ABSTRACT

Title of Document: THE IMPACT OF AN INSTRUCTIONAL INTERVENTION DESIGNED TO SUPPORT DEVELOPMENT OF STOCHASTIC UNDERSTANDING OF PROBABILITY DISTRIBUTION

Darcy Lynn Conant, Doctor of Philosophy, 2013

Directed By: Dr. Patricia Campbell
Teaching and Learning, Policy and Leadership

Stochastic understanding of probability distribution undergirds development of conceptual connections between probability and statistics and supports development of a principled understanding of statistical inference. This study investigated the impact of an instructional course intervention designed to support development of stochastic understanding of probability distribution. Instructional supports consisted of supplemental lab assignments comprised of anticipatory tasks designed to engage students in coordinating thinking about complementary probabilistic and statistical notions. These tasks utilized dynamic software simulations to elicit stochastic conceptions and to support development of conceptual connections between empirical distributions and theoretical probability distribution models along a hypothetical learning trajectory undergirding stochastic understanding of probability distribution.

The study employed a treatment-control design, using a mix of quantitative and qualitative research methods to examine students' understanding after a one-semester course. Participants were 184 undergraduate students enrolled in a lecture/recitation,

calculus-based, introductory probability and statistics course who completed lab assignments addressing either calculus review (control) or stochastic conceptions of probability distribution (treatment). Data sources consisted of a student background survey, a conceptual assessment, ARTIST assessment items, and final course examinations. Student interviews provided insight into the nature of students' reasoning and facilitated examination of validity of the stochastic conceptual assessment.

Logistic regression analysis revealed completion of supplemental assignments designed to undergird development of stochastic conceptions had a statistically significant impact on students' understanding of probability distribution. Students who held stochastic conceptions indicated integrated reasoning related to probability, variability, and distribution and presented images which support a principled understanding of statistical inference.

THE IMPACT OF AN INSTRUCTIONAL INTERVENTION DESIGNED TO SUPPORT DEVELOPMENT OF STOCHASTIC UNDERSTANDING OF PROBABILITY DISTRIBUTION

By

Darcy Lynn Conant

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Patricia F. Campbell, Chair
Professor and Associate Dean Paul J. Smith
Professor Robert G. Croninger
Professor Lawrence M. Clark
Professor Patricia A. Alexander

Acknowledgements

I am very grateful for so many people at the University of Maryland, College Park, who supported me in this endeavor. First of all, I am thankful for the opportunity afforded me by the Mid-Atlantic Center for Mathematics Education to pursue full-time doctoral studies. I am thankful for friends and colleagues from the College of Education who challenged me to think about teaching and learning in new ways and to pursue excellence. I am also thankful for colleagues in the Department of Mathematics, both faculty and staff, who have offered unending encouragement and support.

More specifically, I would like to thank the director of this dissertation, Dr. Patricia Campbell, for her assistance and guidance. I am especially grateful to have had such a wonderful person as my dissertation advisor; someone who was always available to listen and offer encouragement and advice. From initial conversations concerning the design of this study to analyzing and summarizing the results, I learned so much as a result of her valuable expertise and guidance. Dr. Campbell, I am very thankful for everything you did to help me reach this goal; words cannot express my deep gratitude for everything you have done.

The members of my committee supported me throughout this process and I am very grateful for each of you. I would like to thank Dr. Patricia Alexander, who pushed my thinking about cognition and learning. Many of my ideas related to thinking and learning were formed as a result of those challenges and conversations in her doctoral seminar courses. I studied probability and statistics under the direction of Dr. Paul Smith and thank him for all the time he spent with me. Dr. Smith not only helped me to think deeply about these content domains, but also about the college-level curriculum and quantitative analyses. I thank Dr. Robert Croninger for his time and advice regarding quantitative analyses, in addition to his valued expertise and advice offered in mixed research methodologies. I thank Dr. Lawrence Clark for his encouragement and assistance with design of the supplemental lab assignments.

Finally, completion of this dissertation was truly a family endeavor. Anyone who is intimately acquainted with my family knows that there was a time when this endeavor was inconceivable. All the tangible and intangible things ways that my husband and sons assisted me throughout this journey will never be forgotten. Words cannot express how grateful I am for the love and support I received from you. My family knows that completion of this dissertation is a testament to the truth that with God all things are possible.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

The ability to evaluate quantitative information properly is viewed as an important skill, and the study of probability and statistics provides conceptual tools for individuals to deal with quantitative information intelligently. In recent years, discussions regarding statistical literacy, reasoning, and thinking have been prominent in educational, mathematical, and statistical communities (Ben-Zvi & Garfield, 2004). Increased attention has been given to the teaching and learning of probability, data analysis, and statistics at all levels of education.

Large numbers of university students study probability and statistics, and enrollment in introductory statistics courses is increasing, however, what students understand after completing a college-level course varies (Artigue, Batanero, & Kent, 2007; Moore & Cobb, 2000; Shaughnessy, 2007). Investigations of college-level students' understandings show many students exhibit difficulties in learning and applying probabilistic and statistical concepts (Artigue, et al., 2007; Garfield & Ben-Zvi, 2007, 2008; Jones, Langrall, & Mooney, 2007; Shaughnessy, 1992; Zieffler, et al., 2008). These students also struggle with probabilistic and statistical reasoning. Research has repeatedly revealed students' inconsistencies in thinking about concepts in probability and statistics, and this inappropriate reasoning is widespread and persistent across all age levels (Hirsch, L., & O'Donnell, 2001; Metz, 1998). While the desired result of all introductory probability and statistics courses is that students have the ability to think statistically (Garfield et al., 2005), research investigating college-level students' statistical reasoning shows that many students continue to exhibit difficulty applying statistical concepts after instruction (Artigue, et al., 2007; Shaughnessy, Garfield, &

Greer, 1996). Research reveals that it is difficult for students to reason about variability and to recognize the different aspects of variability involved in statistical thinking (Garfield & Ben-Zvi, 2005; Garfield, delMas, & Chance, 2007; Konold & Pollatsek, 2002). Research also shows that after a first course in probability and statistics, many college-level students are able to perform the procedures involved in making a statistical inference, but the majority of students neither understand the reasoning involved in making a statistical inference nor the reasoning required for interpretation of the results (Batanero, Tauber, & Sanchez, 2004; Meletiou-Mavrotheris & Lee, 2002; Reaburn, 2011; Smith, 2008; Thompson, Liu, & Saldhana, 2007).

Background and Rationale

The study of probability and statistics focuses on real-world phenomena that involve uncertainty (Jones, et al., 2007). Notions of probability entail dualistic aspects (Hacking, 1975). One aspect of probability is associated with a degree of belief and the other aspect is associated with the tendency of some chance devices to produce relatively stable frequencies in the long run. Thus, one aspect involves assessing reasonable degrees of belief in propositions and the other aspect encompasses stochastic considerations of processes involving chance. Assessing degree of belief is associated with a subjective view of probability, whereas stochastic considerations are connected to notions of randomness and are associated with classical view of probability and a frequentist view of probability (Batanero, Henry, & Parzysz, 2005). The subjective view of probability underscores the idea that considering the possibility of an event is related to some system of knowledge and not necessarily the same for all people. Bayesian approaches to probability arise from the subjective view and utilize a priori information to determine the probability of an event. Stochastic considerations of probability are statistical in nature and involve connections between a classical view (i.e. theoretical probability) and a frequentist view involving stabilization of long run frequencies of repeated trials (i.e. experimental probability).

A main goal of an introductory college-level probability and statistics course is for students to understand the basic ideas of statistical inference and make appropriate use of statistical inference (Garfield et al., 2005). Statistical inference is multifaceted and involves coordination of probabilistic and statistical thinking. Data analysis techniques focus on describing and representing data and gleaning information about specific collections of data. Statistical inference involves taking a specific collection of data and using information from that collection to make a generalization about the population from which the data was drawn. In order to make a generalization, one must consider the nature of random variability impacting selection of the data, random variability within the data collection, and random variability of possible sample selections.

Statistical methods arise from the omnipresence of variability in data (Cobb & Moore, 1997). Statisticians recognize the role of random variability in involved in making a statistical inference: anticipation of variability when formulating a question, acknowledging variability when collecting data, accounting for variability when analyzing data, and attending to the nature of variability when interpreting the results and looking beyond the data. When making a formal statistical inference, probability is used to modal random variability resulting from underlying processes inherent to the situation at hand.

Coordination of probabilistic and statistical thinking requires making conceptual connections between probability and statistics, which are grounded in the notion of randomness. Outcomes of a stochastic process are synonymous with the outcomes observed as a result of a random process. Furthermore, random sampling can be considered a stochastic process. Thompson et al. (2007) argue, "… to conceive of sampling as a stochastic process is key of all statistical inference" (p. 208). However, understanding sampling as a stochastic process is challenging. Understanding stochastic processes is difficult for students because this notion is fundamentally related to randomness and research shows that understanding notions of randomness is problematic for people of all ages (Batanero & Serrano, 1999; Falk & Konold, 1997; Metz, 1998). Making a statistical inference requires use of probability models, which model outcomes of stochastic processes, and coordination of probabilistic and statistical thinking. This presents a huge conceptual hurdle for students in a first course in probability and statistics (Pfannkuch, 2005).

Jones et al. (2007) describe the complementary role probability has in informing statistics and quantifying chance phenomena: "Statistics uses characteristics of random processes and probability models of such processes to make inferences about problems involving data; … probability focuses on directly describing, quantifying, modeling and illuminating random processes" (p. 910). Research has revealed the critical underpinnings that stochastic reasoning supplies to coordination of probabilistic and statistical thinking. Liu and Thompson (2007) found that stochastic conceptions of probability supported an understanding of statistical inference. Thompson et al. (2007) found high-school mathematics teachers' difficulties in understanding and employing statistical inference was due in part to their compartmentalized knowledge of probability and of statistical inference.

