above.

___ I do NOT give my permission for my interviews and participation in classroom activities to be videotaped.

Signature of student: ___________________________ Date:_________

___ I give permission for my child's interviews and participation in classroom activities to be videotaped and used as described above.

___ I do NOT give my permission for my child's interviews and participation in classroom activities to be videotaped.

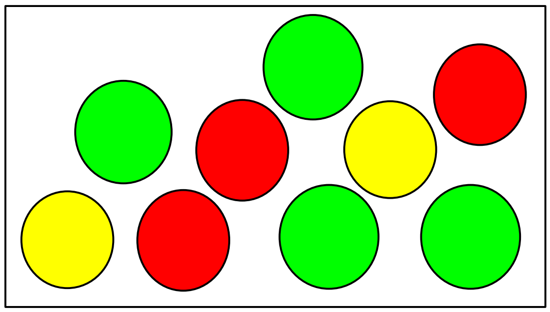
Signature of parent: ___________________________ Date:_________

You will receive a copy of this form for your records.

APPENDIX E

PRE-INTERVIEW TASKS

1. *Marbles from a bag (with and without replacement)*



4 green, 3 red, and 2 yellow marbles
Actual marbles will be used in a see-through container.

a) If you close your eyes and pick a marble out of this bag, what colors could your marble be? Why?

b) Close your eyes and draw a marble. What color is it? Now, put the marble back. If you were to close your eyes and pick again, what colors could your marble be? Why?

c) If you draw a marble without looking, which color do you have the *least* chance of picking? Why? Which color do you have the *best* chance of picking? Why?

d) I am going to draw a green marble out of the bag without putting it back. If you draw another marble without looking, which color will you have the *best* chance of picking? Why?

e) I am going to draw another green marble out of the bag without putting it back. If you draw another marble without looking, which color will you have the *least* chance of picking? Why?

f) Has the chance of getting any of the colors changed since we began picking marbles? Which colors? Why? Can you use numbers to explain your reasoning?

2. Flipping a penny

a) If I flip a penny, what are the possible outcomes?

b) How would you describe the chances of the penny landing on heads? Tails?

c) If you were to flip a penny 10 times, how many heads and how many tails would you predict?

3. *Flipping a penny and nickel*

a) A penny and a nickel are tossed onto a table. What are all the possible ways (in terms of heads and tails) for the coins to land? Draw a diagram, picture or chart that shows all the possibilities for the two coin flips.

b) Are any of the combinations more likely than others? Why or why not?

4. *Flipping a penny nickel, and a quarter*

a) A penny, nickel and a quarter are tossed onto a table. What are all the possible ways (in terms of heads and tails) for the coins to land? Draw a diagram, picture or chart that shows all the possibilities for the three coin flips.

b) Are any of the combinations more likely than the others? Why or why not?

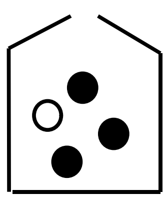
5. Marbles in a Bucket revisited

a) Bring back bucket with four greens, three red, and two yellow cubes. Describe the chances of picking each of the colors.

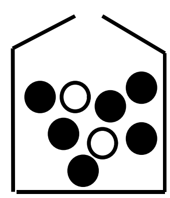
b) Take out two green cubes. Has the chance of getting any of the colors changed? Which colors? Why? Can you use numbers to explain your reasoning?

6. *Which bag is best?*

a) In bag #1, do you have a better chance of picking a white or black marble, or are the chances equal? Why? Same question for bag #2.

