2010

Seventh-grade curriculum in probability (a guide for teachers)

David Nganga Njenga
Louisiana State University and Agricultural and Mechanical College, dnjeng1@lsu.edu

SEVENTH-GRADE CURRICULUM IN PROBABILITY
(A GUIDE FOR TEACHERS)

A Thesis
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Natural Sciences

in

The Interdepartmental Program in Natural Sciences

By
David Nganga Njenga
B.Ed., Moi University - Kenya, 1995
December 2010

ACKNOWLEDGEMENTS

I give thanks to all who have guided and supported me in my journey to this achievement. I especially thank my supervisor and adviser Dr. James Madden for the sacrifice and invaluable guidance and support throughout the process of writing this thesis. Thanks also go to my committee members, Dr. Frank Neubrander and Dr. Padmanabhani Sundar. Their comments have gone a long way in improving the quality of this thesis.

Great thanks go to Sharon Besson and to the Cain Center staff for all the assistance they offered us while we carried out our studies at Prescott Hall. I give thanks to my big brother Jonah Njenga for not only inspiring me to acquire a Masters degree, but by sending me information about this program.

To the mathematics cohort, being with you for the last three years has been a great experience. You have all encouraged me throughout the course. I thank my wife, Fidelis and our children, Rodney, Cephas and Joy for supporting me and understanding whenever I have been long away from them. Lastly but not least, I thank my sister Jane Njenga for giving me the encouragement I needed while writing this thesis. Above all I give thanks to God for enabling me to attain this success and for giving me the strength every time I felt tired and weak while researching and writing this thesis.

This thesis is a special dedication to my loving parents, Geoffrey Njenga Nganga and Rahab Wanjiru Njenga, who tirelessly reminded me that education, is the greatest inheritance I can receive from them! Thank you so much Dad and Mom, and may the Lord bless you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A review of the literature indicates that teachers try to make connections between experimental and theoretical probabilities while teaching their students, but these connections are often not very clear. This greatly increases misconceptions students have about probability. This thesis presents a treatment of teaching probability that is geared for seventh grade and intended to minimize the misconceptions that both teachers and students may have. We present a concise mathematical exposition of finite probability models as well as collection of examples and activities, so as to help teachers and students organize their thinking and minimize misconceptions. The various examples and activities suggested in this thesis hopefully will increase the quality of instruction in probability and ultimately motivate the students in this important field of mathematics.

CHAPTER 1: INTRODUCTION

Probability influences the way we interpret the world around us and shapes many of our decisions at critical points in our lives. People regularly make informal judgments about chance in the context of health, finance, weather, sports, games, and other events. Will it rain today? Should I buy a lottery ticket? How likely is it that I will be stuck in traffic jam? All these are examples of ordinary situations in which people make predictions about uncertain events using probabilistic reasoning. Probability is used to make decisions that affect how ordinary people live their lives.

1.1 Why Teach Probability?

Probability is the language of uncertainty and randomness, undeniable facts of life. Probability is part of our everyday vocabulary in that we use words relating to probability easily and fluently. Examples include: probably, certainly, fairly certain, and not likely. There are many uses of probability in the media and public discourse that we have internalized and now take for granted.

It is now generally recognized that reasoning from uncertainty is an independent and fundamentally different intellectual method that must be taught alongside the other methods such as logical deduction, experimentation in science and reasoning in mathematics (Falk & Konold, 1992; Moore, 1990; Nisbett, Lehman, Fong & Cheng, 1987). Falk and Konold (1992, p. 151) argue that "Probability is a way of thinking. It should be learned for its own sake. In this century probability has become an integral component of virtually every area of thought. We expect that understanding probability will be as important in the 21st century as mastering elementary arithmetic is in the present century".

Probability should also be included in the school curriculum because it will be necessary knowledge for students as they encounter higher level mathematics. Applications of other mathematics in probability, such as ratios, fractions, percents, and decimals, increase the value of other mathematics (NCTM, 1989). To develop probabilistic thinking, it is recommended that students have numerous opportunities to make predictions and test their conjectures.

Probability has become an area of emphasis in school curricula in the past 15 – 20 years (e.g., NCTM, 1989, 2000). However, most teachers have little or no prior experience with probability in their own schooling nor in the teacher preparation programs they have attended. In 1990's, efforts at professional development for practicing teachers began while teacher-preparation programs started to include attention to probability in mathematics methods courses. There is evidence (Vacc, 1995) that many teacher education professionals doubted the usefulness of probability in the elementary grades, despite the recommendation of the NCTM (1989). Besides the importance and relevance of the subject, the nature of the subject itself is a further consideration for deciding how and when it is to be taught.

1.2 Challenges in Learning Probability

The understanding of probability as a school subject presents its own challenges. Garfield and Ahigren (1988) list the following difficulties in learning probability:

Limited experience with uncertainty. While life is conducted amid situations relating to chance and uncertainty of future events, we all have difficulty assessing risk, especially when the situations get more complex. Even a statement as simple as 'there is a 60% chance of rain tomorrow' that is regularly used in the media requires some degree of sophistication in order to articulate its meaning. Gigerenzer (1989) notes the doctor whose patients have difficulty understanding his meaning when he says there is a 30% chance of increased blood pressure with a particular medication. They find it easier when he says that if he were to prescribe the drug for 10 people, about 3would experience increased blood pressure. While there is still need to articulate what is meant by the word ‘about,’ using frequencies rather than probabilities often help people understand. This clearly shows that people use quantitative language while talking about chance, but do not take the meaning deeply analytically.

Belief in determinism and a tendency to look for causes. Adults and children have a strong belief in determinism and hence have a tendency to look for and ascribe causes. We are often reluctant or unwilling to recognize and acknowledge randomness and uncertainty. and often believe that we can manipulate a cause to get a particular effect. Young children, in particular, are reluctant to accept randomness and often think that by manipulation of the situation, a particular effect can be made to happen. They have a belief that an effect can be achieved by personal will or by a mechanism, either personal or external.

We often have difficulty in making sense of situations that do not turn out as we expect. Kahneman & Tversky (1982) write of the conflict of ‘passive expectations and conscious anticipations’ and the surprise we feel when an expectation is violated. In other words while playing a game of chance where a coin is being tossed children often think that, if you get a tail in one event after tossing a coin, then it is automatic that in the next toss, the result will be heads. This is not randomness because while tossing a coin anything is possible, including getting all heads when a coin is tossed ten times.

Lack of representations of cognitive schema. Probability is an abstract concept that is not easily measured, like distance, weight and time, or easily represented, like number, shape and pattern. There are no physical representations of the concept. When we look at a die, there is nothing about the die before it is tossed that indicates that the probability of a particular face is 1/6. There is an equal weight of all sides and an agreement about meaning when we say the probability that a particular side will come up is 1/6. Even after the die is tossed, there is no way we can measure the probability of what just happened.

Complexity. Once we move past simple events, the calculations of probabilities get complex fairly quickly. As an example, tossing 2 dice gives 36 outcomes and tossing 3 dice gives 216outcomes. The numbers get very large very quickly, so that even when *n* is 4 or 5, the sample space is too large to be listed in writing. Thus, much of the reasoning and calculation in more complex situations has to be done abstractly.

The nature of the subject matter. The subject of probability deals with uncertainty and variation. Some authors suggest that this may lead to “epistemological anxiety” and “ambiguity intolerance” (Wilensky, 1994). “Epistemological anxiety”, according to Wilensky, is “a feeling, often in the background, that one does not comprehend the meanings, purposes, sources or legitimacy of the mathematical objects one is manipulating and using (p. 172).” “Ambiguity intolerance” refers to the tendency to feel a discomfort or perceive a threat in situations of incomplete, fragmented, and insufficient knowledge.

1.3 Problem Statement

There are calls for teaching probability in a deeper and different way than it is being done currently. These calls have been made by National Council of Teachers of Mathematics (NCTM), various authors and the Common Core State Standards for Mathematics. This places demand on teachers to improve the teaching of probability but literature shows that their understanding of probability is poor.

This thesis tries to solve this by clearly presenting an elementary description of the theory of probability and developing curricular plans and materials for teachers’ use. This paper will also examine teachers’ knowledge and beliefs about probability, their ability to teach it, and lessons learned from programs in teacher education that have aimed at developing teachers' knowledge about probability. Student-thinking in probability will also be addressed.

Finally, this thesis will propose an approach of teaching probability to seventh grade students using the Common Core State Standards for Mathematics (CCSSM). Classroom activities that can be used to bring the central idea of probabilistic thinking home will also be considered in this thesis.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

In this chapter we will look at the history of probability in an effort to understand the origin of the most basic ideas. This will be followed by a review of the literature on students' probabilistic thinking, of research on teacher knowledge generally and finally of teacher beliefs and content knowledge of probability.

2.2 History of Probability

It is evident from historical texts that probability and related ideas of luck, fate, hazard, and destiny have been a part of the lives of humans from earliest times. Intuition and activity relating to chance, randomness and variation have been with us since earliest civilizations. In the paragraphs below, drawing on Bennett (1983), David (1998), and Hacking (1975, 1990), the historical beginnings of probability are sketched.

The first indication is evidence of astragali (ankle bones of hoofed animals such as deer, calf, sheep or goat) and tali (heel bones) 40,000 years ago in pre-historic sites in Mesopotamia, Egypt, the Indus Valley, Greece, and the Roman Empire. Gaming boards were found in ancient Babylonia around 2700 BC, at the palace of Knossos in Crete around 2400-2100 BC and later at Babylonian and Assyrian sites. These boards were similar to backgammon boards found in Egyptian tombs and to boards in games such as Hounds and Jackals and Snakes and Ladders with pieces or 'men' made of ivory but no accompanying dice.

Besides leisure games, a second important use of chance in antiquity was in divination or seeking divine direction. The use of an objective chance mechanism eliminated human tendencies to judgment, so that the will of the gods could be clearly manifested. Pebbles, arrows, and dice were thrown in temples to probe divine will. Besides divination by priests, people regularly threw astragali, tali and dice to make decisions in their daily lives about love, career and business dealings. However, even though they used chance mechanisms and randomizers, the ancients believed that the gods controlled the outcome.

A dispute in the game of gambling in 1654 led to the creation of a mathematical theory of probability by two famous French mathematicians, Blaise Pascal and Pierre de Fermat. Chevalier de Mere and Antoine Gomband called Pascal’s attention to a contradiction concerning popular game of chance. The game consisted of throwing a pair of dice 24 times. The problem was to decide whether or not to bet even money on the outcome of at least one “double six” in the 24 throws. A seemingly well-established gambling rule led de Mere to believe that betting on a double 6 in 24 throws would be profitable, but his calculations showed him the opposite.

This problem posed by de Mere led to the exchange of letters between Pascal and Fermat in which the fundamental principles of probability theory were formulated for the first time. The Dutch scientist Christian Huygens, a teacher of Leibniz, learned of this correspondence and in 1657 published a book on probability, entitled *De Ratiociniis in Ludo Aleae*, which contained problems associated with gambling. Because of the problems in the game of chance, probability theory soon became popular, and the subject developed rapidly during the 18^{th} century. The major contributors in this period were Jakob Bernoulli (1654 - 1705) and Abraham de Moivre (1667 - 1754).

In 1812, Pierre de Laplace (1749 - 1827) introduced a number of new ideas and mathematical techniques in his book, *Theorie Analytique des Probabilites*. Before Laplace, probability theory was solely concerned with developing a mathematical analysis of the games of chance. Laplace applied probabilistic ideas to many scientific and practical problems. The theory of errors, actuarial mathematics and statistical mechanics are examples of some of the important applications of probability theory developed in the 19th century.

In his book on probability theory, Laplace came up with ten general principles of the calculus of probabilities. The first is the definition of probability itself, which is given as the ratio of the number of favorable cases to that of all cases. He argues in the second principle that this definition supposes the various cases being equally probable. For example, suppose we throw into the air a large thin coin (fair coin). There are four equally probable outcomes, namely, HH, HT, TH, TT. If we are trying to throw heads at least one time in two throws, the first three cases are favorable to the event whose probability is sought. Thus, the probability is ¾. This gives a bet of 3:1 that heads will be thrown at least once in two throws.