This analysis points to the need for research on how to facilitate learners' conceptual connections between probability and statistics. "Research in students informal and formal inferential reasoning would suggest that there are huge gaps in current knowledge about how best to enable learners to make the connection between probability and statistics" (Pfannkuch, 2005, p. 268). Furthermore, research addressing students' understanding of the connections between frequency distributions of empirical data and theoretical probability and the potential for technology tools to support students' understanding of probability concepts is needed (Jones, 2005; Jones, et al., 2007). When summarizing research on students' understanding of probability, Shaughnessy (2003) states:

> Although this chapter focuses on probability, I point out that a separation of research discussions of probability and statistics is artificial, just as artificial as the separation of data and chance when teaching. … I believe the most interesting research questions for the future reside in the joint realm of the areas of probability and statistics, just as the most interesting teaching challenges for the future lie in making interconnections between these two areas (p. 216).

Shaughnessy (2003) recommends an instructional approach that builds on students' primary intuitions and emphasizes notions which support understanding related to probability distribution models. Shaughnessy also recommends making connections between probability and statistics through use of examples and questions that are both statistical and probabilistic as well as connecting two big ideas in stochastics, sample space in probability and the nature of variation. Shaughnessy also advocates introducing probability through data, thus starting with statistics to get to probability. Finally, Shaughnessy suggests a problem-solving approach to learning probability that gives students opportunities to investigate probability problems or chance situations on their own.

Stochastic Conception of Probability

Stochastic reasoning is vital to understanding connections between probability and statistics and is a critical aspect of a conceptual frame connecting notions within probability and statistics. A stochastic conception of probability requires understanding the stochastic nature of random phenomena, understanding how probability is used to model random phenomena, and understanding how probability models are used to make formal statistical inferences.

Stochastic reasoning is grounded in conceptual connections between probability and statistics. One reason many people experience difficulty with stochastic reasoning is that learning about random experiments through simulation or experimentation is not connected to learning about combinatorics and representations in probability (Batanero, Godino, & Roa, 2004). Research has found evidence of disconnections between individuals' intuitive thinking about probability based on experiences with random generators and formal mathematical thinking about probability (Abrahamson, 2007, 2009b, 2009c). Research also shows that the complementary nature of empirical probability and theoretical probability is not salient to learners and that their probabilistic thinking can be compartmentalized (Batanero et al, 2004). This compartmentalization of probabilistic thinking may be related to students' conception of probability and stochastic reasoning. Research shows that individuals exhibit differing perceptions of probability depending on the context of a probability situation and these differing perceptions result in different ways of understanding the problem at hand, as well as different solution approaches to the problem (Liu & Thompson, 2007). This research revealed that a probability situation may be interpreted stochastically or nonstochastically. Other research shows inconsistencies between students' thinking about probability in school versus out of school (Rubel, 2007). Hence, not only the problem context of a given probability situation, but also how and where the problem solving is situated may influence perceptions of probability and approaches to solving the problem.

Many conceptual connections between probability and statistics are connected to stochastic notions of probability. However, development of stochastic reasoning is wrought with challenges. Heitele (1975) advocated for instruction aimed at the development of stochastic reasoning to focus fundamental ideas. He argued that stochastic notions are rarely made concrete to the learner and stressed the importance of connecting instruction in stochastics to students' intuitive experiences: "A large number of paradoxes in stochastics which can be confusing even for experts, show that intuitive pre-establishment is more urgent in stochastics than anywhere else," (Heitele, 1975, p. 189). Hietele took a view of learning aligned with Bruner (1960) that stressed development of foundational ideas and understanding of fundamental concepts. Hietele suggested that instruction in stochastics should provide learners with *explanatory mental models*. At an elementary level, these explanatory mental models connect to intuitive experiences and promote development of normative conceptions along with deeper understanding of stochastic notions. Hietele gives an example of an explanatory mental model for the notion of random variable. At an elementary level, conceptions are developed through playful activities with two dice. An individual intuitively assigns a better chance to obtaining a sum of dots equal to seven rather than two. This model can be developed to arrive at a quantitative model involving outcomes and sample space. A still more elaborate model would consist of interpreting the sums of dots as a stochastic (i.e. random) variable and an image of a probability distribution model. Hence, the concept of probability distribution is an arguably critical concept involved in development of stochastic reasoning.

A stochastic conception of probability and a stochastic understanding of probability distribution are important to development of a deep understanding of connections between probability and statistics. Furthermore, stochastic reasoning is vital to probabilistic and statistical thinking and to development of principled understanding (Greeno, 1978) in both content domains. These ideas point to the need for research regarding development of students' stochastic understanding of probability distribution.

The Importance of Probability Distribution

Conceptual connections between probability and statistics are grounded in an understanding of probability distribution models (Inzunsa, 2008; Wilensky, 1997). A stochastic conception of probability distribution includes understandings that are important to development of conceptual connections between probability and statistics which undergird an understanding of statistical inference. Because the notion of probability distribution affords development of critical connections between probability and statistics, the concept of probability distribution can be considered a key developmental understanding (Simon, 2006) in both the probability domain and the statistical domain.

The concept of probability distribution can be a powerful springboard for development of stochastic reasoning and learning probability in ways that facilitate making deep conceptual connections around probabilistic understandings related to variability, notions of independence versus dependence, notions of sample space, and notions of distribution (Liu & Thompson, 2007). Understanding probability distributions can facilitate thinking about probability in relation to outcomes of stochastic processes rather than simply static values assigned to likelihood. Furthermore, conceptual understanding of probability distribution potentially impacts stochastic reasoning and the ways learners think about the relationship between probability and statistics. A stochastic consideration of probability distributions may help learners think about relationships between experimental outcomes of chance experiments and theoretical probability.

The context of probability distribution facilitates thinking about a probability situation as an expression of stochastic process. A stochastic process is a repeatable process and a random process. Repeating the process will produce a collection of outcomes. Although any given outcome is unpredictable in a stochastic process, the long-run patterned nature of the process is evident. Because of this long-run tendency, general predictions can be made with regards to the behavior of the stochastic process, and outcomes of the process can be mathematically modeled. Understanding the nature of the variability in individual outcomes, variability related to a collection of outcomes, and understanding the long run stability inherent to a stochastic process are important to stochastic reasoning and developing an understanding of connections between probability and statistics. A stochastic conception of probability is related to thinking about probability in terms of a distribution of outcomes.

Liu and Thompson (2007) found that a stochastic conception of probability supports thinking about statistical inference. Making a statistical inference comprises using probabilistic models to make inferences about a population based on sample data. Statistics involves analyzing data, and inferential analyses about data necessarily involve a model. In addition, statistics uses random processes and probability models to make inferences about data. Probability describes, quantifies, models, and illuminates random processes. Conceptions of probability and conceptions of statistics are intertwined in the construct of statistical inference. Making a statistical inference requires a comparison of a data-based understanding of reality with a theoretical stochastic-based probability distribution model.

The concept of probability distribution is foundational to understanding probability as a model used for making statistical inferences. Statistical models developed through analysis of data are framed by an understanding of underlying theoretical probability models. Probability distribution models describe and model the nature of random variability observed in real phenomena. The concept of probability distribution is central to an understanding of how random events can be both unpredictable and modeled. Understanding probability distributions as models is a notion which is important in the domain of probability and the domain of statistics.

Research on Post-Calculus Students' Understanding of Probability

Although not many studies have been conducted with post-calculus students, there is research showing these students exhibit difficulties in understanding probability distribution and its related concepts. Several studies have determined that after instruction in probability, many post-calculus students demonstrated merely instrumental understanding (Skemp, 1976) and presented probabilistic notions that were not aligned with formal probabilistic concepts (Barragues, Guisasola, &Morais, 2007; Batanero, et al, 2004; Giulianno, et al., 2006). This research also showed that although post-calculus students could use formulas and algorithms to work through problems involving probability and statistics, they did not demonstrate an understanding of probabilistic and statistical concepts.

In a study involving 75 post-calculus students in an introductory probability and statistics course, after instruction the vast majority exhibited poor understandings of random phenomena (Barragues, et al., 2006). These students also presented misconceptions of random sequences, insensitivity to sample size, and a deterministic bias. In addition, many of these post-calculus students revealed probabilistic thinking indicative of the same heuristical biases evident in individuals of all ages with lesser knowledge of mathematics (Metz, 1998). Other research indicates that post-calculus students, who were either currently enrolled in or had recently completed an introductory probability and statistics course, demonstrated evidence of probabilistic thinking that was aligned with novice thinking evidenced by high school students and college-level students in algebra-based introductory probability and statistics classes (Abrahamson, 2007; Abrahamson & Wilensky, 2007; Barragues, et al., 2007; Hernandez, Heurta, & Batanero, 2006; Ives, 2007; Lunsford, Rowell, & Goodson-Espy, 2006; Wilensky, 1997).

Research investigating undergraduate post-calculus students' conceptions of probability distribution found that learners exhibited difficulty understanding probability models and struggled to discriminate between empirical distributions and theoretical distributions (Abrahamson, 2007; Batanero, et al., 2004; Lunsford, et al, 2006, Wilensky, 1997). Noll and Shaughnessy (2012) found that despite their strong statistical and mathematical knowledge of theoretical probability distributions, graduate teaching assistants experienced difficulty resolving differences between theoretical models and empirical distributions. The teaching assistants also had difficulty explaining conceptual ideas of probability. There is further evidence that undergraduate post-calculus students have difficulty coordinating notions of random variable and sample space (Hernandez, et al., 2006). Prior research indicates that post-calculus students were comfortable with mathematical procedures and had mastered algorithmic techniques to use probability distributions. However, these students appeared to lack stochastic conceptions of probability and a deep conceptual understanding of probability distribution.