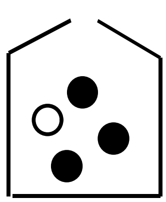


Bag 1

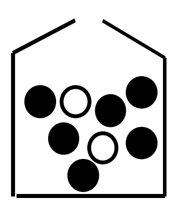


Bag 2

b) Let's say that Bag 3 and Bag 4 were shaken up:



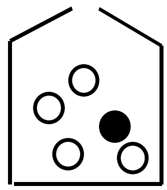
Bag 3



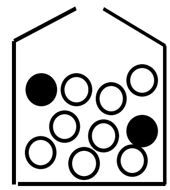
Bag 4

If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

c) These are new bags. Let’s say that Bag 5 and Bag 6 were shaken up:



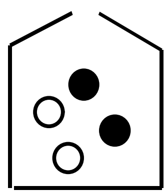
Bag 5



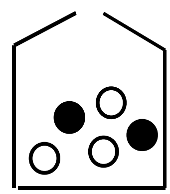
Bag 6

If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

d) These are new bags. Let’s say that Bag 7 and Bag 8were shaken up:



Bag 7



Bag 8

If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

7. Most likely series of coins question

a) If a coin is flipped six times, could you tell me what the six results might be, in order? What are some other possible results? Are any of these results more likely than the others? Why?

b) Of the following sequences, are any more likely than the others? Why?

a. H H H H T T

b. T H H T H T

c. T H T T T H

d. H T H T H T

c) Of the following sequences, are any more likely than the others? Why?

a. H H H T T T

b. H H H H H H

c. T H T H T H

d. H T H T H T

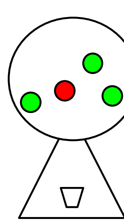
d) A sequence of six coin flips resulted in THTTTT. If I flip a coin again, which outcome is more likely to occur next:

a. Heads

b. Tails

c. Both Heads and Tails are equally likely

8. *Gumball machine*

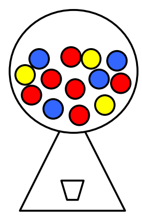


a) If you close your eyes and pick one gumball out of this machine, what color are you most likely to get? Why? What is the chance of getting this color?

b) Suppose you got a green gumball (pick out a green gumball). Now, if you draw again with your eyes closed, what color are you most likely to get? Why? What is the chance of getting this color?

c) Now suppose you pick out a red gumball (pick out a red gumball). If you draw again with your eyes closed, what color are you most likely to get? Why? What is the chance of getting this color?

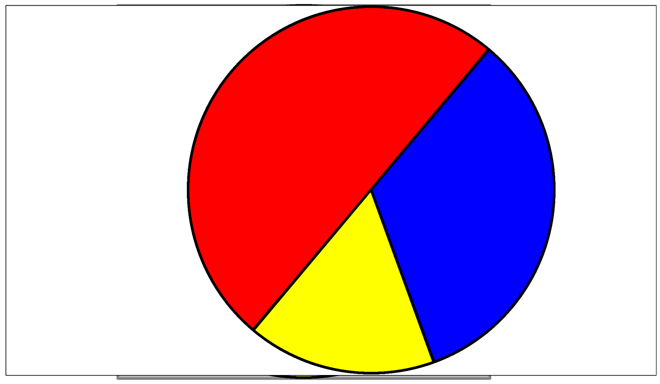
9. Predict gumballs



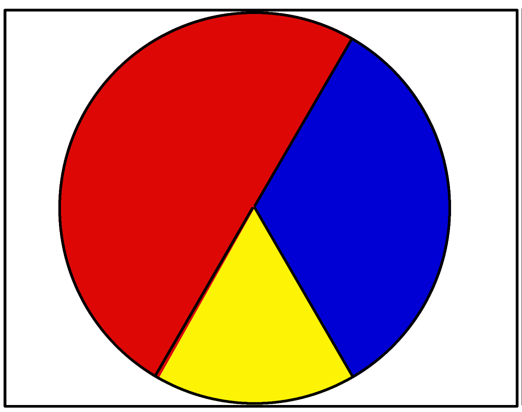
The gum ball machine has 100 gum balls; 20 are yellow, 30 are blue, and 50 are red. The gum balls are well mixed inside the machine. Jenny gets 10 gum balls from this machine. What is your best prediction of the number that will be red?