In the third principle, Laplace talks of the independence of events. He says that probability of several independent events occurring at once is the product of the probabilities of each. For example, the probability of rolling an ace with a single die throw is 1/6, while that of rolling two aces at the same time in throwing two dice is 1/36. The possibly confusing part here is the manner in which these probabilities increase or diminish as the number of rolls increases. The fourth principle is that for dependent events, the probability of the compound event is the product of the probability of the first event and the probability that, this event having occurred, the second will occur. For example given three urns of which two contain only white balls and one contain only black balls, the probability of drawing a white ball is 2/3 (because only 2 urns contain white balls). But when a white ball has been drawn from one of the urns, then the probability of picking a black ball from one of the other two is ½, because only one of the two urns contains white balls. The product $2/3 \times ½ = 2/6 = 1/3$, is the probability of drawing two white balls at one time from two different urns. Here we see the influence of past events to future events.

The fifth principle is a corollary of the fourth one. The probability of a specific case of a general event, after the general event has occurred, is the first probability divided by the second. (For example, the probability of rolling a sum less than or equal to 5 on a green die and a red one (the specific event) is 10/36. But after a number less than or equal to 5 is rolled on the green die—a general event with probability 5/6—the probability of a sum less than or equal to 5 is $10/30 = 10/36 \div 5/6$.

The sixth principle gives the reason why we attribute regular events to a particular cause. Some philosophers have thought that regular events are less probable than others. For example, at the play of heads and tails (with a fair coin), the combination in which heads occurs twenty successive times is less easy in its nature—so they think—than those runs where heads and tails are mixed in an irregular manner. But this is error. In fact, each particular run is equally probable, but we do not distinguish the irregular runs, of which there are many, from one another. So, there are many ways to get an irregular run, but only one way to get all heads.

The seventh principle states that the probability of an event is a sum of terms, one for each cause, each term being the probability of the cause times the probability that, this cause existing, the event will occur. The causes are assumed mutually exclusive. For example, suppose we choose a ball from one of three urns, each with a different proportion of white balls, w_1, w_2 and w_3. If the first urn is twice as likely to be chosen as the other two, then the probability of white is $(1/2)\ w_1 + (1/4)\ w_2 + (1/4)\ w_3.$

The next three principles concern hope. The word hope expresses generally the advantage of that one who expects a certain benefit in situations which are possible. This advantage is called "mathematical hope". In the eighth principle, this advantage is the sum of the products of the probability of each outcome by the benefit attached to its occurrence. For example, in game of heads and tails suppose you receive 2 dollars if you throw heads at first throw and 5 dollars if you throw it in the second draw. In the first throw, (½ × $2.00) + (1/4 × $5.00) = $2.25. This will be your advantage.

The ninth principle is all about benefits and losses. Here, we have the advantage which results from benefits and losses by making a sum of the products of the probability of each favorable event by the benefit it procures. Laplace calls for the understanding that in order to achieve this, appreciate exactly the advantages, the losses and their respective probabilities. Finally, in the tenth principle, Laplace talks about the relative value of an infinitely small amount.

As the development of probability gained momentum from gambling games and the involvement of mathematicians in the 17^{th} and 18^{th}centuries, ideas of statistics were being developed by Galton and his studies of the inheritance of traits from generation to generation, by Fisher for his application of sampling and the design of experiments to agriculture, and in actuarial work relating to censuses, mortality rates, insurance and annuities. While there was still the notion of physical laws, order, design, and determinism in the universe, by the beginning of the 20th century the three main notions of probability that we use today, the classical probability of equally likely outcomes, the long-run frequency notion and the subjective notion of probability, had been formulated.

In addressing why it took so long to formulate a theory of probability, Ian Hacking (1975) argues that the nature of probability is essentially dual: on one hand, aleatory (concerned with chances and contingencies) and on the other, epistemological (concerned with assessing varying degrees of belief of propositions), and that a long incubation was required for addressing that duality.

Like so many other branches of mathematics, the development of probability theory has been stimulated by the variety of its applications. Mathematical statistics is one important branch of applied probability; other applications occur in such widely different fields as genetics, psychology, economics, and engineering. Many workers have contributed to the theory since Laplace's time; among the most important are Chebyshev, Markov, von Mises, and Kolmogorov.

One of the difficulties in developing a mathematical theory of probability has been to arrive at a definition of probability that is precise enough for use in mathematics, yet comprehensive enough to be applicable to a wide range of phenomena. The search for a widely acceptable definition took nearly three centuries and was marked by much controversy. The matter was finally resolved in the 20th century by treating probability theory on an axiomatic basis. In 1933 a monograph by a Russian mathematician A. Kolmogorov outlined an axiomatic approach that forms the basis for the modern theory. Since then the ideas have been refined somewhat and probability theory is now part of a more general discipline known as measure theory."

2.3 Students' Probabilistic Thinking

The seminal work in the area of how the concept of chance develops in children was carried out by Piaget and Inhelder in the 1950s and published in English in 1975 as "The Origin of the Idea of Chance in Children" as part of a number of studies on the cognitive development in children of concepts such as number, space and proportional reasoning. Piaget and Inhelder engaged children in tasks involving random mixtures, distributions (centered and uniform), random drawings, and combinations and permutations. From their findings they proposed three stages in the development of the idea of chance:

a. **Preoperational (four to seven or eight years).** Here the child is unable to differentiate the possible from the necessary (that which has a cause). There is no reference to operations and no idea of chance or deduction but only an intuition of real or imaginary regularity. When faced with a random mixture of objects, the child is intuitively sure that the objects will return to a regular ordering.

b. **Concrete operational (seven or eight years to eleven or twelve years).** At this stage the operations of logic and arithmetic begin to appear and there is some understanding of the difference between necessary and possible events but the child has no systematic approach to generating a list of possibilities due to a lack of combinatorial skills or mathematical maturity.

c. **Formal operational (eleven or twelve years and older).** It is only in this stage that a judgment of probability becomes organized and the evolution of the idea of chance is achieved. The child begins to understand the ideas of experimental and theoretical probability and to have some facility with listing all possible outcomes and other combinatorial analyses.

Piaget and Inhelder concluded that chance is gradually discovered and it is only by constantly making reference to the operations of logic and arithmetic in a parallel development with these operations that children eventually come to an understanding of chance and probability.

These findings were challenged by Fischbein and other researchers in the post-Piagetian period. Fischbein (1975) found evidence of probabilistic thinking after instruction as early as the third grade while others criticized the narrow tasks and settings of the Piagetian experiments and demonstrated that elements of probabilistic thinking were evident in children even as young as preschoolers (Davies, 1965; Goldberg, 1966; Yost et al., 1962).

In a more recent study of children in grades 3 and 4, Jones et al. (1997) developed four levels of probabilistic reasoning (subjective, transitional, informal quantitative, and numerical) for each of four probability ideas of sample space, experimental probability, theoretical probability (including probability comparisons), and independence. Jones and colleagues do not suggest that children demonstrate an ordered progression through the levels but that the framework presents broad guidelines to inform instruction in probability.

We see great variation in emphasis by authors within the various studies of children's probabilistic thinking in the post-Piagetian and contemporary periods, for example in probability ideas, misconceptions, combinations of probability ideas and misconceptions, age groups, tasks, and environments. To narrow down the conceptual field in order to make it manageable for this thesis, I will concentrate on uncertainty and randomness, probability as a ratio, and theoretical and experimental probability.

Uncertainty and randomness. The age at which children understand the idea of uncertainty varies by study but is generally agreed to be between 4 and 10 years, with the inconsistencies in the results of the studies due to the differences in the tasks used (Acredolo, O'Connor, Banks & Horobin, 1989; Byrnes & Beilin, 1991; Falk &Wilkening, 1988; Horvath & Lehrer, 1998; Kuzmak & Gelman, 1986). It has been found that children simultaneously hold ideas of chance and a deterministic outlook, with the deterministic stance often persisting with age (Jones, Langrall, Thornton & Mogill, 1997; Konold, 1991; Metz, 1998a, 1998b; Shaughnessy, 1992). Other studies show that children in the elementary grades have superstitious beliefs, think they can will or influence chance outcomes, and often ascribe outcomes to external, uncontrollable forces (Falk, 1983; Fischbein, Nello & Marino 1991; Green, 1983; Piaget & Inhelder, 1951/1975).

In theorizing about children's understanding of probability, Fischbein (1975, 1987) proposed the concept of "intuitions" which he described as cognitive beliefs that come spontaneously and are self-evident to the believer. He asserted that all people have intuitions which are immediate, holistic, adaptable, and obvious to the believer, including intuitions about number, pattern, and probability. A “primary intuition” is one that a person has from his or her own experience without the benefit of instructional intervention. A “secondary intuition “is a restructured cognitive belief that a person has acquired after instruction and experience in a particular cultural community. Fischbein asserted that replacing a primary intuition by a secondary one is not a gradual process, but one that takes place as a whole, or all at once in a flash of discovery or insight.

Fischbein, Pampu & Minzat (1967) found that pre-school children performed well in probability tasks but older children were less successful. They argued that the pre-schoolers' thinking was intuitive and untutored and that the teaching process in science and mathematics ‘orients the child toward a deterministic interpretation of phenomena, in the sense of looking for and explaining in terms of clear-cut, certain, and univocal relations’(p. 169). Greer (2001) terms this a ‘cultural bias of deterministic thinking’ (p. 20).

According to the studies discussed, it is clear that by age 10—with perhaps some interference—kids have some basic probability ideas. Fischbein says ideas develop as primary intuitions which may be replaced by secondary ones. The purpose of education is to make this important transition. The subject of probability is no exception; the primary intuitions children have can be used by teachers while teaching probability. To be able to tap into this knowledge, teachers need to have a well-planned process so as not to antagonize the development of these useful secondary intuitions.

Probability as a ratio. The research on probability as ratio is contradictory. Falk, Falk & Levin (1980) and Acredolo et al. (1989) claim that children as young as age six have some understanding of probability as the ratio of the number of favorable cases to the total number of cases, using problems that involved changes in the numbers of favorable and total cases. Green (1983, 1988) found that 11-16 year-olds in the UK had a fragile understanding of probability.

Green based his results on a paper-and-pencil multiple-choice survey instrument, while Falk et al. used binary choice tasks with two colors, one of them the payoff color. The child had to choose which of the two colors to draw (from an urn) or land on (spinners and roulette wheels) for a reward. Acredolo et al. used a sliding scale that children used in order to indicate the likelihood of an event. Because of the variety of tasks in these studies, is difficult to gauge the level of children's understanding of the concept of probability as a ratio.

Besides spontaneous responses, there have been studies on the effect of instruction on probability ideas. Fischbein, Pampu & Manzat (1970), in a study of three groups of children aged 5, 9 and 12 years, found that after a short program of instruction the 9 and 12 year olds were able to correctly use comparison of quantitative ratios in a binary-choice task of choosing which bag offered the best chance of getting a particular color.

Theoretical and experimental probability. Metz (1998a, 1998b), in a study involving three groups (undergraduates, children in grade 3 and children in kindergarten), found that both children and adults have difficulties in probability determinations related to knowing and assessing relative magnitude, comparing numbers, and understanding part-part and part-whole relations. Metz (1998a) also comments that the unschooled primary-grade children's conceptions "fail to reflect the mathematician's framing of both randomness and probability over large numbers of events. Thus instruction will presumably need to address the subtle idea of phenomena whose individual outcomes may be undetermined but that reveal patterns of outcomes over the long haul" (p. 169).

Research indicates that perhaps no other branch of mathematical sciences is as important for all students as probability (Schultz, 1989; Shaughnessy, 1992). Among many attempts to encourage an increased emphasis on probability, the NCTM Curriculum and Evaluation Standards for School Mathematics recommends including probability in the core curriculum at all levels (NCTM, 1989). There is a wide gap between these ideals and reality, however. Probability is still a weak spot in school mathematics curriculum and instruction.

According to the standards, the study of probability in the middle school should not focus on developing formulas or computing the likelihood of events pictured in texts. Students should actively explore situations by experimenting and simulating probability models (Konold, 1991). Therefore, devising and carrying out experiments or simulations to determine probability is important for all students.