A study of post-calculus, engineering students' conceptions of probability revealed that conventional teaching can have a poor effect on students' probabilistic reasoning (Barragues, et al., 2007). In spite of the research-based evidence of students' difficulties in understanding probability distribution and its related concepts, some research shows that particular kinds of instruction can have a positive impact on students' understandings. Although not conducted in a classroom, the work of Abrahamson (2007) indicates that post-calculus learners can consolidate their intuitive notions of probability with their formal mathematical knowledge in the context of probability distribution. Abrahamson (2007) found that individuals were able to coordinate their thinking about relationships between empirical distributions and theoretical distributions as a result of engaging with interactive models in a computer environment.

Batanero and Diaz (2007) suggest, "A genuine knowledge of probability can only be achieved through the study of some formal probability theory and the acquisition of such theory should be gradual and supported by the students' stochastic experience" (p. 124). Virtual environments offer simulations tools which can support stochastic experience. Research related to instruction in probability and statistics points to the promise of learners' engagement in tasks utilizing a computer-based, dynamic statistical environment as a means towards facilitating development of notions of sampling distribution, variability, and inferential reasoning (Meletiou-Mavrotheris, 2003; Sanchez & Inzunsa, 2006). Prodromou (2012) found that specific software tools enabled students to operationalize variation and subsequently to coordinate a data-centric perspective of distribution with a modeling perspective of distribution. Instructional supports in a virtual environment helped students connect theoretical probability to simulated phenomena.

Learning with Understanding

A synthesis of research concerning how people learn revealed that students at all levels come to classrooms with preconceptions (Bransford, Brown, & Cocking, 2000). This research also found that effective teaching elicits students' pre-existing understandings and builds on that understanding. If students' pre-existing understandings are not engaged in their learning experiences, learners tend to compartmentalize the new knowledge and subsequently revert to their original preconceptions when encountering problem situations outside of the classroom. An implication of these findings is that those who aim to teach for understanding must recognize that students bring a plethora of understandings to a learning situation and elicit those pre-existing understandings. Effective teaching for understanding in probability and statistics begins with instruction that draws on students' intuitive probabilistic and statistical understandings and builds on these notions to support development of formal understanding in probability and statistics.

Bransford et al. (2000) found that in order to develop competence in a content domain, students need a deep foundation of factual knowledge, an understanding of facts and knowledge within the context of a conceptual framework, and the ability to organize knowledge in ways that facilitate retrieval and application. These findings imply that supporting learners' movement toward more formal understanding in probability and statistics involves development of foundational understandings of big ideas in these content domains as well as development of a conceptual frame for connecting these big ideas. Formal understanding in probability and statistics also requires a conceptual frame that includes connections between probability and statistics.

Teaching for understanding in probability and statistics in ways which undergird development of these aspects of competence described by Bransford, Brown, and Cocking (2000) should aim to support development of principled knowledge (Greeno, 1978; Spillane, 2000). Greeno (1978) described *principled knowledge* as mathematical knowledge that includes understanding of mathematical skills and procedures integrated with understanding of the ideas and concepts which support mathematical procedures. Thus, principled knowledge in probability and statistics refers to understanding the ideas and concepts which support mathematical and statistical procedures, as well as understanding the connections between and within probability and statistics. In order to support development of principled knowledge, learning supports in probability and statistics should not focus on algorithms, formulas, and memorization. Rather, learning supports should focus on development of conceptual connections between and within probability and statistics with an aim of deepening those connections.

Curriculum and Instruction in Probability and Statistics

The intended curriculum of introductory college-level probability and statistics courses typically culminates with instruction in formal statistical inference. Making formal statistical inferences requires application of advanced stochastic thinking for correct interpretation (Batanero, 2006), which includes making conceptual connections between probability and statistics and a stochastic understanding of probability distribution. Probability distribution is an important curricular topic in probability and statistics because understanding statistical inference requires a stochastic conception of probability distribution. A stochastic understanding of probability distribution is essential to development of deep understandings in probability as well as development of formal understanding of statistical inference. Thus, instruction in probability and statistics should support development of stochastic reasoning and lay foundations for formal inferential reasoning.

Along with the development of technological tools, statistical practice and instruction in probability has changed in recent years. In a report published by the Mathematical Association of America, Cobb (1992) recommended that instruction in probability and statistics should emphasize statistical thinking, emphasize data and concepts as opposed to theory and calculations, and foster active learning. Moore (1997) argued for curricula reforms focused on more data analysis and less probability, pedagogical reforms focused on fewer lectures and more active learning, and technological reforms focused on data analysis and simulations. In 2005, the American Statistical Association endorsed curricular guidelines for assessment and instruction in college-level courses, as well as for K – 12 education (Franklin, et al., 2005; Garfield, et al., 2005). As a result, reformed-based curricula, which favor a data analysis approach with less probability, have been developed and are widely used, while traditional curricula, which favor more mathematics and probability, also remain in use. At the present time, there are generally two different approaches to instruction in probability and statistics at the college level: one favors a data analysis approach to instruction and the other favors a more traditional, mathematical approach to instruction.

There is widespread dissatisfaction with a traditional, mathematically oriented approach to teaching probability and statistics, which emphasizes teaching of formulas for calculating statistics and lacks interpretive activities and simulations (Artigue, et al., 2007). On the other hand, there is also dissatisfaction with a data analysis approach to instruction that tends to mitigate probability and omits many topics in probability (Biehler, 1994; Meletiou-Mavrotheris, 2007). Research indicates that despite recent curricular reforms, at the end of an introductory course in probability and statistics, most college-level students demonstrate an instrumental understanding of statistics inference and struggle with understanding key concepts related to statistical inference.

Recommendations for a data analysis approach to learning statistics include a lack of emphasis on probability. The reform curriculum tends to include one chapter addressing formal probability, and this is the only consideration of probability in the curriculum. Reform curricula focus on data and distributions of data and aim for students to develop an understanding of statistics based on data analysis. However, these curricula do not support development of conceptual connections between probability and statistics. A serious concern with the data analysis approach to learning statistics is that it requires students to make a major conceptual leap from analyzing data sets to considering the data as random samples from population. Biehler (1994) argued that both positions, an instructional approach aimed at purely formal probabilistic conceptions and an instructional approach aimed at a probability-free conception of data analysis, present obstacles to learning probability and statistics.

Meletiou-Marvotheris (2007) argues that students' persistent difficulties with stochastic reasoning may be the result of traditional mathematics curriculum that leads to compartmentalization of knowledge and fails to communicate the interconnectedness of probabilistic and statistical notions. Developing an understanding of probability is challenging for most learners, and there is substantial research documenting evidence of persistent probabilistic misconceptions from childhood through adulthood, misconceptions that are resistant to change (Artigue, et al., 2007; Batanero, et al., 2005; Jones, et al., 2007; Metz, 1998; Shaughnessy, 1992). Research addressing instruction in probability for young learners advocates supporting development of connections between data and chance, and this research could inform an instructional approach that is viable for adult learners as well (Abrahamson, 2007, 2009a, 2009b Konold & Kazak, 2008). Research indicates that building notions of chance which are connected to data can help children develop connected understandings of probability and variability (Konold & Kazak, 2008; Shaughnessy, Cianetta, & Best, 2004; Shaughnessy & Cianetta, 2002). These studies showed that by building on intuitive understandings of chance, and then analyzing the data collected from probability experiments, young learners were able to perceive the empirical law of large numbers and to begin building notions of probability as a distribution.

Research involving high school and college-level students has demonstrated that particular instructional models which include students' utilization of dynamic software simulations can support development of conceptual connections between empirical distributions of data and probability distribution models (Biehler & Prommel, 2010; Budgett, Pfannkuch, Regan, & Wild, 2012; Maxara & Biehler, 2006; Prodromou, 2012; Wild, Pfannkuch, Regan, & Horton, 2010). The implication is that a curricular approach, which is neither purely formal nor probability-free but engages the learner in connecting intuitive probabilistic notions with distributions of data and with formal ideas and theorems, could promote stochastic reasoning and foster deeper understanding of conceptual connections between probability and statistics.

Statement of Purpose

The purpose of this study was to investigate the impact of an instructional intervention designed to support development of stochastic reasoning in the context of probability distribution. Probability distribution encompasses foundational stochastic ideas and is a key developmental understanding in both statistics and probability; a stochastic conception of probability distribution affords critical connections between probability and statistics. As described earlier in this chapter, research investigating individuals' understandings related to the notion of probability distribution indicates that many learners experience difficulties thinking and reasoning about stochastic events. This research also points to challenges learners experience conceptualizing and applying the concept of probability distribution.

The analysis presented in this chapter suggests the need for a study which investigates the impact of instruction on students' understanding of probability distribution. The analysis also points to the need to develop an instructional model which promotes the development of stochastic reasoning in the context of probability distribution and to design instruction validating a hypothetical learning trajectory. Very limited research exists addressing probabilistic conceptions held by students who have strong backgrounds in mathematics, specifically those who have completed calculus. Among these studies, only a few investigated students' stochastic conceptions of probability and only a few involved participants who had earned college-level credits for calculus. This study aims to contribute to knowledge of post-calculus students' stochastic understandings of probability distribution after instruction in a one-semester, calculus-based, introductory probability and statistics course.