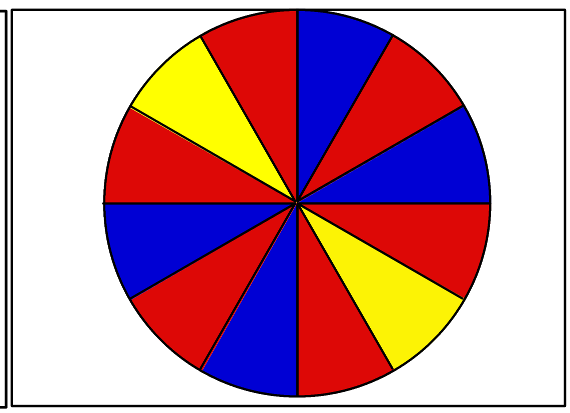
10. *Probability with Spinners (continuous areas)*

a) This spinner is used to play the penny game. You and a friend each choose a different color and take turns spinning. If the pointer lands on your color, you get a penny. If it does not land on your color, you lose a penny. Which color would you choose?



b) You are going to play the penny game again. Which spinner would be best for you? Why? Can you use numbers to tell me about the chances of getting a red on each spinner?





APPENDIX F

BRIEF OUTLINE OF INSTRUCTIONAL TASKS

Teaching Session	**Tasks Posed by Teacher/Researcher (Many tasks were extended and adapted individually by students during the sessions. The tasks listed reflect the *original tasks*)**
1	• Flipping a coin 10 times with and without technology and examining the order of results • Flipping a coin 20 times and analyzing the results in multiple representations (table, stacking columns, pie and bar graph) • Flipping a coin a large number of times to observe trends in variability
2	• Simulating rolling a standard die in the microworld both a small and large number of trials. • Comparing variability in results from a die to a coin toss • Free play with students designing their own experiment with six possible outcomes, predicting results with small and large number of trials, experimenting, and comparing results with predictions.
3	• Simulating picking out a marble with replacement from a 2B2W bag of marbles both a small and large number of trials. • Simulating picking out a marble with replacement from a 5B5W bag of marbles both a small and large number of trials. • Comparing results and theoretical probability from the 2B2W and 5B2W bags • Simulating picking out a marble with replacement from a 3B1W bag

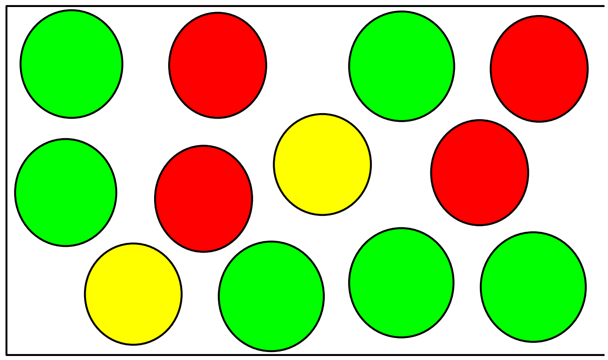
	of marbles both a small and large number of trials. • Comparing the "evening out" process (EOP) with equiprobable results to that with the 3B1W bag of marbles (unequiprobable).
4	• Playing a coin toss game with weights secretly changed to 5/6 heads and 1/6 tails to see if children are concerned with results that do not match their expected "even" distribution and if the can explain the seemingly unequal theoretical probabilities.. • Once the weight tool is discovered, challenge students to use a variety of different weights that would make the coin toss game more "even." • Students have a lot of free choice to design experiments and use the weight tool to model, predict, and experiment with a variety of different weights.
5	• With a 3-outcome experiment, the outcomes are secretly weighted in a 2-2-1 relationship. The children are challenged to experiment and use whatever tools they can to gather enough evidence to convince them that they can estimate the secret weights and justify their prediction based on experimental evidence. • Given a pie graph of ½red, 1/3 blue, and 1/6 yellow, students have to design an experiment, assign weights and run a simulation that they believe could result in a pie graph similar to the one given.
6	• With a two-outcome experiment, the children have to design weights so that one outcome is "twice as likely" to occur as the other outcome. They are challenged to find several ways to model the "twice as