2.4 Research on Teacher Knowledge

Regardless of the subject being taught, the knowledge base of the teacher plays a role in student learning (Shulman, 1987). In order to comprehend issues related to the teaching and learning of probability, it is important to have an understanding of both teacher's content knowledge of probability and teacher's knowledge of how to teach probabilistic concepts (Stohl, 2005). Therefore, it is necessary to first explore teacher knowledge in general in order to have a broader lens for interpreting a teacher's knowledge specifically related to probability. I will first explore research regarding teacher knowledge and followed by review of literature relating to teacher's knowledge of probability.

According to Shulman, teaching starts with the teacher's understanding of knowledge and the teacher's conceptions of how this knowledge is to be taught to students' (Shulman, 1987). In the classroom, the teacher is ultimately responsible for designing and providing opportunities for students to acquire new knowledge. This task is not easy. A teacher makes a multitude of decisions and engages in mathematical reasoning prior to and during instruction in order to create learning experiences for students; the nature of these decisions have an impact on what can be learned by students (Ball, 2000; Fennema & Franke, 1992; Wilson, Shulman, & Richert, 1987). For example, a teacher must decide which parts of the curriculum to highlight in his or her classroom and which types of tools and activities will be presented to students. During instruction, the teacher must also make many decisions that affect the outcome of learning in his or her classroom. For example, the teacher must determine appropriate questioning techniques and suitable interactions with his or her students. All pre-active and interactive decisions made by teachers are dependent on the knowledge base of the teacher.

In the 1980's, Shulman and his colleagues studied the knowledge base of teachers. Shuman (1986, 1987) specifically studied the transition of student teachers to first year teachers in order to identify how new teachers learn to teach. This research focused on the development of secondary teachers in California in the subject areas of English, biology, mathematics, and social sciences. Shulman and his colleagues also studied veteran teachers to compare their practices to the novice teachers. Shulman identified that there are three kinds of content knowledge that teachers possess: subject matter content knowledge, pedagogical content knowledge, and curricular knowledge. Shulman (1986) emphasizes the importance of subject matter content knowledge, stating that teachers must be capable of defining for students the accepted truths in a domain. They must also be able to explain why a particular proposition is deemed warranted, why it is worth knowing, and how it relates to other propositions, both within the discipline and without, both in theory and in practice (p.9).

Pedagogical content knowledge is the knowledge that allows teachers to communicate conceptions of their discipline to their students. Shulman (1986) defines pedagogical content knowledge as useful forms of representation of those ideas, powerful analogies, illustrations, examples, explanations, and demonstrations – in a word, the ways of representing and formulating the subject that make it comprehensible to others (p. 4). According to Shulman, the teacher must have an understanding of the different types of representations, founded in research and in the practice of teaching, that are needed to help students develop an understanding of the content. Teachers must also be able to identify what makes the learning of topics easy or difficult for students.

The third category of content knowledge that teachers possess is curricular knowledge. This type of knowledge guides teachers in their making decisions concerning instructional materials such as textbooks, technology and manipulatives. This type of knowledge is not limited to the selection of instructional materials and tools. It also includes an understanding of how to sequence these materials within a curriculum.

Wilson et al. (1987) argue that although it is necessary for teachers to have an understanding of the subject matter that they teach, subject matter content knowledge alone is insufficient to foster understanding in students. They claim that teachers must also possess pedagogical content knowledge, the knowledge that allows teachers to communicate conceptions of their discipline to their students. Although both types of knowledge are equally important to teachers, they are often treated as separate domains.

Also agreeing with Shulman, Ball (2000) confirms the dual importance of the understanding of theoretical knowledge about subject matter and the development of pedagogical methods. Ball argues that understanding content knowledge is crucial to creating worthwhile learning opportunities for students in mathematics. Ball states that "No matter how committed one is to caring for students, to taking students' ideas seriously, to helping students develop robust understandings, none of these tasks of teaching is possible without making use in context of mathematical understanding and insight" (p. 243).

Ball also claims that knowing how to use this knowledge of mathematics coupled with the understanding of mathematics itself is the heart of teaching all students. She also asserts that issues regarding teacher understanding of subject matter and their ability to use this understanding in teaching should be viewed from the perspective of the practice of teaching.

2.5 Teacher Beliefs and Content Knowledge of Probability

In a 2001 study, Watson gained insight into teachers' subject-matter content knowledge and pedagogical content knowledge by developing an instrument to profile teacher achievement in probability and statistics. This instrument was designed to assess Australian teachers in terms of their ability to implement a new mathematics curriculum. Watson's study included 43 Australian public and private school mathematics teachers of grades 1 – 6 and grades 7 – 12. She found a need to improve teachers' content knowledge of probability as well as their pedagogical content knowledge generally.

Watson found several critical issues relating to the teaching of probability. She indicated that there was no evidence that teachers were planning coherently with respect to chance and data at the primary level. Primary teachers needed experience with probability in equally likely outcomes, basic probability outcomes, odds, and sampling. Teachers were generally more familiar with the concept of average than with the concept of sample, meaning they had handicaps in dealing with the probability curriculum.

When questioned about topics that they would include in preparing a unit on chance and data, primary teachers chose topics such as surveys, graphs, chance in general, and probability contexts that involved rolling dice. In contrast, when secondary teachers were questioned, the most common topic mentioned was probability. More than half of the teachers (both primary and secondary) reported that they liked teaching the probability topics they listed. Almost 30 % of the teachers indicated that the most enjoyable topics to teach were graphing, normal distributions, surveys and data collection, analyzing and interpreting data, and different facets of probability. These teachers indicated that these topics were enjoyable for their students as well. Watson (2001) stated that: “these data suggest that many teachers are either comfortable describing the teaching of topics they enjoy or come to enjoy the topics they are more familiar with teaching” (p. 317).There is danger that important ideas in probability may be ignored or not taught because teachers find them difficult to understand or not enjoyable.

Studies investigating teachers’ conceptual knowledge of probability have also been fruitful in unveiling teachers’ attitudes towards the use of technology and manipulatives for teaching probability. These studies have produced mixed results. Begg and Edwards (1999) found that technologies such as calculators and computers were not widely used by teachers. When technology was used, it was used for tasks like computing percentages or for word processing. In contrast, Watson in the above study revealed that teachers are using concrete materials and technology in their instruction. 70% to 81% of the teachers used calculators or computers, concrete materials, materials with chance outcomes, and sources of data to teach topics that involved chance or data.

Literature regarding the implementation of probability lessons by teachers is scarce (Stohl, 2005). Steinbring (1991) studied teaching episodes focusing on the concept of chance in German classrooms. In these episodes, students conducted some type of chance experiment, described the outcome with a probability model, and tried to explain the difference between the experimental results and their theoretical predictions. Steinbring found that the concept of “chance emerged as a universal object for explaining the connections between the actual outcome of an experiment and theoretical prediction” (p. 518). Students explained the differences in terms of “magic” or “luck.” Steinbring states that reexamining experimental conditions and assumptions of a theoretical model is important in the development of probabilistic reasoning. The teacher's role in this analysis is pivotal in terms of questioning and fostering social interaction in the classroom.

Haller (1997) also studied classroom teachers' instructional practices of probability. Haller selected 4 cases from 35 participants who were involved in a professional development project during the summer. Each of the 4 participants in her study was an in-service middle school teacher in the United States. Based on pretests from the summer institute, teachers initially exhibited a weak understanding of probability. Some participants specifically showed weaknesses in their knowledge of two-stage events, multi-stage events, and conditional probability. Through classroom observations, interviews, and other assessment instruments, Haller found that these teachers held misconceptions which adversely affected the implementation of their probability lessons. They also failed to use multiple representations of probability (i.e. fractions, decimals, percents) while teaching probability. These teachers also relied on their textbooks for instruction.

Due to these misconceptions and difficulties, the next chapter will concentrate on the theory of probability. Various examples using coins, dice, marbles and cards will be used to explain this theory.

CHAPTER 3: PROBABILITY THEORY

3.1 Definition and Introduction

Probability theory is that part of mathematics that aims to provide insight into phenomena that depend on chance or on uncertainty. It is the branch of mathematics that measures and models the frequencies of events in the real world in situations that can be repeated over and over indefinitely. The most prevalent use of the theory comes through the frequentists' interpretation of probability in terms of the outcomes of repeated experiments, but probability is also used to provide a measure of subjective beliefs, especially as judged by one's willingness to place bets.

The origin of probability theory is found in the interest taken by certain early modern mathematicians in games of chance. They include Gerolamo Cardano in the sixteenth century, Pierre de Fermat and Blaise Pascal in the seventeenth century. The kind of phenomena we observe when tossing of coins, throwing dice, picking cards from a shuffled pack, or randomly selecting a marble from a bag containing many marbles of the same color are the motivating experiences for the theory. Let us now view some examples.

Consider what happens when we toss a symmetrical coin. After a large number of tosses the relative frequency of heads, compared to the total number of throws, will be about ½. We are not able to predict the exact relative frequency, and it is even possible that no heads occur, yet when the tosses are repeated over a long time, the relative frequency of 'heads' tends to be near ½. Therefore, it is natural to say that the probability of throwing a head in a single toss is ½.

For another motivating example, imagine that we have a jar containing 100 marbles, all of the same size and weight. Assume that 20 are red, 35 are black and 45 are white. We draw marbles one at a time, and after each draw we return the marble to the jar and shake the jar thoroughly. After a large number of draws, we will find that the relative frequency getting red is 20/100, black is 35/100 and white 45/100. As in the case of the coin, it is natural to view this fraction as a property of a single draw. We say that the probability of drawing red is 20/100, black is 35/100 and white is 45/100.

A third example involves tossing a coin 5 times. In this case, there are $2^5 = 32$ different possible outcomes. The list of all of them begins HHHHH, HHHHT, HHHTH, HHHTT, etc. (see table 1 on page 22). If you write out all 32 possible outcomes, you will find only 10 outcomes in which there are exactly two tails. For instance, a few ways are TTHHH, THTHH, THHHT, etc. Since there are 10 ways to get exactly two tails and 32 possible outcomes all of equal probability, the probability of getting exactly two tails is 10/32.

3.2 Finite Probability Models

We will formalize the main features of the examples. In all cases, we have an action that we can repeat indefinitely, and each time we perform the action, we have a specific outcome from a well-defined set. For example, the coin may land on heads or tails and the marble in the second example maybe red, black or white. In the third example, we get a sequence of 5 heads or tails, and there are exactly 32 possibilities. Prior to performing the action, we do not know which outcome we will get, but we do know the full set of possibilities explicitly and in complete detail.

The first part of a finite probability model is a finite set containing all the possible outcomes. The outcomes are exhaustive (include all possibilities) and mutually exclusive (cannot have two occurring at once). By tradition, this set is called the “sample space” of the model. For example, if you toss a coin the sample space contains {H, T} and if you throw a die, the sample space S is {1, 2, 3, 4, 5, 6}. In another example, suppose five marbles, each of a different color, are placed in a bowl. The sample space for choosing one marble, from above, is {red, blue, yellow, green, and purple}. In a third example in this case, suppose there are 3 red marbles and 2 blue marbles in a bowl. An individual picks three marbles, one at a time, from the bowl. The sample space for picking three marbles, one at a time, is all of the possible ordered combinations of three marbles, S = {(red, red, red), (red, red, blue), (red, blue, red), (blue, red, red), (blue, blue, red), (blue, red, blue), (red, blue, blue)}.

The second part of a finite probability model is an assignment of probabilities to the elements of sample space. (Before stating the details, we note that the procedure we will outline here does not generalize to continuous probability models. We will discuss the differences later). The probability of an element x of a sample space S is a number denoted p(x). The assignment of probabilities must obey the following rules:

- Axiom 1. For all x, $0 \le p(x) \le 1$.
- Axiom 2. If sample space $S = \{x_1, x_2, \ldots, x_n\}$, then $p(x_1) + p(x_2) + \ldots + p(x_n) = 1$

Examples:

1. If all the points in sample space have equal probability, and there are n points, then $p(x) = 1/n$ for all x. In a coin flip, p(H) = ½ and p(T) = ½. When a die is rolled, p(1) = p(2), ………, p(6) = 1/6. (This coin and die are fair).
2. If we had a biased coin and felt that a head was twice as likely to appear as a tail, then we would have, p(H) = 2/3, p(T) = 1/3.
3. If the experiment is spinning a spinner of various colors, and sample space is the set of colors, then p(x) = the proportion of the spinner dial containing color x. Using spinners, we can set up an experiment with any desired number of outcomes and any

desired probabilities for the outcomes, provided they satisfy the outcomes, by dividing and coloring the spinner dial as appropriate.