Theoretical Perspective

The theoretical perspective for this study frames a view of how understanding of probability distribution develops in a learning environment and provides a hypothetical model of connections between student learning and development of understanding of probability distribution (see Figure 1.1). Central to development of understanding is a viewpoint of how students' learn. This viewpoint draws on constructivist and situated perspectives and focuses on individual understanding that is built through learning experiences, which are impacted by the learner, the teacher(s), the instructional material, and peers. Learners' understandings are developed throughout the process of learning. This learning process is dependent on particular kinds of reasoning. In the case of probability, stochastic reasoning is crucial for development of understandings, which ground probabilistic and statistical thinking that is essential to understanding statistical inference. Notions of stochastic reasoning are built on previous experiences with stochastic processes, which also inform development of current stochastic reasoning. As notions of stochastic reasoning develop, learning results in the formation of probabilistic and statistical concepts. Thus, reasoning about the stochastic informs an understanding of probability distribution, and understanding of probability distribution informs stochastic reasoning. This research seeks to measure and describe students' understandings of probability distribution, which result from instruction designed to support development of stochastic conceptions of probability.

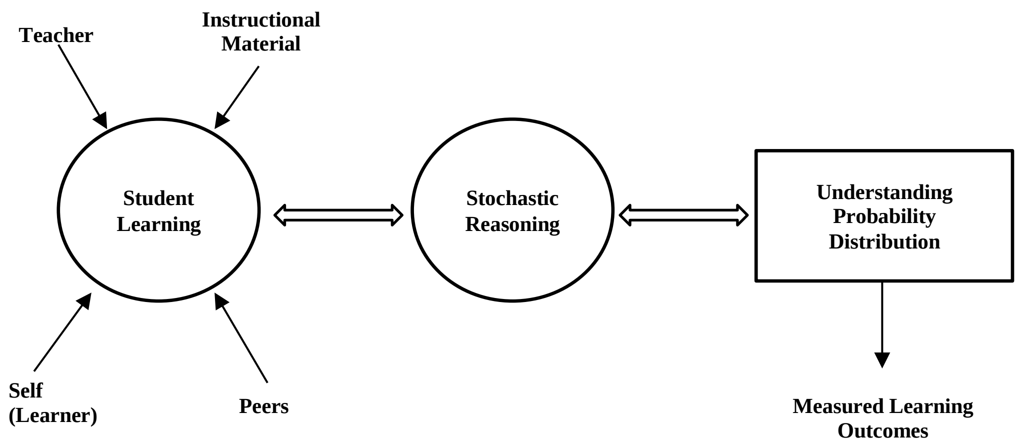


Figure 1.1. Model of theoretical framework for development of stochastic understanding of probability distribution.

Student Learning

Learning experiences shape the development of understanding. This study draws on two perspectives of learning, a constructivist view and a situated view. A situated view of learning informs knowledge of how learners interact with the learning environment and offers implications for design of instructional interventions. Student learning is influenced by engagement with instructional material, and this study sought to impact student learning via supplemental tasks designed to support development of stochastic understanding of probability distribution. Research investigating development of post-calculus students' understanding of probabilistic concepts indicates that teaching is an important factor related to students' understandings of probability (Barragues, Guisasola, & Morais, 2006; Sanchez & Inzunsa, 2006). Teaching, which emphasizes procedures tends to result in instrumental understanding. Teaching which facilitates learner explorations of conceptual notions of probability as a distribution and its connection to mathematical the theorems offers opportunities for students to build relational understanding (Skemp, 1976) in probability and statistics.

A constructivist view of learning informed the development of student understanding for purposes of this study. Student learning is viewed as an active and recursive process whereby understanding is developed as a result of previous understanding, and new understandings are built from current understanding (Martin, 2008; Piaget, 2001; Pirie & Kieren, 1994). In this study, student learning was examined from an individual viewpoint and was understood to be impacted through interactions with self, interaction with peers, and interactions with instructional materials, and interactions with teachers. A *hypothetical learning trajectory* (Simon, 1995) for understanding probability distribution was adapted from Liu and Thompson's work (2007). The learning goal for the hypothetical learning trajectory was a stochastic understanding of probability distribution.

Development of Stochastic Reasoning

Development of stochastic reasoning is an important component for understanding probability distribution because stochastic reasoning is grounded in conceptual connections between probability and statistics. This study assumes that development of stochastic reasoning will impact learners' understandings of probability distribution. Stochastic reasoning provides conceptual tools for individuals to deal intelligently with random phenomena. To reason stochastically means conceiving of an observed outcome as but one expression of an underlying repeatable process that will produce a stable distribution of outcomes in the long run (Liu & Thompson, 2007).

Heitele (1975) suggested that instruction aimed at the development of stochastic reasoning should focus on fundamental stochastic ideas rooted in intuitions of random phenomena. Heitele maintained that stochastic notions link conceptions of empirical distributions with theoretical probability distributions, and it is necessary to distinguish and coordinate thinking about an empirical law of large numbers from a purely mathematical law of large numbers. These ideas are in agreement with those purported by Steinbring (2006) who proposed that developing stochastic conceptions of probability are manifested by structural connections that arise from gradual changes in stochastic meaning, resulting from interplay between an empirical object/reference context and a formal mathematical sign/symbol. In this characterization, the empirical object/referent context consists of "real" situations (i.e., chance experiments, sampling, simulations) which result in an empirical distribution of relative frequencies and the formal mathematical sign/symbol refers to probability distribution models (see Figure 1.2). The meaning of a stochastic conception of probability cannot be conceived by only considering of one of the vertices (i.e. only empirical distributions or only theoretical distributions), but requires a balance among all vertices of the epistemological triangle. The implication of this characterization is that development of stochastic reasoning requires coordination of thinking about empirical probability and theoretical probability, as well as coordination of thinking about relationships between empirical objects and formal mathematical signs, which should be developed simultaneously. Thus, a stochastic conception of probability distribution includes coordination of thinking about empirical probability distributions and theoretical probability distributions.

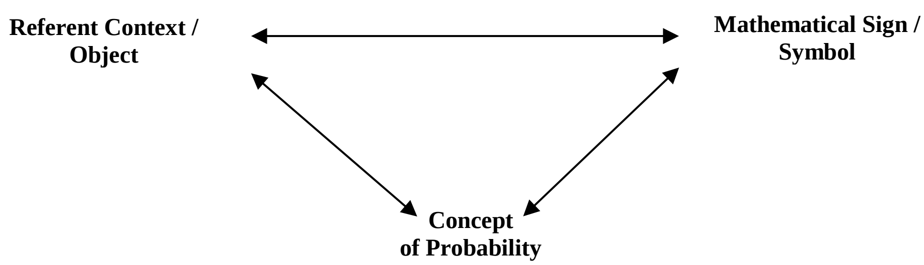


Figure 1.2. Epistemological triangle illustrating a stochastic conception of probability (adapted from Steinbring, 2006).

Research confirms that development of stochastic conceptions involves coordination of relational thinking about empirical probability and theoretical probability. Liu and Thompson (2007) investigated probabilistic understandings of eight secondary mathematics teachers and found that stochastic conceptions were developed through a series of ways of thinking that include: (1) conceiving of an underlying repeatable process, (2) understanding the conditions and implementations of this process in such a way that it produces a collection of variable outcomes, and (3) imagining a distribution of outcomes that are developed from repeating this process. This research illustrates that developing notions of probability as a distribution are central to building stochastic conceptions.

Understanding Probability Distribution

Understanding probability distributions can be difficult for students because this construct is a complex mathematical entity. Probability distributions are mathematical models of results of random processes. At one level of the model, outcomes of random phenomena are modeled by random variables. The next level of the model is comprised of random variables, which are functionally mapped onto the probabilities associated with the outcomes of the random variable. Random variables are the independent variables, and probability is the dependent variable in this functional relationship. The function represents probabilities associated with a distribution of outcomes of the random variable. These outcomes may be discrete or continuous.

Three overarching constructs frame an understanding of probability distribution: probability, variability, and distribution (Table 1.1). Notions of probability, which are important to an understanding of probability distribution, are understandings of sample space, independence, random variable, coordination of empirical and theoretical probability, and models for inference. Important notions and theorems related to variability, which connect to an understanding of probability distribution, are randomness, the law of large numbers, unit-to-unit variability, sampling variability, and the central limit theorem. Distributions of random variables, parameterization of distribution models, distributions of sample data, population distribution, and sampling distribution are notions of distribution that are associated to understanding of probability distribution.

Table 1.1

Framework for Understanding of Probability Distribution

Probability Distribution		
Probability	Variability	Distribution
Coordination of empirical and theoretical probability	Randomness and random variability	Distribution of random variable
Random variable	Law of large numbers	Parameterization of distribution model
Sample space	Unit-to-unit variability	Distribution of sample versus population distribution
Independence versus dependence	Sampling variability	Sampling distribution
Model for inference	Variability of sample statistics and the central limit theorem	

Understanding probability distribution means understanding connections between the constructs of probability, variability, and distribution, as well as understanding connections among notions within each construct and across the constructs. This means that understanding variability is not exclusive of understanding probability. Furthermore, understanding variability involves an understanding of randomness and random variability that is connected to the law of large numbers, unit-to-unit variability in a set of data/values, sampling variability, and variability of sample statistics. In addition, a deep understanding of randomness involves understandings associated with the notions found within the constructs of probability and distribution, and this deep understanding is fostered by connected conceptions of probability, variability, and distribution. Deep conceptual connections, such as those described by this theoretical framework for understanding random variable, are cultivated through development of stochastic reasoning. Thus, development of stochastic reasoning builds understandings about probability distribution.

Liu and Thompson's (2007) research about the development of stochastic conceptions of probability informs a hypothetical learning trajectory (Simon, 1995) for developing stochastic understandings of probability distribution:

1. Conceiving of a probability situation as stochastic process that has an underlying in repeatable process: developed through notions of randomness, the law of large numbers, and unit-to-unit variability of random phenomena;
2. Understanding the conditions of a stochastic process: developed through notions of randomness and independence versus dependence;
3. Understanding implementations of a stochastic process and anticipating that repeating a stochastic process would produce a collection of outcomes: developed through notions of sample space, random variable and sampling variability;
4. Imagining a distribution of outcomes that are developed from repeating a stochastic process: developed through notions of distribution of random variable and of coordination of empirical and theoretical distribution;
5. Conceiving of probability distribution as a model: developed through parameterization of distribution, distribution of data, population distribution, and sampling distribution; and through the central limit theorem, and a model for inference.