	likely" relationship and simulate the experiment with both small and large number of trials to observe and discuss the EOP with this situation. • Students design an experiment with 3 possible outcomes that occur two at a time. They use the Make It tool to construct all possible outcomes (sample space). Discussions about whether order matters are expected to arise. • Students run experiments with small and large number of trials with the 3-outcome, 2-stage experiment and analyze the data both as "ordered" and "not ordered" and discuss differences between the expected results and theoretical probabilities of each situation.
7	• Students design an experiment with 4 possible outcomes and choose their own weights. They are asked to predict the results for 10, 100, and 1000 trials. They then run several sets of simulations with each number of trials (10, 100, and 1000) and compare their predictions with the actual results. • Students design a 3B1W bag of marbles and do many simulations with replacement. They then design a 6B2W bag of marbles, run simulations, and are asked to compare the results as well as the theoretical probabilities.

APPENDIX G

POST-INTERVIEW TASKS

1. *Marbles from a bag (with and without replacement)*



6 green, 4 red, and 2 yellow marbles
Actual marbles will be used in a see-through container.

a) If you close your eyes and pick one of these marbles out of this bag, what possible colors could your marble be? Why?

b) Close your eyes and draw a marble. What color is it? Now, put the marble back. If you were to close your eyes and pick again, what colors could your marble be? Why?

c) Close your eyes and draw a marble. What color is it? If you were to close your eyes and pick again, what colors could your marble be? Why?

(Put all marbles back in bag)

d) If you draw a marble without looking, which color do you have the *least* chance of picking? Why? Which color do you have the *best* chance of picking? Why? Can you use numbers to describe the chances of picking each of the colors?

e) I am going to take a green marble out of the bag without putting it back. If you pick another marble without looking, which color will you have the *best* chance of picking? Why?

f) I am going to take another green marble out of the bag without putting it back. If you draw another marble without looking, which color will you have the *least* chance of picking? Why?

2. Coin Tosses

a) Of the following sequences, are any more likely than the others? Why?

a. H H H H T T

b. T H H T H T

c. T H T T T H

d. H T H T H T

b) Of the following sequences, are any more likely than the others? Why?

a. H H H T T T

b. H H H H H H

c. T H T H T H

d. H T H T H T

c) A sequence of six coin flips resulted in HTHHHH. If I flip a coin again, am I more likely to get a head, a tail, or are they equally likely?

3. Is this coin fair?

a) I tossed a coin 10 times and got eight tails and two heads. Is this coin fair?

b) I tossed a coin 100 times and got 41 heads and 59 tails. Is this a fair coin?

c) I tossed a coin 500 times and got 175 tails and 325 heads. Is this a fair coin?

4. Marbles in a Bucket revisited

a) Bring back bucket with six greens, four red, and two yellow cubes. Describe the chances of picking each of the colors.

b) Take out two green cubes. Has the chance of picking any of the colors changed? Which colors? Why? Can you use numbers to explain your reasoning?

5. *Which bag is best?*

a) In bag A, do you have a better chance of picking a white or black marble, or are the chances equal? Why? Same question for bag B.

If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

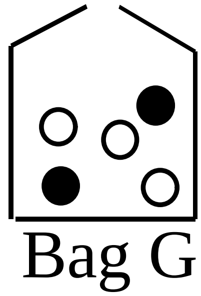
b) Let's say that Bag C and Bag D were shaken up:

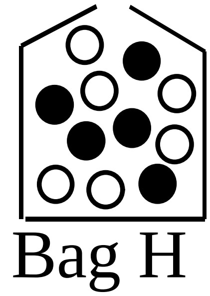
If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

c) These are new bags. Let's say that Bag E and Bag F were shaken up:

If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

d) These are new bags. Let's say that Bag G and Bag H were shaken up:





If you had to close your eyes and pick a marble from one bag or the other, and you wanted to pick a **black** marble, which bag would you choose from, or does it matter? Why?