The third part of a model is the description of events. Any subset of sample space is called an “event”. The probability of an event is the sum of the probabilities of all outcomes in it. Let us motivate this idea by an example (See Table 1). Consider tossing a coin 5 times. We may be interested in determining the probability of getting more heads than tails. Now there are several ways this can happen. For example, we may get no tails at all. There is only one outcome like this. Or, we may get exactly one tail. There are five outcomes in which this happens. Or we may get two tails, and there are 10 outcomes in which this happens. This gives a total of 16 ways to get more heads than tails. There are 16 outcomes that satisfy the condition. We call the set of all these outcomes “the event of getting more heads than tails”.

Table 1. Outcomes of 5 coin flips. The outcomes with 2 tails are in bold. The outcomes with more heads than tails are shaded.

HHHHH	HHHHT	HHHTH	**HHHTT**	HHTHH	**HHTHT**	**HHTTH**	HHTTT
HTHHH	**HTHHT**	**HTHTH**	HTHTT	**HTTHH**	HTTHT	HTTTH	HTTTT
THHHH	**THHHT**	**THHTH**	THHTT	**THTHH**	THTHT	THTTH	THTTT
TTHHH	TTHHT	TTHTH	TTHTT	TTTHH	TTTHT	TTTTH	TTTTT

Let us close this section with some comments. Above, we described a finite model for a colored spinner, where the outcomes were the colors. It is also possible to make a continuous probability model for a spinner, where the outcomes are the possible final positions of the arrow. In this case, there are infinitely many outcomes, all of which are equally likely. The different colors, in this model, are events rather than outcomes. The important observation is that within this model it is impossible to define probabilities by assigning probabilities to outcomes, since all outcomes must have equal probability, so any non-zero assignment violates Axiom 2. The probabilities of specific point-outcomes are not used. This is why we chose to use small p for probabilities of outcomes in the finite case, out of consideration for the fact that in non-finite models, we do not pay any attention to the probabilities of specific outcomes, but only events. The distinct notation also has a pedagogical purpose, in that it emphasizes the different status of outcomes (<u>elements</u> of sample space) and events (<u>subsets</u> of sample space).

3.3 Some Probability Laws

Let A be an event. By definition, $P(A) = p(a_1) + p(a_2) + p(a_3) + \ldots + p(a_n)$, where a_1, a_2, a_3, …, a_n are the elements of A listed one time each, and $p(a_i)$ is the probability of a_i. Note $P(\emptyset) = 0$. From this we immediately deduce the following:

- If $A \cap B = \emptyset$, then, $P(A \cup B) = P(A) + P(B)$.
- For any events A, B, $P(A \cup B) = P(A) + P(B) – P(A \cap B)$.
- Let $A^c = \{b \in S | b \notin A\}$. Then $P(A \cap A^c) = P(A) + P(A^c) = 1$, so, $P(A^c) = 1 – P(A)$

At this point we summarize ideas and terminology that we have introduced in Table 2; see next page.

3.4 Extended Example

Consider a single roll of two dice, a red one and a green one. This is a repeatable action, and whenever it is performed we get an "outcome". The table below shows the set of outcomes in the sample space, S. Each outcome is a pair of numbers (the number appearing on the red die and the one on the green die). There are 36 possible outcomes. Note that as required, the set of outcomes is exhaustive (each time we roll the dice, we must get one of these 36) and mutually exclusive (we cannot).

The second part of modeling this situation is assigning probabilities to outcomes. In this case, we assume the dice are fair. No outcome is any more or less probable than any other.

Table 2: Summary of basic terminologies used in probability

Term	Meaning	Example 1	Example 2	Example 3
Experiment	A probabilistic experiment is simply the act of doing something and noting the outcome	Flip a fair coin three times and observe the pattern of heads and tails.	Flip a coin and throw a die at the same time and observe the face of coin sides and the number on the die.	Blindly throw a dart at a wall that is painted part red, part green, part blue, and part white.
Sample Space	The set of all possible outcomes.	{HHH,HHT,HTH,HTT, THH,THT,TTH,TTT}	{H1,H2,H3,H4, H5,H6,T1,T2,T3, T4,T5,T6}	{Red, Green, Blue, White}
Probability of an outcome	The relative frequency of the outcome, when the experiment is performed many times.	Assuming the coin is fair, each of the outcomes will occur roughly the same number of times in many repeats of the experiment. Therefore, each outcome has probability of 1/8.	Each outcome is equally likely. So probability of each is 1/12.	Probabilities of different outcomes will depend on how much of the wall is covered by each color.
Event	A set of outcomes	"two tails" = {HHT,HTH,THH}	"Head and a prime number" = {H2,H3,H5}	"Dart hits primary color" = {Red, Green, Blue}
Probability of an event	The relative frequency of the event, when the experiment is performed many times.	The probability of "two tails" is 3/8, since there are 3 outcomes in the event, and each has probability 1/8	The probability of getting "head and a prime number" = 3/12	Probabilities of different events will depend on how much of the wall is covered by each color.

So, all outcomes have the same probability. Since there are 36, each outcome has probability 1/36.

p((1,1)) = p((1,2)) = p((1,3)) = p((1,4)) = p((1,5)) = p((1,6)).

An event is a subset of the entire sample space. The event that consists of the whole sample space is the event that some one of the outcomes occurs. This event is certain to happen; if we roll the dice, the outcome cannot be something other than the 36 outcomes listed in the table. Therefore, the probability associated with the event S is P(S) = 1. Also the 36 possible outcomes are mutually exclusive, that is, only one of the outcomes can happen on any roll of the dice.

Table 3. Set of outcomes in the sample space, S.

		Number on Red Die					
		1	**2**	**3**	**4**	**5**	**6**
Number on Green Die	**1**	(1, 1)	(1,2)	(1,3)	(1, 4)	(1, 5)	(1, 6)
	2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
	3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
	4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
	5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
	6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Let A be the event that the sum of the numbers on the dice is 7. This event consists of 6 of the possible outcomes, which lie in the diagonal of Table 3.

A = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}.

By definition, the probability of event A is:

P(A) = p(1,6) + p(2,5) + p(3,4) + p(4,3) + p(5,2) + p(6,1)

= 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 = 1/6.

Let B be the event that the number showing on the green die is 1. This event (the first row in Table 3) may also be described by listing the outcomes that it contains:

$B = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)\}$

Again, by the sum law for the probability of mutually exclusive events,

$$P(B) = P(1,1) + P(1,2) + P(1,3) + P(1,4) + P(1,5) + P(1,6)$$

$$= 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 = 1/6.$$

Let C be the event that the number showing on the green die is 6. This event may also be described as the event that the outcome of the roll of the dice is:

$C = \{(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$

As before, we find P(C) = 1/6.

Finally, let D be the event that neither number appearing on the dice is greater than 4 (the region inside the double lines in Table 3). Then,

$$D = \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4)\},$$

The probability of this event is:

$$P(D) = P((1,1)) + P((1,2)) + P((1,3)) + P((1,4)) + P((2,1)) + P((2,2)) + P((2,3)) + P((2,4)) + P((3,1) + P((3,2)) + P((3,3)) + P((3,4)) + P((4,1)) + P((4,2)) + P((4,3)) + P((4,4)) = 16/36 = 4/9.$$

3.5 Conditional Probability

Suppose that E and F are events in some model. We may consider a scenario where we know that F has occurred, and we are interested in the probability that E occurred, conditional on the occurrence of F. For example, suppose a person has flipped two coins, and he tells us that he has at least one head (i.e., F = {HH, HT, TH}). Now he asks us to bet on the proposition that at least one tail has occurred (i.e., E = {HT, TH, TT}). If both coins are fair P(F) = ¾ = P(E). But if we know that F has occurred, then there are two ways for E to occur out of the three ways for F to occur. So knowing F, the probability of E = 2/3.

When E, F are events, we write P(E|F) to denote the probability of E, given that F has occurred. The rule for computing P(E|F) is: $P(E|F) = P(E\cap F)/P(F)$. The rationale for the definition is as follows. If we assume that F has occurred, we are no longer dealing with all the outcomes in the sample space, but only those in F. So, F itself becomes a new sample space. But now we need to normalize all probabilities so they add up to 1. The way to do this is to divide all the original probabilities by P(F). Conditional probability is simply moving to a new sample space, consisting of the outcome in the condition event.

Here is another example. Suppose someone has rolled two fair dice. He reports that he has at least one 6. The probability that he has rolled 7 can be computed as follows. The event of rolling 7 is A:= {(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6,1)}, and P(A) = 1/6. The event of at least one six is B:= {(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6), (6, 5), (6, 4), (6, 3), (6, 2), (6, 1)}, and P(B) = 11/36. Note that $A\cup B$ = {(1, 6), (6, 1)}, so $P(A\cup B)$ = 2/36. Applying the definition of conditional probability gives:

$$P(A \mid B) = P(A\cup B) / P(B) = 2/36 \div 11/36 = 2/11$$

Hence the desired probability is 2/11 (and not 1/6).

3.6 Independence

Two events of non-zero probability are said to be "independent" if the probability that one of them occurs is not affected by the occurrence or non-occurrence of the other. Here is an example. Refer back to the Extended Example (3.4), and consider events A (getting a sum of 7) and D (numbers on both dice are at most 4) defined there. Note that $A\cup D$ = {{3, 4), (4, 3)} These events are not independent, because P(A) = 1/6, while

$$P(A \mid D) = P(A\cup D)/P(D) = 2/36 \div 16/36 = 1/8$$

Hence $P(A) \neq P(A \mid D)$, so A and D are not independent. The event D affects the probability of A, in the sense that if we know that D has occurred, then our expectation of A declines.

Note that when we checked independence, the roles of the events were different. In fact, it does not matter what role the events play, as we now show. Suppose that A and B have non-zero probability. Then:

$$
\begin{aligned}
P(A) = P(A \mid B) \quad &\Leftrightarrow \quad P(A) = P(A \cup B)/P(B) \\
&\Leftrightarrow \quad P(A)\, P(B) = P(A \cup B) \\
&\Leftrightarrow \quad P(B) = P(A \cup B)/P(A) \\
&\Leftrightarrow \quad P(B) = P(B \cup A)/P(A) \\
&\Leftrightarrow \quad P(B) = P(B \mid A).
\end{aligned}
$$

This also shows that A and B are independent if and only if $P(A)\, P(B) = P(A \cup B)$.

3.7 Probability and the Law of Large Numbers

The law of large numbers has an informal, common sense version and a rigorous mathematical version. The former states that if you repeat a random experiment, such as tossing a coin or rolling a die, many, many, many times, the relative frequency of each outcome tends to be equal to the probability of the outcome. For example, suppose you have a fair coin, such that the probability of getting heads is the same as that of getting tails and you toss this coin 100 times. You cannot predict exactly how many tails and heads you will get, but you can give an estimate: about 50 heads and about 50 tails. Maybe there will be a little more on one side or the other, say maybe 48 heads and 52 tails, but somewhere in that range. What would you think if you got 90 heads and only 10 tails? Even worse, what if you got all heads and no tails at all? Certainly it is conceivable that by a great coincidence you could toss the coin 100 times and get all heads, but then you would have a very strong reason to suspect that your coin was not fair at all.

The common sense law of large numbers, says that if you toss a fair coin many times, say 100, you should expect to get about the same number of heads and tails. Furthermore, if you toss it many more times than 100, say 10,000, you would expect to get even closer to a 1:1 ratio of heads to tails. The more times you repeat the experiment, the closer the relative frequency of heads is to the probability of getting a head.