This hypothetical learning trajectory anticipates that learners will connect their intuitive notions of randomness and variability to more formal conceptions of these ideas. Each phase of the learning trajectory connects to development of stochastic conceptions, which were framed by Liu and Thompson (2007), and aims to move the learner toward a modeling perspective of probability in various contexts of probability distribution. The hypothetical learning trajectory builds on an intuitive knowledge base and progresses towards building connections of more formal mathematical understandings of probability distribution. This research-based trajectory informed the design of instructional tasks for supporting students' learning with understanding. These tasks were aimed at development of stochastic reasoning in the context of probability distribution.

This hypothetical learning trajectory was based on a small-scale study, which the researchers described as *highly exploratory* (Liu &Thompson, 2007, p. 157). The sample size in Liu and Thompson's (2007) study does not support making broad claims related to emergent understanding of probability. The implications are that a larger study involving more mathematically-advanced learners is needed to illuminate understandings of probability in the general population of mathematically-advanced learners, and to inform a hypothetical learning trajectory for developing stochastic understanding of probability distribution.

Research Questions

Given the evidence that many students in probability and statistics are not learning with understanding, additional research is needed to investigate students' understandings resulting from instruction designed to support development of principled knowledge. A quasi-experimental study with comparison groups was designed to investigate the impact of instruction on post-calculus students' understandings of probability distribution. The study employed mixed methodologies to investigate the impact of an instructional intervention aimed at supporting a stochastic understanding of probability distribution. This study focused on investigating students' understandings of probability distribution as a result of an instructional intervention in a control-treatment design.

The aim of this study was to contribute to current knowledge of college students' understandings of probability by addressing the following research question: What is the impact of an instructional intervention designed to support the development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory, calculus-based, probability and statistics course? The study also addressed four research sub-questions. Qualitative research methodologies were used to address research sub-question 1 and three components of this question. Quantitative research methodologies were used to address research questions 2, 3, and 4.

1. What is the nature of students' reasoning when confronted with a probability situation?
    a) How do students characterize a probability situation in terms of an image of a repeatable process?
    b) How do students characterize a probability situation in terms of specification of conditions of a repeatable process?
    c) How do students characterize a probability situation in terms of an image of a distribution of outcomes?
2. Does instruction designed to support development of stochastic understanding of probability distribution impact students' stochastic conceptions of a probability situation as evidenced on a conceptual assessment?

3. Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding of confidence intervals as measured by the ARTIST assessment?
4. Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding as evidenced on final course examinations administered in an introductory, calculus-based, probability and statistics course?

Significance

The study aims to contribute to knowledge of post-calculus students' understanding of probability distribution after instruction in a one-semester, introductory, calculus-based probability and statistics course and may inform mathematics and statistics educators as to the probabilistic conceptions held by students who have earned college-level credits for two semesters of calculus and have strong backgrounds in mathematics. While the field has some evidence of post-calculus students' understanding of probability distribution, only one study of 117 students majoring in pedagogy, psychology, and economics was conducted in conjunction with course instruction (Batanero, et al., 2004). Other studies involving post-calculus students were either case studies (Abrahamson, 2009b; Hernandez, et al., 2006; Wilensky, 1997) or studies involving only a small number of students, which took place outside of the classroom (Abrahamson, 2007; Inzunsa, 2008).

The study may also provide insights into characteristics of students' understanding resulting from the instructional intervention designed to support development of stochastic reasoning. The study could also provide evidence related to the impact of this kind of instructional intervention on students' understanding. Prior research suggests that tasks designed to foster the development of stochastic reasoning in the context of probability distribution could potentially promote students' thinking about connections between probability and statistics.

The study potentially offers important contributions to knowledge about how best to foster the development of students' stochastic understandings of probability distribution in the context of instruction. This study aims to utilize a hypothetical learning trajectory positing development of a stochastic understanding of probability distribution. Because statistical inference is the capstone of a typical college introductory course in probability and statistics and statistical inference builds on a stochastic conception of probability, a learning trajectory that intentionally supports development of stochastic reasoning appears to be crucial to the development of probabilistic understandings that ground statistical inference. This study could inform the usefulness of the proposed hypothetical learning trajectory and provide validity for a framework to support development of stochastic understanding of probability distribution.

Additional research is needed to inform educators about how best to approach teaching for understanding in probability and statistics in ways that build on students' pre-existing understandings and support development of principled knowledge of concepts and connections between and within probability and statistics. Research, which informs statistics educators about students' understandings of probability distribution, could help instructors understand how best to help learners make connections between probabilistic and statistical concepts. This study may inform development of an instructional model which promotes the development of stochastic reasoning in the context of probability distribution. As a result, this investigation of students' understanding of probability distribution in the context of instruction aimed at the development of stochastic reasoning may also provide valuable evidence about effective curricular approaches for learning probability and statistics with understanding. Knowledge of how instruction aimed at supporting and fostering students' stochastic understandings of probability distribution impacts that understanding could inform development of curriculum and instruction in probability and statistics.

Overall Design of the Study

The study employed a treatment-control design to investigate the impact of an instructional intervention designed to support students' stochastic understanding of probability. The study involved 184 students enrolled in two large-lecture classes of a calculus-based, introductory probability and statistics course at a large, public university. Participants in the study were students who had completed at least two semesters of college-level calculus and were majoring in computer science, engineering, mathematics, economics, and other scientific fields. Based on prior research it was assumed that prior to instruction, students' thinking would likely be more aligned with statistical novices than experts (Abrahamson, 2007; Hernandez, et al., 2006; Lunsford, et al., 2006; Wilensky, 1997). The study extended research investigating the impact of learning supports (Abrahamson & Wilensky, 2007) on college-level students' understanding of probability distribution in an interview setting (Abrahamson, 2007, 2009c) to a large-scale investigation of the impact of learning supports in an actual course setting.

All students enrolled in the calculus-based, introductory probability and statistics course received supplemental lab assignments. These assignments were completed outside of regular classroom instructional time. For purposes of this study, students were assigned to either a treatment or control group. Students in the control group received lab assignments that consisted of a review of calculus content which students would be re-encountering in this course. Students in the treatment group received technology-supported lab assignments designed to support development of stochastic reasoning. These learning supports were designed as anticipatory tasks (Simon, 2013) to elicit stochastic conceptions of probability along a hypothetical learning trajectory aimed at development of a stochastic understanding of probability distribution. The aim of the anticipatory tasks was to promote stochastic conceptions of probability and impact students' stochastic understanding of probability distribution. These supplemental lab assignments utilized bridging tools (Abrahamson, 2007, 2009a, 2009c) to facilitate development of connections between tacit and formal mathematical knowledge related to stochastic conceptions of probability distribution. In this study, students engaged in the anticipatory tasks (supplemental lab assignments) prior to class-based instruction in probability theory. The aim of these tasks was development of anticipations related to stochastic conceptions of probability, thus preparing students to learn from their lecture and discussion classes (Schwarz & Bransford, 1998).

Both groups completed written work for their respective lab assignments, and this written work was due prior to receiving lectures on the specific course topic related to lab assignment tasks. After completing the written portion of each lab assignment, students in both the treatment group and the control group received explicit instruction via video lessons connecting knowledge that their respective lab assignment was designed to initiate with course specific content related to the lab assignment. For students in the treatment group, this explicit instruction connected foundational stochastic conceptions related to probability distribution with topics in probability theory that students would encounter in the course. Students in the control group received explicit instruction which connected calculus-review practice problems with calculus content students would be using in this particular introductory probability and statistics course.

The instructional intervention for the treatment group consisted of supplemental lab assignments aimed at the development of stochastic reasoning in the context of probability distribution. The design of lab assignment tasks for this intervention was based on a hypothetical learning trajectory of students' stochastic conceptions of probability developed by Liu and Thompson (2007) which was adapted by the researcher for use in the context of probability distribution. These lab assignments required students to create and analyze virtual simulations which were designed to elicit students' prior understanding of probability and then deepen and extend that understanding. Lab assignment tasks required students to consider juxtaposed constructs relevant to understanding probability distribution, such as theoretical versus empirical probability and independent versus dependent events (Abrahamson & Wilensky, 2007; Batanero, Biehler, Maxara, Engel, & Vogl, 2005). The aim of the stochastic reasoning lab assignments was to have the learner decompose domain constructs into idea components and then use conceptual bridging tools to recompose the constructs using their intuitive and analytic resources.

This instructional intervention was designed to impact development of students' stochastic conceptions. Lab assignments for the treatment group were designed to engage students in contrasting and analyzing probabilistic and statistical notions connected the development of stochastic reasoning. The instructional intervention was designed to prepare students to learn from the lectures and therefore to provide greater opportunity for students to make deeper conceptual connections and to develop knowledge of probability distribution (Schwartz & Bransford, 1998; Schwartz & Martin, 2004). During the latter part of the course, all students' understandings were measured by a conceptual assessment of their stochastic understanding, an assessment of their understanding of confidence intervals, and the final course examination.