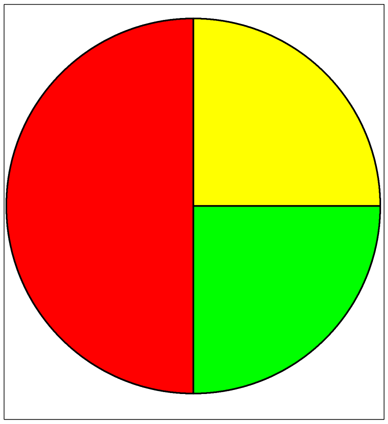
6. *Family of Three kids.*

The Lindburg family has three children. The kids ages are 9, 5, and 3.

a) Could you make a list of all possible arrangements of boys and girls by their ages?
b) Are any of these arrangements more likely to occur than the others? Explain.
c) What is the chance that the arrangement is BGB?
d) What is the chance that the family has two boys and one girl?
e) What is the chance that the family has all boys or all girls?
f) What is the chance the family has two girls and one boy?

7. *Given results design bag of marbles*

A. I designed a bag of marbles and ran an experiment (with replacement). This is a pie graph of my results.

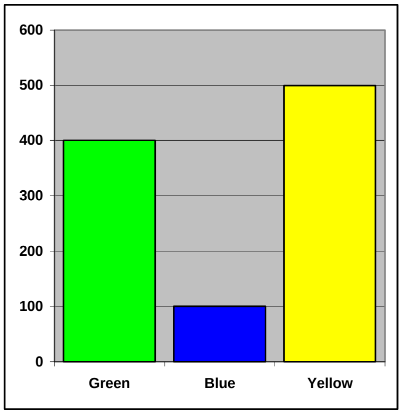


1. How many times did I pick marbles from the bag?

2. Based on these results, what can you tell me about the bag of marbles?

3. Design a bag of marbles that you think would give these results?

B. I designed a different bag of marbles and ran another experiment (with replacement). This is a bar graph of my results.

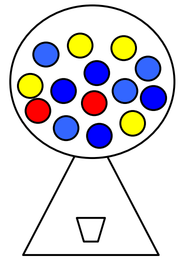


1. How many times did I pick marbles from the bag?

2. Based on these results, what can you tell me about the bag of marbles?

3. Design a bag of marbles that you think would give these results?

8. Predict gumballs



The gum ball machine has **100** gum balls; **30** are yellow, **60** are blue, and **10** are red. The gum balls are well mixed inside the machine. Jeff gets 10 gum balls from this machine.

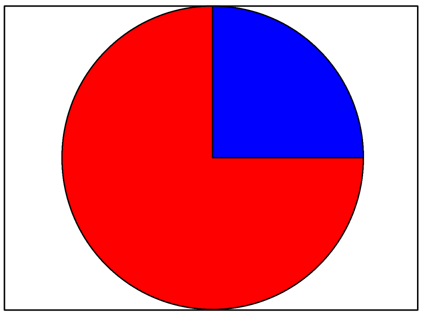
What is your best prediction of the number that will be red? Why?

What is your best prediction of the number that will be yellow? Why?

What is your best prediction of the number that will be blue? Why?

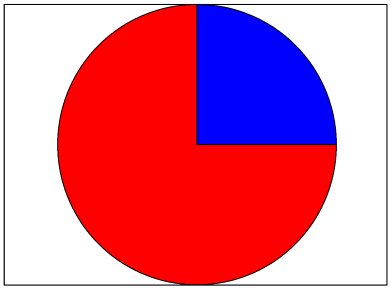
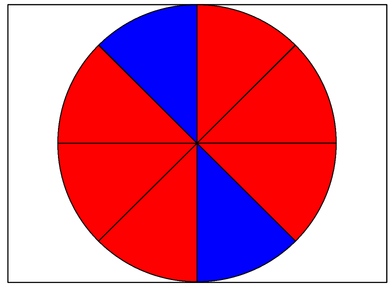
9. Probability with Spinners (continuous areas)

a) This spinner is used to play the penny game. We each choose a different color and take turns spinning. If the pointer lands on your color, you get a penny. If it does not land on your color, you lose a penny. Which color would you choose?



b) Is this a fair game? Explain.

c) You are going to play the penny game again. Which spinner (below) would be best for you? Why? Can you use numbers to tell me about the chances of getting a red on each spinner?

d) Can you think of a way to change this game so that it is fair? Explain.