In some sense, the informal law of large numbers is just a description of what it means to be a fair coin. In another sense, it is an empirical observation about the way that symmetrical coins behave. If a coin is flipped only a few times, the results may not indicate there are equal chances of it landing on heads and tails. For example, flipping a fair coin four times may yield three heads and one tail. It could even yield four heads and no tails. However, the law of large numbers says that as the sample increases, the relative frequency of the outcomes will better and better reflect their probabilities. If a coin is flipped 200 times, there is a good likelihood the number of times it lands on heads and tails will be near 100 each. However, the law of large numbers does not predict it will be exactly 100 each, only that it will likely be more representative of the true range of possibilities than a smaller average.

The formal law of large numbers is a theoretical version of the informal law. It is a statement about a sequence of models. To give a full treatment of the formal law of large numbers, we need to use probability models in a more sophistical way. We can formulate the law as follows, “If a certain chance experiment is repeated an unlimited number of times under exactly the same conditions, and if repetitions are independent of each other, then the fraction of times that a given event A occurs will converge with probability 1 to a number that is equal to the probability that A occurs in a single repetition of the experiment.” (Tijms H.C., 2007). The theoretical law of large numbers can best be understood in the context of a random process where a fair coin is tossed an unlimited number of times. An outcome of this random process can be described by an infinite sequence of heads and tails, recording whether a head or a tail turns up with each toss. Since the informal law is as much as we can hope to convey in 7th grade, we do not discuss the formal law.

Having looked at probability theory and its laws, the next chapter will look at the Common Core State Standard for Mathematics (CCSSM) and the Louisiana Comprehensive Curriculum (LCC) for 7th grade probability. A highlight of the activities included in the LCC for 7th grade will be presented followed by activities that are in line with the CCSSM curriculum.

CHAPTER 4: SEVENTH GRADE PROBABILITY CURRICULUM

4.1 Introduction

This chapter is about teaching probability in a manner that students can understand. A good approach is to relate the concept of the probability of an event to the relative frequency of the event in numerous trials. Students have had many experiences that can be used as a basis of thinking about probability. These experiences can be used in building an intuitive understanding of probability of events. For example, if a coin is tossed 100 times, the resulting relative frequency of getting heads or tails can be used to estimate their respective probabilities.

The current Louisiana Comprehensive Curriculum introduces probability in seventh-grade by using activities that involve the fundamental counting principle (the number of ways to do two things is the number of ways to do the first times the number of ways to do the second). This combinatorial idea is independent of the central concepts of probability, as spelled out in Chapter 3. The curriculum then draws comparisons between "experimental probability" (relative frequency) and "theoretical probability" (probability of an event), but one wonders about the extent to which this drives the idea of a probability model home. Only in the last lesson (based on the story of game of *Jumanji*, read to the students by the teacher) do we begin any deep work to determine various probabilities. Perhaps the curriculum could be strengthened by including more activities that call for finding probabilities via models.

The Common Core State Standards for Mathematics (CCSSM) expect students to be much involved in computing the probabilities of various events by coming up with a sample space of outcomes, assigning probabilities to these and then finding the probabilities of various events from the sample space, in parallel to the description of the theory that we presented in

Chapter 3. The CCSSM explicitly advocates developing models and using them in finding probabilities. The standards do talk about experimental and theoretical probability, addressing the idea of relative frequency of an event and the relationship of this to probability, but the treatment of probability goes well beyond this. The systematic arrangement of ideas in these standards leave no doubt that it is a carefully-considered and conceptually-based plan for teaching seventh-grade probability.

I begin this chapter by reviewing the probability lessons from the 7^{th}-grade Louisiana Comprehensive Curriculum. Then I present the CCSSM standards. Following this, I suggest a series of examples that teachers might use. This list of examples is in no way exhaustive; it's only a guide, so teachers are at liberty to come up with more examples. I present a sample final that teachers might use, and finally I include a sample introductory lesson as a guide to teaching probability using the CCSSM.

4.2 Louisiana Seventh Grade Comprehensive Curriculum

The Louisiana Comprehensive Curriculum (LCC) has probability as the seventh unit in 7^{th} grade Mathematics. According to the LCC, this unit "solidifies basic counting as students compute probabilities from collected data and record the data in tables and charts to help analyze the outcomes of experiments. These experiments are both theoretical and experimental in nature."(LCC Revised 2008).

Apart from describing data, it is evident from Table 4 (see page 36) that the 7^{th} grade LCC has the relationship between experimental probability and theoretical probability as the main focus. It is expected that students already have the foundational knowledge from previous grades, as evidenced by the fact that students in their first and second activities (see below) are asked about the fundamental counting principle. The third activity goes straight into finding probability from data in tables. The most efficient way to describe what the curriculum does is simply to summarize the activities, as we now do.

The LCC 7th grade probability Grade-Level Expectations (GLEs) are summarized in the state document as follows:

Table 4. Grade-Level Expectations (GLEs) for Louisiana Comprehensive Curriculum, Grade 7 Unit 7: Data Analysis, Probability, and Discrete Mathematics

GLE#	**GLE Text and Benchmarks**
32.	Describe data in terms of patterns, clustered data, gaps, and outliers (D-2-M)
36.	Apply the fundamental counting principle in real-life situations (D-4-M)
37.	Determine probability from experiments and from data displayed in tables and graphs (D-5-M)
38.	Compare theoretical and experimental probability in real-life situations (D-5-M)

Activity 1: It's fundamental (GLE: 36). In this activity, the fundamental counting principle is recalled by using three different colored books to demonstrate the number of ways the books can be arranged or ordered when packed in a bookshelf. This is actually permutations, because the order of the individual items is important in determining how many arrangements can be made.

Activity 2: How many choices? (GLE: 36). This activity involves the application of the fundamental counting principle to real-life situations. Students are to use the items in menus to list the different orders that can be made by customers. The greater the variety of items that one can order, the more tedious the listing becomes. It is at this point that the idea of the fundamental counting principle kicks in and the students realize that they do not have to list all the possible options in the menu, but rather simply multiply. Example: there are 4 flavors of yogurt and 8 toppings. How many choices are available if you can have only one topping?

Activity 3: Determine probability from data (GLE: 37). This activity involves finding probabilities using basketball statistics about national players or players from the local high school. Here students are asked how to calculate players' free throw percentages. Students are expected to understand that these percentages represent the probabilities that can be used to predict future scorings by the players.

Activity 4: It's Theoretical! (GLE: 38). Students are led to the idea that the probability of a head on a coin toss is ½ and are told that this is called "theoretical probability". Then they are asked to make numerous coin tosses, record the outcomes and compute the ratio of heads to tosses. Students are told that this ratio is called "experimental probability"**.** Students are expected to observe that experimental probability approaches theoretical probability as the number of trials increases.

Activity 5: How do the chips fall? (GLEs: 37, 38). In this activity, students are given a bag containing colored chips and are asked to find the theoretical probability of picking each color. Students are then asked to conduct an experiment where they are to pick chips at random from the bag so as to find their respective experimental probabilities. Students are then asked to compare the two probabilities.

Activity 6: Birthdays (GLEs: 32, 37). Here, students are asked to collect data from class on days of the month that each one was born. Students are then asked to organize the data in a line plot. Then they are supposed to analyze the data and identify any clusters, gaps, outliers or patterns. From this data, students also find probabilities of different birth dates.

Activity 7: Probability with *Jumanji* (GLEs: 36, 37, 38). This is the last activity in the probability unit. In this activity, the teacher reads to the class the book *Jumanji*. While reading, the teacher is supposed to stop at various points to ask probability questions. At various points in the story characters roll two dice, and this determines their fate and whether they win or lose (according to some complex set of rules). After the reading, the students are given a worksheet where they compute the theoretical probabilities of different sums (on two dice) and find experimental probabilities for the same. The worksheet emphasizes the contrast of experimental and theoretical probability, in line with the Louisiana GLEs.

4.3 Common Core State Standard for Mathematics (CCSSM)

The Common Core State Standards for Mathematics were developed by The Common Core State Standards Initiative (CCSSI), a state-led effort coordinated by the National Governors Association Center for Best Practices (NGA Center) and the Council of Chief State School Officers (CCSSO). The standards claim to be informed by the highest, most effective models from states across the country and countries around the world, and to provide teachers and parents with a common understanding of what students are expected to learn. Teachers, school administrators and experts collaborated in the writing to provide a clear and consistent framework to prepare children for college and the workforce.

The CCSSM recommend that investigations of chance processes and the development, use, and evaluation probability models be concentrated in the 7th grade. In contrast, the Louisiana Grade-Level Expectations (GLEs) for probability are spread from 6th to 8th grade. According to the CCSSM, students in 6th grade should get involved in data collection and analysis and in summarizing and describing distributions. Probability is covered in 7th grade. The CCSSM proposes a considerably deeper 7th grade probability curriculum than the Louisiana Comprehensive Curriculum. It starts off the topic with the idea of quantifying likelihood or expectation. This idea is developed in the context of both theoretical and experimental probabilities. The CCSSM go on to recommend the use probability models to find probabilities and to calculate the probabilities of compound events using organized lists, tables, tree diagrams and simulations. Let us discuss the standards one at a time.

Standard 7.SP.5 says that students should understand that probability is a number that quantifies likelihood. The probability of a chance event is a number between 0 and 1. Larger numbers indicate greater likelihood. A probability near 0 indicates an unlikely event, a probability around ½ indicates an event that is equally likely (an event that is neither unlikely nor likely), and a probability near 1 indicates a very likely event.

Standard 7.SP.6 requires the approximation of probability of a chance event by collecting data on the chance process that produces it, observing its long-run relative frequency, and predicting the approximate relative frequency given the probability. This brings in the idea of experimental probability which is defined as the ratio of the number of times the event occurs to the total number of trials. It is in this standard that experimental and theoretical probabilities are compared.

Standard 7.SP.7 calls for the understanding of probability models that are both uniform and those that are not. To understand these two types of models here are examples. Suppose you have a spinner that is divided into 5 proportional parts labeled 1, 2, 3, 4, 5. Finding probabilities here is an example of a uniform model. But if you have a spinner labeled 1, 2, 2, 2, 3, then using this to find probabilities of spinning any of these numbers is an example of a non-uniform model. Sample problems on these two types of models are given later in the collection of problems.

Finally, **standard 7.SP.8** is all about finding probabilities of compound events. This standard recommends the use of tables, tree diagrams and Venn diagrams to represent compound events. Using simulations in finding probabilities is also emphasized here. A Venn diagram, named after John Venn, a 19^{th}-century English mathematician, consist of a rectangular region containing several circles. Using Venn diagrams make students visualize the information given thus simplifying the process of finding the required probabilities.

4.4 Teaching Probability

In the present section, I will make various recommendations about the teaching of probability, based on the material I presented earlier. Teaching probability calls for careful choice of problems to be solved. These problems should be attractive, meaningful and suitable for the students.

Experiments such as tossing a coin, rolling a dice, spinning a roulette wheel, taking a card from a well-shuffled deck are suitable for acquiring experience with the phenomena of chance, the idea of a probabilistic experiment and the notion of sample space. In the early lessons on probability, students should be equipped with coins, dice, a deck of cards, and cardboard for making spinners. Before asking about the probabilities, the teacher should be sure that students have a clear idea of the possible outcomes, at the level of an exhaustive and mutually-exclusive list, in order to bring home the importance of sample space. What can we see if we flip a coin? What if we flip a penny and a nickel? What are the possibilities? What about a penny a nickel and a dime? These are similar to the kinds of questions that the LCC begins with, but the emphasis is not on counting, but more simply and more relevantly on listing.

After sample space has become an idea that is clear enough to work with, we can begin to think about probabilities, determined experimentally or theoretically. The teacher may ask questions such as:

- What sum of the points is most often obtained if two dice are rolled repeatedly?
- What is the chance of obtaining three heads if three coins are tossed?
- What are the odds that two cards of the some denomination appear in five-card poker, hand taken from a well-shuffled deck?
- What is the chance of spinning an even number on a uniform spinner labeled 1 to 7?

After students are used to talking about and discussing the probabilities of outcomes, the teacher may begin to develop the idea of a complex event as a set of outcomes and show how the probabilities of events may be found.