Limitations of the Study

This study has five main limitations. First, this study was conducted in a mathematics department at a large public university in conjunction with traditional lecture-recitation instruction, limiting the generalizability of these findings to other instructional settings. A second limitation is the duration of the supplemental probabilistic and statistical learning experiences in this study. Changes in understanding may take years to develop rather than weeks or months. As such, it would be impossible to assess the full impact of an instructional intervention which occurred during the course of a one-semester class (Dunbar, Fugelsang, & Stein, 2007). Student engagement with the lab assignments could be another limiting factor. Students may or may not have perceived the lab assignments as being worthwhile because lab assignment questions differed from the type of questions typically asked on homework or examinations. All lab assignments were completed outside of regular class time, but students in the treatment group were required to go to an on-campus computer lab to complete the stochastic reasoning labs, while completion of the calculus review labs did not require students to go to a special location. The fourth limitation of this study related to implementation was the role of peer interaction. Some students may have worked on the lab assignment individually, while others may have worked on the lab assignments in small, self-selected groups. Finally, beliefs are related to understanding and can be very difficult to change, and prior research indicates these beliefs may be of particular concern in the domain of probability (Fischbein, 1997; Greer, 2001).

Definitions of Terms

Explanatory mental model. An explanatory mental model is a developmentally appropriate model which connects intuitive notions and experiences to deeper conceptual understanding and supports development of more integrated, normative conceptions (Heitele, 1975).

Heuristic. A heuristic is mental shortcut which eases the cognitive load and aids problem solving (Kahneman, Slovic, and Tversky, 1982).

Inferential reasoning. Inferential reasoning supports' thinking about statistical inference and involves coordination of probabilistic and statistical thinking which includes images of data, distributions of data, randomness and random variability, probability distribution models, sampling and sampling variability, and sampling distribution models, as well as notions of the law of large numbers and the central limit theorem (Pfannkuch & Wild, 2012).

Key developmental understanding. A key developmental understanding is a conceptual advance that changes one's ability to think about and/or perceive particular mathematical relationships; and without completing the developmental process, the learner lacks a particular mathematical ability (Simon, 2006).

Principled understanding. Principled understanding includes understanding of mathematical skills and procedures integrated with understanding of the ideas and concepts which support mathematical procedures (Greeno, 1978).

Stochastic conception. Stochastic conceptions involve: (1) understanding the nature of random phenomena, (2) understanding how probability is used to model random phenomena, and (3) understanding connections between probability and statistics. (Steinbring, 1991).

Stochastic process. A stochastic process is a repeatable process that produces a collection of outcomes, and although any given outcome is unpredictable, the long-run patterned nature of the process is evident.

Stochastic reasoning. Stochastic reasoning involves conceiving of the probability of an outcome in relation to the process that produced that outcome and coordination of intuitive notions of chance with formal mathematical knowledge of probability (Steinbring, 1991).

CHAPTER 2: BACKGROUND THEORY AND LITERATURE REVIEW

This chapter provides a review of literature and the theoretical background for the study. The first section presents a conceptual analysis of stochastic reasoning. It describes conceptions of stochastic processes, stochastic conceptions of probability, and development of stochastic reasoning. The second section presents analyses of three conceptual components comprising an understanding of probability distribution and a review of literature addressing understanding for each of these conceptual components. This section begins with an overview of the historical development of mathematical models and notions related to probability distribution. The historical overview is followed by subsections which present analyses of aspects of conceptual components related to understanding probability distribution: probability, variability, and distribution. This subsection is includes a review of literature addressing understanding for each of the three conceptual components. The final subsection included under the section addressing understanding of probability distribution provides a review of literature addressing inferential reasoning. The third section of Chapter 2 presents an analysis and review of literature addressing the theoretical perspective of learning used for this study. The final section addresses use of technology in learning probability and statistics.

Conceptual Analysis of Stochastic Reasoning

Stochastic reasoning is grounded in the conceptual connections between probability, statistics and data, and offers conceptual tools for individuals to deal intelligently with quantitative information. The foundation of stochastic reasoning involves consideration of both probabilistic and statistical notions. Statistics uses random processes and probability models to make inferences about data, and probability describes, quantifies, models and illuminates random processes (Jones, et al., 2007). Hacking (1975) states the statistical side of probability concerns itself with “stochastic laws of chance processes” (p. 12). Statistical notions of probability are exemplified by experiments in which long run frequencies stabilize on repeated trials. Outcomes from these experiments result in data, which are produced by a random or stochastic process. Thus, stochastic reasoning means conceiving of an observed outcome as but one expression of an underlying repeatable process that will produce a stable distribution of outcomes over the long run (Liu & Thompson, 2007).

Stochastic reasoning is an important notion which undergirds a principled understanding of probability distribution that incorporates consideration of both empirical and theoretical distributions. Stochastic reasoning includes stochastic conceptions of probability that are characterized by the coordination and linking of intuitive notions of chance with formal, mathematical knowledge of probability (Steinbring, 1991). A stochastic conception of probability means conceiving of the complementary relationship between empirical probability, which results from real-world observations, and theoretical probability, which is an abstract mathematical model.

Statistical inference is grounded in a stochastic conception of probability (Liu & Thompson, 2007). The observed statistic, which results from data, is interpreted in light of a probabilistic model, and an inference is made based on a probabilistic statement. Thus, engaging in statistical inference involves understanding of probability distribution as a model and requires making conceptual connections between probability, statistics and data.

Conception of Stochastic Process

A stochastic process is random process that may be described in terms of probability. The outcomes of a stochastic process are non-deterministic and may be modeled by random variables. The distribution of a random variable is a probability distribution; thus, a stochastic process can be mathematically modeled as a function of a random variable. The probability distribution is an abstract model of observed phenomena.

An observed phenomenon is a result of reality, is concrete, and can result in data. Thus, although stochastic process is an abstract notion, concrete data may result from this random process, and the data may be collected and analyzed. These data are complex and involve aspects, which are both stochastic and patterned (Metz, 1998). Empirical data, such as these, may be modeled using probability or analyzed using statistical techniques. Inferences may be made about the phenomenon, and probabilistic models are used to make formal statistical inferences about this stochastic situation. Hence, thinking about a stochastic process involves interplay between data, probability, and statistics.

Stochastic Conception of Probability

A stochastic conception of probability provides a foundation for stochastic reasoning and is a critical affordance for making connections between probability and statistics. Thus, a stochastic conception of probability can be considered to be a *key developmental understanding* in probability and statistics (Simon, 2006). An individual who holds a stochastic conception of probability exhibits the ability to perceive the complementary relationship between empirical probability (which results from real-world observations) and theoretical probability (which is an abstract mathematical model). This person is also able to discern stochastic notions present in statistical problems, which makes these problems understandable and solvable.

A stochastic conception of probability does not develop through merely hearing explanations, but is built through multiple experiences and is a result of active reflection on those experiences. As *a key developmental understanding*, a stochastic conception of probability transforms the way one understands situations involving probability and random events. Such a conception is essential to understanding probability. "A genuine knowledge of probability can only be achieved through the study of some formal probability theory; however, the acquisition of such formal probability theory by the students should be gradual and supported by their stochastic experience" (Batanero, et al., 2005).

Liu and Thompson (2007) describe a stochastic conception of probability as a notion that entails conceiving of the probability of an outcome in relation to the process that produced that outcome. Similarly, a stochastic conception of an event is conceived as but one expression of an underlying repeatable process. Liu and Thompson (2007) investigated probabilistic understandings of eight secondary mathematics teachers and found that a stochastic conception is developed through a series of ways of thinking that include: (1) conceiving of an underlying repeatable process, (2) understanding the conditions and implementations of this process in such a way that it produces a collection of variable outcomes, and (3) imagining a distribution of outcomes that are developed from repeating this process. This research revealed that a stochastic conception of probability incorporates five aspects:

1. Conceiving of a probability situation as the expression of a stochastic process;

2. Taking for granted that the process could be repeated under essentially similar conditions;
3. Taking for granted that conditions and implementation of the process would differ among repetitions in small, yet perhaps important ways;
4. Anticipating that repeating the process would produce a collection of outcomes;
5. Anticipating that the relative frequency of outcomes will have a stable distribution in the long run (Liu & Thompson, 2007, pg. 122).

These researchers also found that a stochastic conception of probability supported thinking about formal statistical inference.

Development of Stochastic Reasoning

One reason people may experience difficulty with stochastic reasoning is because learning about random experiments through simulation, or experimentation, is not connected to formal mathematical learning about probability and to representations in probability (Batanero, et al., 2004). Research indicates there are disconnections between an individual's intuitive thinking about probability that is based on experience with random generators and formal mathematical thinking about probability (Abrahamson, 2007, 2009c, 2009d). For example, Rubel (2007) observed these disconnections were evident in high school students' thinking about probability in school versus out of school. The complementary nature of empirical (frequentist) probability and classical (theoretical) probability was not salient to learners, and as a result, their probabilistic thinking was compartmentalized. Stochastic thinking involves coordination of empirical and classical perspectives of probability, thus compartmentalization of probabilistic thinking could be related to difficulty in thinking about the random processes (Batanero, et al., 2004).

Making statistical inferences requires application of stochastic thinking in order to correctly interpret data in light of probabilistic models. However, many people with little formal training in mathematics or stochastics are using statistical tools and engaging in statistical analysis (Batanero, 2006). This statement implies that instruction in probability and statistics aimed at helping learners develop conceptual connections that will lay foundations for inferential reasoning is crucial. Having stochastic conceptions, which foreground statistical inference, means the learner will understand probability statistically and understand statistics probabilistically (Liu & Thompson, 2009). This reciprocity of thinking is critical to understanding formal statistical inference. Instruction should support learning aimed at developing connected notions of a stochastic conception of probability and statistics in the context of data.

Stochastic reasoning involves both probabilistic and statistical thinking, but topics in probability and statistics are typically compartmentalized in the curriculum. Teaching should support learning, which conceptualizes the complementary relationship between probability and statistics. Batanero and Diaz (2007, p. 124) state, “A genuine knowledge of probability can only be achieved through the study of some formal probability theory, and the acquisition of such theory should be gradual and supported by students’ stochastic experience.” Thus, a primary goal of statistics education should be to support stochastic conceptions of probability, which are also conceptions that support thinking about statistical inference (Liu & Thompson, 2007).