There are many text books currently being used in the 7th grade classroom that address basic probability concepts like coin tossing, spinning spinners, and picking marbles from bags and so on. Most of these examples involve the use of theoretical probability, which is easily found when all possible outcomes of a simple experiment are "equally likely". At this point I present a collection of problems that can be used in exploring the above standards for deeper understanding. Later in this chapter, I have presented two games for further exploration of probability. The chapter ends with a presentation of a sample test that can be used to assess the grasping of the ideas learned in probability.

4.5 A Collection of Basic Problems and Solutions

Problem. Two standard dice are thrown. What is the probability that the sum of the two dice is 14? What is the probability that the sum of the two dice is less than 13?

Solution. Since the maximum number of dots on a die is 6, it is impossible to get a sum of 14 with two standard dice. Therefore, P(Sum of 14) = 0. Since the sum of two standard dice is always less than 13, the event is a certain event. The probability of an event that is certain is 1. Therefore, P(Sum less than 13) = 1

Problem. A coin is tossed 60 times. Heads appear 27 times. Find the experimental probability of getting heads. Assuming the coin is fair, how do the experimental and theoretical probabilities of getting heads compare?

Solution. Experimental probability = (number of times the event occurs) / (total number of trials). Heads appeared 27 times in 60 tosses, so the experimental probability is 27/60 = 9/20.

The theoretical probability of getting heads is ½. In this case, the experimental probability is slightly less than the theoretical probability.

Problem. You take a card at random from a pack of 52 cards. What is the probability that the card is: (a) a king, (b) a diamond, (c) a black card?

Solution. (a) P(king) = 4/52 = 1/13, since there are 4 kings in 52 cards, all equally likely to be chosen. (b) P(diamond) = 13/52 = ¼, since there are 13 diamonds. (c) P(black) = 26/52 = ½, since there are 26 black cards.

Problem. Suppose two players use a six-sided die that has the sides numbered 1, 1, 2, 2, 3, and 4. There are four numbers that can come up: 1, 2, 3, and 4. The first player wins when 1 or 2 shows and the second player if 3 or 4 shows. Is the game fair?

Solution. Even though each player has two winning numbers, these numbers are not equally likely! Of the 6 outcomes corresponding to the different faces, four produce a win for the first player.

Problem. A fair die is renumbered so that it has the following numbers on its faces: 1, 1, 2, 3, 4, 6. What is the probability that the outcome is: (a) 1? (b) 3? (c) 5? (d) Greater that (>) 2?

Solution. (a) P(1) = 2/6 = 1/3. (b) P(3) = 1/6. (c) P(5) = 0. (d) P(greater than 2) = 3/6 = ½.

Problem. Two dice, red and blue, both unbiased, are rolled at the same time. Use a table to show all the possible outcomes. If the scores on the two dice are then added together, what is the probability of obtaining (a) a sum of 5; (b) a sum greater than (>) 3, (c) a sum that is an even number?

Solution. (Refer to Table 3 in page 28)

(a) P(a sum of 5) = 4/36 = 1/9 (b) P(a sum greater than 3) = 33/36 = 11/12

(c) P(a sum that is an even number) = 18/36 = 1/2

Problem. A coin is tossed twice. Use a tree diagram to show the possible outcomes and calculate their probabilities

Solution. We first note that, for a single toss, P(Head) = ½ and P(Tail) = ½. We now go ahead and put these probabilities on the branches of the tree diagram as shown:

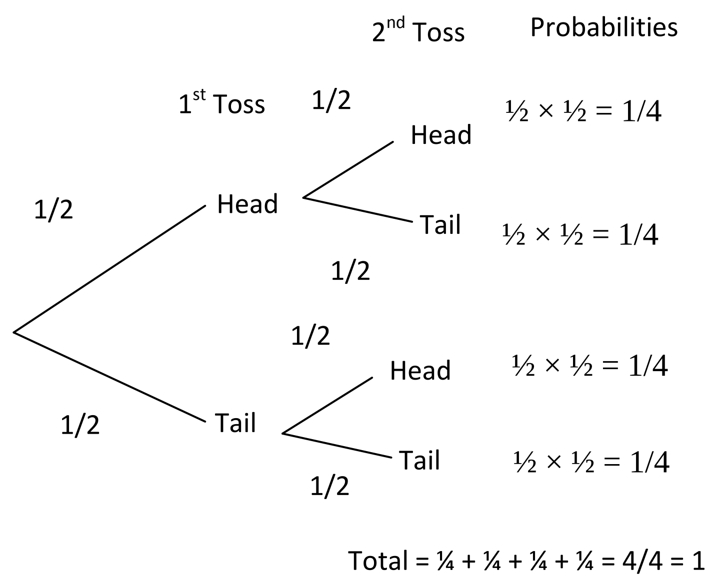

Figure 1: Tree diagram representation of two coin tosses

Problem. In a group of 100 people, there are 18 people with red hair, 14 people with blue eyes, and 10 people with both red hair and blue eyes. Find the probability of a random person in this group having:

(a) either red hair or blue eyes.

(b) red hair but not blue eyes.

(c) both red hair and blue eyes.

(d) red hair or blue eyes but not both.

Solution. First, put the information in a Venn diagram. There are a total of 18 people with red hair and 10 people with both red hair and blue eyes, so there must be 8 people inside the circle R but outside circle B. There are a total of 14 people with blue eyes, but 10 of them are in the intersection, so there are 4 inside the circle B but outside circle R. There are 100 people in total and we have already placed 8 + 10 + 4 = 22 of them, so there must be 100 – 22 = 78 people outside of the circles.

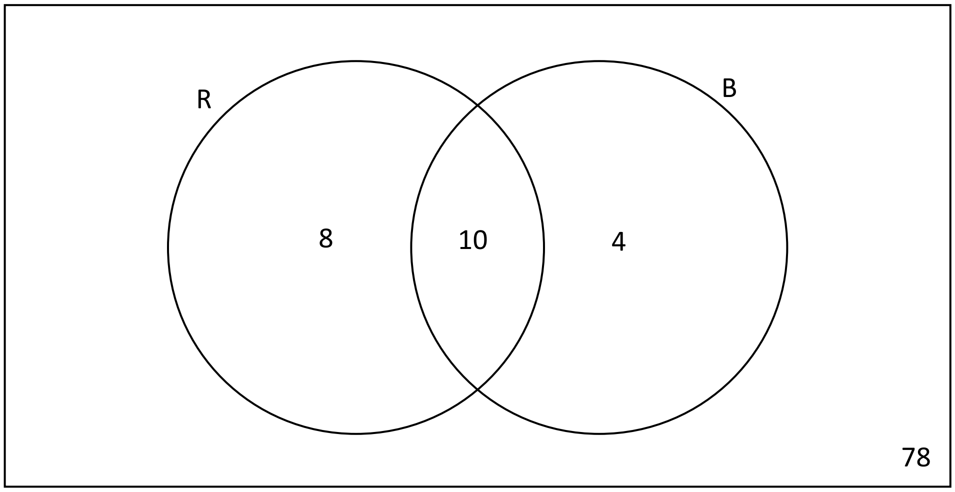

Figure 2: Venn diagram on hair and eye color.

(a) P(R or B) = P(R $\cup$ B) = $\frac{8 + 10 + 4}{100}$ = 22/100 = 22%

(b) P(R ∩ B^c) = 8/100 = 8%.

(c) There are 10 people with both red hair and blue eyes; we can represent this as the intersection of two circles, R and B. P(R∪B) = 10/100 = 10%.

(d) The people who have red hair or blue eyes but not both are read directly from the diagram. There are 8+4 of them. P(R or B but Not Both) = 8/100 + 4/100 = 12/100 = 12%.

4.6 Sample Final Assessment

Answer ALL the questions in this test.

Show All the explanations to earn full points.

1. Decide whether or not each of the statements below is reasonable.

(a) The probability that it will snow on Christmas Day in Baton Rouge is 0.9.

(b) The probability that you will win a raffle prize is 0.5.

(c) The probability that you will go to bed before midnight tonight is greater than 0.8.

(d) The probability that your pocket money is doubled tomorrow is less than 0.2.

2. Describe something that is:

   (a) Very unlikely, (b) Unlikely, (c) Likely, (d) Very likely.

3. Explain why the probability that you will be the first person to walk on the moon is zero.

4. Describe something that has a probability of 1.

5. When you toss a fair coin, the probability of obtaining a head is ½ and that of obtaining a tail is ½. Describe something else that has a probability that is equal to or close to ½.

6. A pack of candy contains mostly red candies, a few green candies and only one yellow candy. You take a candy at random from the pack. The events A, B, C and D are listed below.

   A – You take a yellow candy.

   B – You take a green candy.

   C – You take a red candy.

   D – You take a blue candy.

   Write these outcomes in order of probability, with the most likely first.

7. The faces of a regular tetrahedron are numbered 1 to 4. When it is rolled it lands face down on one of these numbers. What is the probability that this number is:

   (a) 2, (b) 3, (c) 1, 2 or 3, (d) an even number?

8. A card is taken at random from a full pack of playing cards. What is the probability that it is:

   (a) A red card? (b) A “queen”?

   (c) A red “ace”? (d) The “seven of hearts”?

   (e) An even number?

9. You roll a fair die 120 times. How many times would you expect to obtain:

   (a) A 6? (b) An even score? (c) A score of less than 5?

10. A spinner in the shape of a regular octagon has its faces numbered as follows; 1, 1, 1, 1, 2, 2, 3, 4. Each score is equally likely to occur. What is the probability of scoring:

    (a) 1, (b) 2, (c) 3, (d) 2 or larger, (e) 5, (f) A number less than 6?

11. A spinner has numbers 1 to 5, so that each number is equally likely to be hit. How many times would you expect to get a score of 5, if the spinner is spun:

    (a) 10 times, (b) 250 times, (c) 400 times?

12. Kevin has 5 white beads and 1 black bead in his bag. He asks two friends about the probability of picking a black bead without looking in the bag.

    Owen says: "It is 1/5 because there are 5 white beads and 1 black bead."

    Maria says: "It is 1/6 because there are 6 beads and 1 is black."

    (a) Which of Kevin's friends is correct? Explain why the other friend is wrong.
    (b) Joy has a different bag of black beads and white beads. The probability of picking a black bead from Joy's bag is 7/13. What is the probability of picking a white bead from Joy's bag?
    (c) How many black beads and white beads could be in Joy's bag?
    (d) Cephas has a different bag of black beads and white beads. Cephas has more beads in total than Joy. The probability of picking a black bead from Cephas' bag is also 7/13. How many black beads and how many white beads could be in Cephas' bag?

13. An unbiased coin is tossed twice.

    (a) List all the possible outcomes. (b) What is the probability of obtaining two heads?
    (c) What is the probability of obtaining a head and a tail in any order?

14. A red die and a blue die, both unbiased, are rolled at the same time.

(a) Use a table to show all the possible outcomes.

(b) If the number on the dice are added, what is the probability of obtaining:

(i) A score of 5, (ii) A score which is greater than 3,

(iii) A score that is an even number?

15. A card is taken at random from a pack of 52 playing cards. Without replacing it, a second card is then drawn at random from the pack. Use a tree diagram to determine the probability that:

(b) Both cards are diamonds, (b) At least one card is a diamond,

(c) Exactly one card is a diamond, (d) the second card is a diamond.

16. Tess will toss a fair coin 3 times. The possible results are illustrated in the tree diagram below. Based on the information given in the tree diagram, in how many ways (outcomes) can Tess toss at least 2 heads?

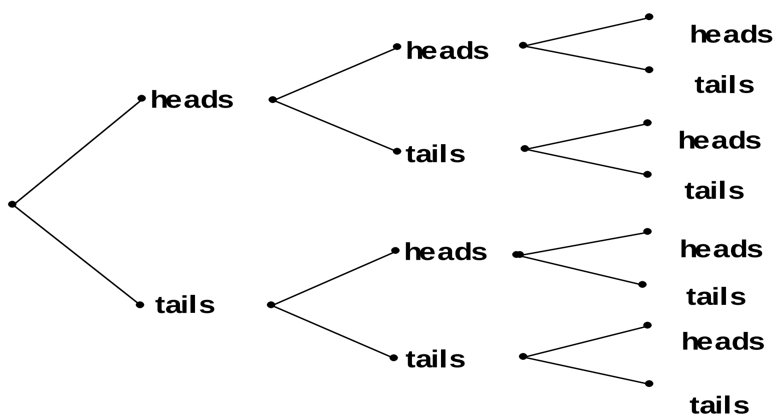

Figure 3: Results of a coin tossed 3 times.