Understanding Probability Distribution

An understanding of probability distribution includes understanding along three important and interrelated conceptual components: probability, distribution, and variability. Because probability distribution comprises a key developmental understanding in the domain of probability and the domain of statistics, each of the three conceptual components includes aspects of probability which support thinking about probability distribution across these domains. Stochastic conceptions of probability distribution include coordinated images of probability, distribution, and variability which undergird understanding statistical inference.

Historical Development of Probability Distribution

Historically, early notions of probability distribution emerged through study of observations of data. Early mathematicians struggled with ideas regarding coordination of empirical observations, mathematical representation, and formalization of models for probability. The origin of probability theory is connected to the binomial probability distribution.

The work of Jacob Bernoulli is regarded as the beginning of the mathematical theory of probability (Hacking, 1975). Bernoulli's interest in games of chance involving equally likely outcomes stimulated him to attempt formal mathematization of empirical notions specifying that the accumulation of more evidence about an unknown proportion resulted in being closer to certain knowledge about the that proportion. Bernoulli asserted that even "the most stupid of men" (Stigler, 1986, p. 83) understand that uncertainty decreases as the number of observations increases. The result of Bernoulli's early work in this area is referred to as *Bernoulli's weak law of large numbers.* It was the first mathematical approach toward measurement of uncertainty.

Around 1721, De Moivre began efforts towards approximating terms of the binomial expansion. Building on Bernoulli's work, De Moivre found a function for a large sample approximation $P\left(X = {}^{n}\!/_{2}\right)$ where X was binomial distribution (symmetric with n trials, n even). Although De Moivre did not develop the concept of a density function, he viewed this exponential function as a *curve* that approximated the binomial distribution. De Moivre's approximation to the binomial distribution is taken to be the original appearance of the *normal curve* (Stigler, 1986, p. 76). In this work, De Moivre emphasized the law of large numbers rather than measurement of irregularity. His approach was mathematical, and he perceived that chance lay in the data, not in the underlying theoretical probabilities. However, during that period of time, the perspective of probability was only from cause to effect, which thwarted application of the binomial distribution as a model used for inference.

The idea of inversion of probability, which is probabilistic reasoning from effect to cause, was developed by Laplace. This idea provided ways of thinking about probability that could be applied to inference. Laplace took a more philosophical and analytic approach to thinking about probabilistic ideas. Laplace found the tools he needed for inference involving the binomial distribution and applied the notion of inverse probability to binomial situations (Stigler, 1986).

Other developing mathematical ideas connected to binomial distribution were rooted in interest about data. Quetelet was interested in fitting normal curves to data observations. He viewed the normal curve as the end product of a binomial mechanistic model. Quetelet was interested in a connection between the normal curve and sums of independent accidental causes, and this was an important link towards application of the normal model. Quetelet constructed a table that was based solely on the binomial distribution, with 1000 outcomes (p = 0.5, n = 999). Without using an analytic approximation to the normal integral, he found an approximation to the normal distribution.

In 1875, Galton devised the quincunx as a device for illustrating lectures. The quincunx was an analogue for binomial experiments. Shot was dropped through a glass funnel and cascaded through rows of equally spaced pins, which formed an array. At each level (row), the shot had an equal probability of falling left or right. The shot was collected in bin at the bottom of the devise. Dropping many shot resulted in a normal shaped curve at the bottom. Development of the quincunx resulted in an extraordinary visual representation of the binomial distribution and provided an analogue proof related to ideas about independence.

Understanding Probability

One of three major conceptual components of probability distribution is probability. The probability of an event is defined as a numerical measure of the likelihood that an outcome for a chance experiment with a finite sample space will occur. In a chance experiment, an event corresponds to a subset of the sample space, and a subset of the sample space consists of one or more outcomes. Thus, we state that an event A occurs if any one of the outcomes in A occurs. The probability of event A *is* denoted by $P(A)$. $P(A)$ is a numerical value that gives a measure of the chance that A will occur.

Assigning a value to measure probability involves a different kind of thinking than the logical or causal approach typical in other areas of mathematics (Kapadia & Borvcnik, 1991). Probability is multifaceted and consists of meanings which are dialectically and experimentally intertwined (Batanero & Diaz, 2007). Meanings of probability include: (1) a priori mathematical degree of uncertainty, (2) evidence supported by data, (3) a propensity, (4) a logical relation, (5) a personal belief, and (6) a mathematical model. These meanings can be subsumed under three views of probability: subjective, empirical, and theoretical. Subjective probability is often based on intuition or experience and reflects a measure of personal belief (Ross, 2006).

Empirical probability is based on evidence supported by data. It is formalized by estimation of the frequency of occurrence of events derived from relative frequencies. The relative frequency of an event *A* is defined as the number of times the event occurs divided by the number of repetitions of the experiment represented by $f_n(A) = \frac{n(A)}{n}$. A frequentist definition of probability is rooted in the law of large numbers, which describes the long run stability of the relative frequency. According to the law of large numbers, variations in the relative frequency with which event *A* occurs will fluctuate less as the experiment is repeated, and the limiting number to which the relative frequency converges as the number of repetitions increase is the probability of event *A*. The law of large numbers confirms intuitions that the probability of an event in a repeatable experiment can be estimated by the relative frequency of event *A*.

Theoretical probability is anchored in mathematical theory and can be defined via a classical approach or via axiomization. Probability for finite sample spaces can be determined using the classical approach, which is also called the Laplace model. The classical definition assumes all outcomes are equally likely. For example, if there are *k* outcomes in the sample space, then each outcome is assigned the probability of $\frac{1}{k}$ (Devore, 2004; Ross, 2006; Tijms, 2007). The classical definition of the probability of event *A* is the number of outcomes in event *A* divided by the number of outcomes in the sample space $(A) = \frac{m}{k}$, where *m* represents the number of outcomes in event *A*. The axiomatic approach to probability was not formalized until 1993 when Kolmogorov laid a satisfactory mathematical foundation of probability theory (Tijms, 2007). The axioms of probability hold for an experiment with a finite or countable infinite sample space, as well as for an experiment with an uncountable sample space.

For a finite sample space, probability is defined as a measure and this measurable function *P* assigns a numerical probability $P(A)$ to each subset *A* of the sample space. The probability measure *P* must satisfy the following conditions (axioms):

Axiom 1: $P(A) \geq 0$ for every event *A*

Axiom 2: $P(A) = 1$ when *A* is equal to the sample space

Axiom 3: $P(A \cup B) = P(A) + P(B)$ for disjoint events *A* and *B*

These axioms require that probability values can only range between 0 and 1. A sample space together with an assignment of probabilities to events in the sample space is called a probability space (Koralov, 2007; Tijms, 2007).

The concept of probability is multifaceted and related to ideas about variability and distribution. For the purpose of this study, five aspects of probability are envisioned as being important to an understanding of probability distribution for a post-calculus introductory probability and statistics course. These aspects are: coordination of empirical and theoretical probability, notions of random variable, notions of sample space, notions of independence versus dependence, and understanding the use of probability as a model for making statistical inferences. An understanding of probability distribution requires understanding these distinctions and coordinating thinking about empirical and theoretical probability.

Coordination of empirical and theoretical probability. Historical foundations of probability began with games of chance and gambling, and are grounded in the empirical dimension of probability. Evidence of this history is found in the use of random devices throughout the historical record, some of which date as early as 3500 B.C. These devices point to the prevalent nature of chance games throughout human history. However, it was not until the sixteenth century that the Italian physician and mathematician, Cardano, quantified chance outcomes based on the concept of outcomes of an experiment for cases when all of the outcomes are equiprobable (David, 1998; Hacking, 1975). This quantification of probability was fore grounded by empirical evidence based on observations of dice games, and Cardano's subsequent naming of the outcomes. The approach that Cardano used marks a great leap in the development of probability theory and exemplifies the importance of the complementarity of the dual natures of probability, empirical and theoretical.

A rich understanding of probability must acknowledge the complementarity between empirical and theoretical natures of probability and requires an acceptance that "probability is neither strictly an empirical nor mathematical property, but the combination of these aspects bound to context" (Kapadia & Borvcnik, 1991, p. 20). This kind of understanding includes a notion that a purely experimental approach to probability is not sufficient (Batanero & Diaz, 2007). Empirical probabilities are evaluated in light of probabilistic models, and these models are theoretical constructs. Additionally, the work of Piaget and Inhelder (1975) points to the importance of context, in addition to combination of the complementary empirical and theoretical aspects of probability, in order for individuals to development formal mental organization of probabilistic concepts.

Random Variable. Quite often, we are interested in describing, summarizing, modeling, and representing outcomes of chance experiments, and these endeavors are more readily accomplished with numerical values. A numerical value is determined from some characteristic pertaining to the outcome. For example, we may be interested in the number of heads obtained in tossing four coins, which is a function of the outcome of the experiment, rather than the actual outcome. Furthermore, the association of a number with a chance outcome enables the employment of a host of mathematical tools. The rule of association is called a random variable. It is a variable because different numerical values are possible, and it is random because the values depend upon the result of the experiment.

A random variable is defined as a mathematical function that maps (associates) each outcome onto the set of real numbers. For each outcome, there is only one value of the random variable. Thus, a random variable is a function, which assigns a numerical value to each outcome of the chance experiment (Devore, 2004; Korolov, 2007). The domain of this mathematical function is the sample space and the range is the set of real numbers. A random variable is also a mathematical model comprised of real numbers, which are dependent upon random outcomes. In this sense, a random variable models "reality" given by the random events.