17. List the 8 possible outcomes when 3 fair coins are tossed at the same time. How many times would you expect to get:

(c) 3 heads, (b) 2 heads, (c) 1 head, (d) 0 heads?

18. 100 students were asked whether they studied French or German. 27 students studied both French and German.

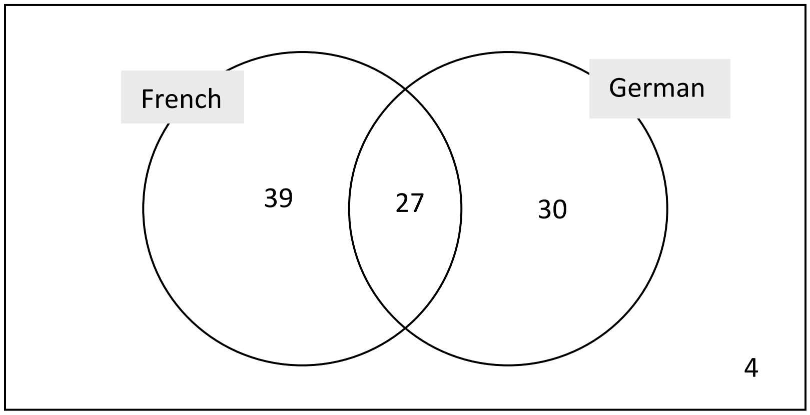

Figure 4: Results of subjects studied.

(c) What is the probability that a student chosen at random will study only one of the languages?

(d) What is the probability that a student who is studying German is also studying French?

(e) Two of the 100 students are chosen at random. From the following calculations, select one which shows the probabilities that both students study French and German.

(i) $27/100 \times 26/100$ (ii) $27/100 + 26/99$ (iii) $27/100 + 27/100$

(iv) $27/100 \times 26/99$ (v) $27/100 \times 27/100$

Answers to test questions

1. (a) Very unlikely (b) Very unlikely (c) Very likely (d) Very likely.

   Note: Answers may vary therefore students should explain their answers.
2. Answers will vary. Refer to question 1 for possible answers.
3. Because Neil Armstrong has already done so.
4. The probability that next year is 2011.
5. Picking a red marble from a bowl that contains only two marbles, 1 red and 1 blue.

6. Order from most likely. C, B, A, D. Note that the probability of picking a blue candy is 0 because there are no blue candies in the pack.
7. (a) P(2) = ¼ (b) P(3) = ¼ (c) P(1, 2 or 3) = P(1) + P(2) + P(3) = ¼ + ¼ + ¼ = ¾

   (d) P(an even number) = 2/4 = 1/2
8. (a) P(red card) = 26/52 = ½ (b) P(a queen) = 4/52 = 1/13

   (c) P(a red ace) = 2/52 = 1/26 (d) P(seven of hearts) = 1/52

   (e) P(an even number) = 20/52 = 5/13
9. (a) P(6) = 1/6 × 120 = 20 times. (b) P(an even number) = 3/6 × 120 = 60 times.

   (c) P(a score of less than 5) = 4/120 = 1/30
10. (a) P(1) = 4/8 = ½ (b) P(2) = 2/8 = ¼

    (c) P(3) = 1/8 (d) P(2 or larger) = 4/8 = ½

    (e) P(5) = 0/8 = 0 (f) P(a number less than 6) = 8/8 = 1
11. (a) P(5) = 1/5 × 10 = 2 times (b) P(5) = 1/5 × 250 = 50 times

    (c) P(5) = 1/5 × 400 = 80 times
12. (a) Maria is correct. Owen is wrong because he only counted the number of white beads.

    (b) P(W) = 1 – P(B) = 1 – 7/13 = 6/13

    (c) 7 black beads and 6 white beads.

    (d) Cephas could be having 14 black beads and 12 white beads. Here the students are being tested on equivalent fractions too. From this number of beads that Cephas has, the P(black bead) = 14/26 which reduces to 7/13.
13. (a) outcomes = {HH, HT, TH, TT}

    (b) P(two heads) = ¼

    (c) P(a head and a tail in this order) = ¼
14. (a) See Table 3 on page 28.

(b) (i) Outcomes with a score of 5 are (4,1), (3,2), (2,3), (1,4).

So, P(a score of 5) = 4/36 = 1/9

(ii) P(a score greater than 3) = 33/36 = 11/12

(iii)P(a score that is an even number) = 18/36 = ½

15.

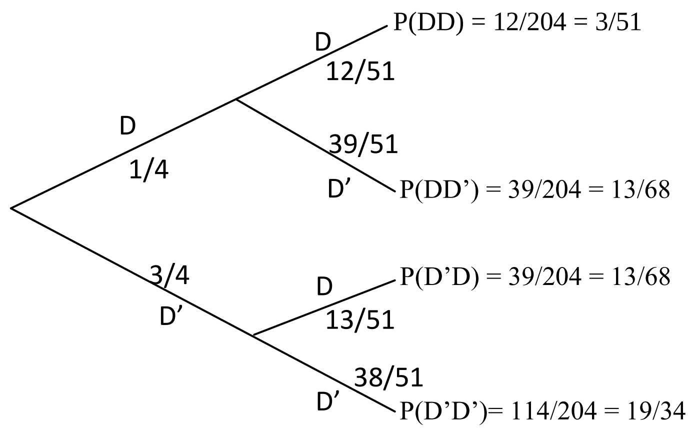

Figure 5: Picking two diamond cards without replacement.

(a) P(both cards are diamonds) = P(DD) = ¼ × 12/51 = 3/51

(b) P(at least one card is diamond) = P(DD) or P(DD’) or P(D’D)

= 3/51 + (1/4 × 39/51) + (3/4 × 13/51) = 3/51 + 13/68 + 13/68

= 3/51 + 26/68 = 15/34

16. P(at least two heads) = p(HHH) + p(HHT) + p(HTH) + p(THH)

= 1/8 + 1/8 + 1/8 = 3/8

17. Outcomes of 3 coin tosses are {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

(a) P(HHH) = 1/8 (b) P(2 heads) = P(HHT) + P(HTH) + P(THH) = 3/8

(c) P(1 head) = P(HTT) + P(THT) + P(TTH) = 3/8 (d) P(0 heads) = P(TTT) = 1/8

18. (a) P(only one of the languages) = P(French) or P(German) = 39/100 + 30/100 = 69/100

(b) P(German & French) = P(German ∩ French) = 27/100

(c) The correct answer is (ii) = $27/100 \times 26/99$. This picking is without replacement. Thus after picking the first student who studies both German and French, there will be 26 students who study both languages left in the class and the total number of students in the whole class will be 99.

4.7 Lesson Proposals

In an effort to provide guidance on teaching experimental and theoretical probabilities this thesis proposes some lessons.

a. From experimental to theoretical probability. This lesson will introduce students to the concepts of experimental probability and use this as a basis for introducing the idea of theoretical probability. Students will collect data in an experiment involving a spinner and observe trends in the data as the number of spins increases.

Prerequisites. Students should know how to write ratios and convert those ratios to percent and decimals.

Learning Objectives. On completion of this lesson students will be able to:

- Conduct an experiment with repeated trials and determine the relative frequency of a specific outcome.
- Estimate the probability of a specific outcome experimentally and write the estimate as a ratio, percent, and decimal.
- Understand the theoretical probability of an outcome as an abstraction that describes the tendencies in data from repeated experiments.
- Recognize that repeated trials will result in a closer match between the theoretical probability and relative frequency.

Materials/resources.

- Pencils and paper
- Color spinners: a large one for display and small copies for each group

Activities.

1. The first phase develops a qualitative sense of chance outcomes, upon which we will soon build a quantitative sense. Display a large spinner that is 60% red, 30% blue and 10% green. (Don't mark or say the precise sizes of the regions.) Spin it several times and record outcomes on the board. Ask students to answer the following questions:
    - "Can we tell before a spin what color we will get?
    - "If we spin 10 times, what color will we get most often? Can we be sure? How about least often?
    - "If we spin 10 times, will we get red more than half of the time? Can we be sure"
    - "What will happen if we spin more times?"
2. In this phase, transition to a more quantitative understanding, but we note that with small number of trials (10, in this case), there is much uncertainty. Distribute small color spinners with the same coloring as the large one to each group. Instruct the groups to spin 10 times and record results. Ask each group what percent of times (of the 10 spins) that they got red. Compile all the results from all the groups on the board, and calculate what percent of times the whole class got red.
3. In this phase, we provide the experiences students need in order to have an experimental base for understanding the law of large numbers. Ask each group to spin 100 times and record the results. Ask each group to calculate the percent of each color. Write the results on the board. (Every group should have a percent of red near 60, but there will be

variation. Similarly, every group will have a percent blue near 30 and a percent green near 10, but there will be variation.)

4. In this phase, we develop the idea of the law of large numbers further. Combine all the results from the whole class, and compute percentages of each color. At this point, we can reveal what the percentages on the original display and the small spinners have been, and we can see that these percentages help to explain the combined results of the class.
5. In the last phase of this lesson, the teacher summarizes the experiences of the class and draws conclusions, such as, "a) the result of a single spin is unknowable. It is a chance event. b) In a large number of spins a certain pattern appears. The pattern becomes clearer and clearer as the number of spins increases. The ratios that we find in the data may be viewed as experimental estimates of a trend which would be clearer and clearer in vast amounts of data. A particular set of data give us an 'experimental probability'. 3) We may be able to explain the pattern and even predict it, if we know the properties of the spinner. Such a prediction, in the form of a number between 0 and 1, is called a 'theoretical probability'."

<u>Assessment.</u> The assessment can be informal. Ask students to summarize their own thoughts on the activity. They should include how they think you find theoretical probability, experimental probability, and explain how the two are related.

b. What is a sample space? Introduce the idea of sample space, and find the sample spaces for experiments of tossing several coins. Use this to compute the probabilities of events.

<u>Prerequisites.</u> None, beyond 7^{th}-grade maturity.

<u>Learning Objectives.</u> On completion of this lesson students will be able to:

- List the possible outcomes in a variety of probabilistic experiments.

- Use the list to compute theoretical probabilities.
- Understand that the probabilities computed in this manner predict the patterns in the outcomes when a specific experiment is repeated.

Materials/resources.

- Pencils and paper
- Five pennies, a nickel and a dime for each group

Activities.

1. Ask students what can happen if a penny and a nickel are spun or flipped. What are the possible results? Students may say several things, but ultimately the teachers should make sure that everyone understands that each coin can land on either a head or a tail. Have students list the four possible outcomes. This is an example of a "sample space."
2. Ask students: "If two coins (penny and nickel) are flipped, is any one of the four possible outcomes more likely than any other?" Lead them to the agreement that the answer is, "No. Each of the four is equally probable." Ask how many times the coins will land on different faces, if the pair is flipped 100 times.
3. Ask, "If 2 pennies are tossed, what are the possible outcomes?" Some students may say, "Zero heads, one head or two heads." Ask the class whether there are three possible outcomes, or four (as there would be if a penny and a nickel were tossed). Ask, "What is the probability of getting one head?" Some students may say 1/3 (which, of course, is not correct, but don't say this). If no student suggests this, then the teacher should. Ask students to justify their expectation, whether it be 1/3, 1/2 or something else.
4. Perform an experiment with the class. Have each group flip two pennies 50 times. Gather data from at least 200 flips of the pair, preferably more. Compile the data from

the class. With near certainty, the data will support the assertion that P(one head in two flips) is 1/2.

5. Explain the result using the sample space for two pennies.
6. Ask the class, “If three coins (penny, nickel and dime) are tossed, what are the possible outcomes?” Have the students create a list. Ask them, “What is the probability there are exactly two heads in one toss of the three?”
7. What is the sample space for 3 pennies? Create a list.
8. Write out the sample space for 4 pennies? For 6?
9. Close the lesson by summarizing the main ideas: the sample space for the experiment of tossing n coins has 2^n elements. The outcomes can be listed explicitly for n up to about 6 or 7. All the outcomes are equally likely, if the coins are fair. The sample space can be used to calculate the probabilities of events, such as “tossing 2 heads.”

<u>Assessment.</u>

1. Write out the sample space for the experiment of tossing 5 coins?
2. What is the probability of 1 head when 5 pennies are tossed? Of 2 heads? Of 3?

c. Using dice games to introduce the concept of events. One relatively common approach to introducing probability concepts involves the rolling of a pair of dice. Using multiple trials of the experiment, the sum of the numbers appearing on the two dice can be determined in order to promote the idea that some events are more likely to occur than others. Through successive rolls, students may discover, for example, that the event ‘‘7’’ is more likely than the event ‘‘12.’’ Many other events can be described, and independent events and conditional probabilities can be investigated. The discussion of dice games earlier in this thesis provides a basis for this

Games are another important strategy that is engaging and interesting and therefore can greatly be used to drive the idea of probability home. In this paper two types of games have been suggested. As we all know students in this generation are greatly advanced in the computer and other hand-held electronic games; thus including them in probability will highly motivate the students.

One interesting possibility is the Race Game, described in (shodor interactive activities). The in-class activity sheet for The Race Game appears in Figure 6. In this game, each group works with token race-cars numbered 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12. The group rolls a pair of dice, one red and one green, and determines the sum of the spots. Each roll determines which race car advances one space toward the finish line. In essence, the ‘‘race track’’ serves as a line plot to represent the frequency with which each event (2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12) has occurred. The winner of The Race Game is the player whose race car is first to cross a finish line located 10 spaces from the starting line. Before beginning play, students are asked to predict which car they thought would win.

As the game is played, the whole-class results may be tabulated on a ‘‘race track’’ on the board. Students are asked to announce which race car is ahead at each milepost; that is, which race car number is first to come up twice, which is first to occur three times. A list of racecars leading at each milepost is listed on the board.

Students should be able to see that race cars with numbers such 6, 7 or 8 have a greater probability of winning. This is due to the fact that these outcomes are more probable than outcomes like 2 or 12.

Yet another activity that can play a key role in fostering students’ probabilistic reasoning is the “Sum It Up “dice game. In this activity, students work in small groups and carry out numerous trials of a simulation. This game is played with a pair of dice. Prior to commencement of the game, students divide the numbers from 2 to 12 between the members, and then a member scores on a roll if one of his numbers comes up. After numerous plays, the questions of what numbers are most valuable may be discussed and solved using sample spaces.

10											
9											
8											
7											
6											
5											
4											
3											
2											
1											
Mile	Race Car #2	Race Car #3	Race Car #4	Race Car #5	Race Car #6	Race Car #7	Race Car #8	Race Car #9	Race Car #10	Race Car #11	Race Car #12

Figure 6: The Race Game (the finish line)

CHAPTER 5: SUMMARY AND CONCLUSIONS

In middle schools students are generally able to understand the purpose of tasks that are given to them. This is no less the case in mathematics. Students can appreciate the concept of uncertainty, but they need experience and instruction to overcome possible misconceptions and to build a conceptual framework that enables them to use modern probability ideas. There are enormous opportunities for teachers to link fractions, ratios, and proportional reasoning from elsewhere in the mathematics curriculum to probability problems. This may increase the motivation and understanding of students in relation to both areas of the curriculum as connections begin to be made.

Overall the middle school provides fertile ground for the development of probabilistic ideas that can lay a foundation for the formal work with complex sample spaces and probability distribution in high school years. This should be combined with a wide range of applications to create an appreciation for usefulness of probability for the "real world" (Jane Watson- 2005) as well as theoretical mathematical problem solving.

This explorative and descriptive study of the grade 7 probability curriculum presented in this thesis may be useful for a revised curriculum that is based on the Common Core State Standards for Mathematics. The tasks that have been presented in this thesis were chosen for mathematical relevance and pedagogical purposefulness, based on a review of mathematics and pedagogical advice gleaned from the literature.

REFERENCES

Ball, D. (2000). *Working on the inside: Using one's own practice as a site for studying teaching and learning.* In A. Kelly & R. Lesh (Eds.), Handbook of research design in mathematics and science education (pp. 243). Mahwah, NJ: Lawrence Erlbaum Associates.

Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning (pp. 15-37).* Netherlands: Kluwer Academic Publishers.

Begg, A., & Edwards, R. (1999, December). *Teachers' ideas about teaching statistics*. Paper presented at the combined annual meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education. Melbourne, Australia.

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

Common Core State Standards Initiative (2010). *Common Core State Standards for Mathematics.*

Conference Board of the Mathematical Sciences. (2001). *The Mathematical Education of Teachers*. Providence, RI: American Mathematical Society.

Consortium for Mathematics and Its Applications. *For All Practical Purposes: Introduction to Contemporary Mathematics*, 4th ed. New York: W. H. Freeman and Company, 1997.

Devlin, Keith. *Life by the Numbers*. New York: John Wiley & Sons, Inc., 1998.

Douglas B. Aichele, Arthur F. Coxford (1994*). Professional Development For Teachers Of Mathematics*: 1994 Yearbook (Yearbook (National Council Of Teachers Of Mathematics)) (Hardcover)

Fennema, E., Franke, M. L., Carpenter, T. P., & Carey, D. A. (1993). *Using students' mathematical knowledge in instruction.* American Educational Research Journal, 30, 555-583.

Frank A. Friday (1967). *The Elements of Probability and Sampling.* Barnes & Noble, New York.

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

Haller, S. K. (1997). *Adopting Probability Curricula: The Content and Pedagogical Content Knowledge of Middle Grades Teachers.* Unpublished doctoral dissertation, University of Minnesota.

Hodges J. L & Lehmann E.L. (2005). *Basic Concepts of Probability and Statistics, 2nd Edition*.

Society for Industrial and Applied Mathematics. Philadelphia.

Jefferson Hane Weaver (2003). *The Math Explorer: A journey Through the Beauty of Mathematics*. Prometheus Books, 59 John Glenn Drive, Amherst, New York.

John Haigh (2002). *Probability Models*, Springer Undergraduate Mathematics Series. Athenaeum Press Ltd, Gateshead, Tyne & Wear. (pp 3 – 4)

Jones, G. A. (2005). Reflections. In G. A. Jones (Ed.), Exploring probability in school: Challenges for teaching and learning (pp. 375-379). Netherlands: Kluwer Academic Publishers.

LeCoutre, V. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 55 7-568.

Lee, H. S., Rider, R. L., & Tarr, J. E. (2005). *Making connections between empirical data and theoretical probability: Student's generation and analysis of data in a technological environment.* Manuscript submitted for publication.

Louisiana Comprehensive Curriculum (Revised 2008).

Lui, Y. (2005). *Teachers' understanding of probability and statistical inference and their implications for professional development.* Unpublished doctoral dissertation, Vanderbilt University.

Moore, David S. (1995). *The Basic Practice of Statistics.* New York: W. H. Freeman and Company.

National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards*. Reston, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.

Pierre Simon, Marquis De Laplace (1995). *A Philosophical Essay on Probabilities* (with an Introductory Note by E.T. Bell), Dover Publications, INC. New York.

Ronald Staszkow, Robert Bradshaw, (2005). The Mathematical Palette (pp 176 -181), 3rd Edition, Annotated Instructor's Edition. Thompson Learning Inc.

Shaughnessy, J. M. (1981). Misconceptions of probability: From systematic errors to systematic experiments and decisions. In A. P. Shulte& J. R. Smart (Eds.), *Teaching statistics and probability* (pp. 90-100). Reston, Va.: National Council of Teachers of Mathematics.

Shaughnessy, J. M. (1992). Research in 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.

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

Stohl, H. (2005). Probability in teacher education and development. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 345-366). Netherlands: Kluwer Academic Publishers.

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

Watson, J.M. (2005). The probabilistic reasoning of middle school students. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 145-169). New York: Springer.

Wilson, S. M., Shulman, L. S., & Richert, A. E. (1987). '150 different ways' of knowing: Representations of knowledge in teaching. In J. Calderhead (Ed.), *Exploring teacher's thinking* (pp. 104-124). Eastbourne, England: Cassell.

APPENDIX: COMMON CORE STATE STANDARDS FOR MATHEMATICS (CCSSM)

Seventh Grade Statistics &Probability

Use random sampling to draw inferences about a population.

1. Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.
2. Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions. For example, estimate the mean word length in a book by randomly sampling words from the book; predict the winner of a school election based on randomly sampled survey data. Gauge how far off the estimate or prediction might be.

Draw informal comparative inferences about two populations.

3. Informally assess the degree of visual overlap of two numerical data distributions with similar variability, measuring the difference between the centers by expressing it as a multiple of a measure of variability. For example, the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable.

4. Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book.

Investigate chance processes and develop, use, and evaluate probability models.

5. Understand that the probability of a chance event is a number between 0 and 1 that expresses the likelihood of the event occurring. Larger numbers indicate greater likelihood. A probability near 0 indicates an unlikely event, a probability around 1/2 indicates an event that is neither unlikely nor likely, and a probability near 1 indicates a likely event.

6. Approximate the probability of a chance event by collecting data on the chance process that produces it and observing its long-run relative frequency, and predict the approximate relative frequency given the probability. For example, when rolling a number cube 600 times, predict that a 3 or 6 would be rolled roughly 200 times, but probably not exactly 200 times.
7. Develop a probability model and use it to find probabilities of events. Compare probabilities from a model to observed frequencies; if the agreement is not good, explain possible sources of the discrepancy.

    a. Develop a uniform probability model by assigning equal probability to all outcomes, and use the model to determine probabilities of events. For example, if a student is selected at random from a class, find the probability that Jane will be selected and the probability that a girl will be selected.

b. Develop a probability model (which may not be uniform) by observing frequencies in data generated from a chance process. For example, find the approximate probability that a spinning penny will land heads up or that a tossed paper cup will land open-end down. Do the outcomes for the spinning penny appear to be equally likely based on the observed frequencies?

8. Find probabilities of compound events using organized lists, tables, tree diagrams, and simulation.

   a. Understand that, just as with simple events, the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs.
   b. Represent sample spaces for compound events using methods such as organized lists, tables and tree diagrams. For an event described in everyday language (e.g., “rolling double sixes”), identify the outcomes in the sample space which compose the event.
   c. Design and use a simulation to generate frequencies for compound events. For example, use random digits as a simulation tool to approximate the answer to the question: If 40% of donors have type A blood, what is the probability that it will take at least 4 donors to find one with type A blood?

VITA

David Nganga Njenga was born in April 1971 to Geoffrey Njenga Nganga and Rahab Wanjiru Njenga, in Makutano, Rift Valley Province, Kenya. He attended Igure Primary School through eighth grade and graduated from Kabarnet High School in 1989. He attended Moi University in Eldoret, Kenya, from October 1990 through 1995, where he graduated with a Bachelor of Education in mathematics and geography in October of the same year.

After graduation, Njenga taught mathematics and geography in grades nine through twelve for four years at Kipsangui High School. In 1999 he moved to Kapsiliot High School where he taught the same subjects. While in Kapsiliot, Njenga got married to Fidelis Nganga in August of 1999. In June 2000, the couple was blessed with their first born, Rodney Nganga, followed by Cephas Nganga in 2002. Before their daughter, Joy Nganga was born in 2005, Njenga taught mathematics in Burnt Forest, Kenya, at Rukuini High School.

In March 2007, the family moved to Baton Rouge, Louisiana, where Njenga started his present teaching career at Staring Education Center in August 2007. He teaches 7th grade mathematics. He has also been tutoring 10th grade geometry at Dunham School's McKay Academic Center (MAC) since September 2007.

Njenga was admitted into graduate school at Louisiana State University in 2008 as part of National Sciences cohort. He plans to obtain his degree in December 2010.