Sample Space. A random variable is a function defined on the sample space of a chance experiment (Devore, 2004; Korolov, 2007). The sample space is a set of elements that corresponds with the set of all possible outcomes of the experiment (mathematically, this is a one-to-one correspondence). In the example of the experiment described as tossing four coins, the sample space consists of the union of the distinct outcomes, i.e. the set of all possible combinations of four elements, where each element is either heads or tails. Each outcome provides the input (domain) on which the random variable function operates. Thus, an understanding of sample space is intimately connected to the concept of random variable.

Piaget and Inhelder (1975) articulate the importance of sample space in addition to an image of the distribution of all possible outcomes for a chance experiment:

> Now with chance…, it is precisely a question of describing the sample space, that is, all the possibilities, so that, on the one hand, the distribution of the total number of cases becomes predictable and, on the other hand, each isolated case acquires a probability expressed as a fraction of the whole.(p. 228).

This research shows that a rich conception of sample space is related to a consideration of the distribution of possible outcomes, including the important notion of variability. Piaget and Inhelder (1975) claim that one important aspect of formal probabilistic thinking is the notion of the composite of individual outcomes coordinated with a whole that is defined as all possible outcomes. The whole is determinable by a system of multiple possibilities. A deep understanding of probability consists of judging isolated outcomes in comparison to the whole set of possible outcomes. Thus, conceptions related to outcomes of a random variable must be considered simultaneously with a conception of the distribution of the random variable. Further implications that ensue from the work of Piaget and Inhelder (1975) are: (1) sample space in an important construct in making probability judgments and assignment of probabilities, and (2) sample space is a composite concept of individual outcomes along with the distribution of possible outcomes must be considered simultaneously. Both of these implications are also related to notions of empirical probability and theoretical probability. Experimental outcomes of chance experiments result in frequencies, which may be transformed into relative frequencies describing empirical probability.

Independence versus dependence. Understanding probabilistic notions of independence versus dependence is important to understanding notions of randomness, sampling, random variable, and distributions of random variables. Many probability models assume trials are independent. Random sampling assumes that selection of any element is equally likely and independent, i.e. the selection in any given trial or occurrence does not depend on any other selection. The concept of independent random variable is very important to understanding probability distribution. Independent random variables are uncorrelated and this property undergirds many probability distribution models. A sample statistic can be considered a random variable and sampling distribution models are based assumptions of independence and random sampling.

Research addressing understanding of probability. There are differing standards in terms of elements of abstraction that constitute the meaning of understanding probability that are expected for individuals in various institutional settings or situations. All students who are undertaking formal study of concepts of probability need to distinguish between personal and institutional meanings of probability and appropriate these meanings in a given context (Batanero & Diaz, 2007). Students in this study were preparing for undergraduate degrees and subsequent scientific work, which requires knowledge of probability distribution and stochastic processes. Thus, during their study and when accessing their knowledge of probability in applied contexts, the students will need to integrate thinking about empirical distributions involving relative frequencies and empirical probabilities with thinking about theoretical probability and probability distribution models.

Personal meanings of probability arise from intuitions; studies show that intuitive probabilistic notions may interfere with the development of *normative* mathematical probabilistic notions (Fischbein, 1975). Fischbein (1987) defines probabilistic intuition as a cognition that appears subjectively as self-evident. Fischbein (1997) argues that one reason intuitive cognitive beliefs may conflict with reality is because experiences may be limited. Individuals tend to seek coherence for cognitive organization and tend to integrate information that is more easily available, ignoring information that requires a more sophisticated effort. Fischbein (1997) investigated the evolution of intuitively-based, probabilistic conceptions with age. Participants in this study were students in grades 5, 7, 9, 11, and college-level students who were preparing to specialize in teaching mathematics. Fischbein (1997) found that conceptions related to representativeness and the independence of outcomes of a stochastic process improved with the age of the student. However, other studies indicate that misconceptions related to representativeness are common among college-level students from a variety of majors (Barragues, et al., 2006, 2007; Batanero, et al., 1996; Hirsch & O'Donnell, 2001). Studies show that students demonstrate inconsistent reasoning across a variety of problems contexts, and misconceptions related to determining the probability of compound and simple events are common and stable across all age groups (Fischbein, 1997; Rubel, 2007). Fischbein (1997) reported that misconceptions related to the effect of sample size were prevalent across all age groups and these misconceptions became stronger with age for students "not trained in stochastics" (pg. 103). The findings from other studies confirm that students indicate insensitivity to sample size and a tendency towards belief in the law of small numbers (Barragues, et al., 2006, 2007). Similarly, other research indicates that post-calculus students, who were either currently enrolled in or had recently completed an introductory probability and statistics course, exhibited probabilistic thinking that was aligned with the level of novice thinking evidenced by high school students and college-level students in algebra-based introductory probability and statistics classes (Abrahamson, 2007; Abrahamson & Wilensky, 2007; Barragues, et al., 2007; Hernandez, et al., 2006; Ives, 2007; Lunsford, et al., 2006; Wilensky, 1997).

Studies demonstrate the benefits of supporting students' intuitive expectations in probability situations (Abrahamson, 2007, 2009c). Abrahamson (2009c) investigated the capacity of 24 self-selected, mathematically trained, undergraduate and graduate students (post-calculus) who were requested to model a simple binomial probability problem situation. Abrahamson (2009c) examined students' connections between combinatorial analysis and simulated probability experiments and was interested in whether these post-calculus students understood how the binomial function is grounded in combinatorial analysis. These students worked with innovative learning tools, which were designed by Abrahamson and Wilensky (2007) to develop perspectives of modeling probability. The researcher guided and observed students interactions with the tools so as to elicit participants' understanding (Abrahamson, 2007). The majority of participants alluded to formulas for computing expected frequencies and referred to mathematical concepts, such as limit. Of the 24 participants, 18 described the nature of probability experiments with examples, and 12 cited the law of large numbers. Only four participants correctly employed combinatorial analysis before explicit prompting, while another six participants used some form of combinatorial reasoning, but abandoned the procedure or could not correctly complete the procedure. The remaining students did not mentioned combinatorial analysis in any form. However, once students were prompted to construct the sample space, all students expressed the implication of the sample space for the experimental outcomes and indicated coordination of thinking about empirical and theoretical probability. Abrahamson (2007, p. 22) asserted, "… with their metacognitive and 'school-wise' development, students come to regard their intuitive/experimental and formal knowledge as separate resources, as though the former constitutes a possible impediment to 'really learning.'" Abrahamson (2007) argued that learning environments which allow students to experience the "power of modeling" should not be limited to middle school, but rather should be sustained through secondary and post-secondary education.

Shaughnessy and Cianetta (2002) assert that the concept of statistical variation is closely connected to the concept of sample space for a probability experiment. The concept of sample space is also connected to conceptions of random variable. Hernandez, et al. (2006) found post-calculus students exhibited predominantly deterministic thinking along with a tendency to focus on algebraic procedures and did not link mathematical conceptions of random variable to the problem context. These students demonstrated difficulties distinguishing between model and reality in a probability situation and did not accept random variable as an aspect of a probability model (Hernandez, et al., 2006). Other studies also indicate students' engagement with probability simulation tasks can promote development of conceptual connections between empirical and theoretical probability models (Meletiou-Mavrotheris, 2003; Konald & Kazak, 2008). Corter and Zahner (2007) found that many students enrolled in an introductory statistics course for graduate students spontaneously used visual representations to reorganize information and list outcomes while solving probability problems. Research indicates that appropriate tools can support development of an outcome-based notion of sample space and coordination of notions of empirical and theoretical probability (Abrahamson, 2009a, 2009c; Abrahamson & Wilensky, 2007; Kuntze, Engel, Martignon, & Gundlach, 2010; Lee, Angotti, & Tass, 2010).

Statisticians must consider probability distribution models when making inferences from data. However, research addressing post-calculus students' conceptions of probability distribution reveals that learners exhibit difficulty understanding probability models and discriminating between empirical distributions of data and theoretical probability distribution models (Abrahamson, 2007; Batanero, et al., 2004; Lunsford, et al., 2006, Wilensky, 1997). Noll and Shaughnessy (2012) found that despite their strong statistical and mathematical knowledge of theoretical probability distributions, graduate teaching assistants experienced difficulty resolving differences between theoretical models and empirical distributions. The teaching assistants also had difficulty explaining conceptual ideas of probability.

In general, research suggests that many post-calculus students indicate thinking that is not aligned with formal probabilistic concepts and lack a principled knowledge of probability (Artigue, et al., 2007). A number of studies have found that after instruction in probability, a large majority of post-calculus students may demonstrate proficiency with the use of formulas and procedures, but not demonstrate understanding probabilistic and statistical concepts (Barragues, et al., 2007; Batanero, et al., 2004; Giulianno, et al., 2006). However, other studies indicate that particular kinds of instruction may support development of conceptual connections between empirical distribution of data and theoretical probability distributions, conceptions that undergird a principled knowledge of probability (MacGillivray, 2006; Meletiou-Mavrotheris, 2003). This review of literature points to the need for a study investigating the impact of instruction designed to support development of stochastic understanding of probability distribution.

Understanding Variability

Conceptions of variability and variation are essential aspects of statistical reasoning (Moore, 1997; Wild & Pfannkuch, 1999). Reading and Shaughnessy (2004) define *variability* as the varying characteristic of an entity that is observable and define *variation* as a term used to measure variability or describe variability. Makar and Confrey (2004) argue that developing a conception of variation requires some understanding of distribution. Thus, notions of variability and variation are important aspects related to an understanding of probability distribution. Understanding sampling distributions, statistical inference, and *p*-values requires an understanding of statistical variation that involves conceptions of variability and of variation connected to a stochastic understanding of probability distribution (Chance, delMas, & Garfield, 2004; Saldhana & Thompson, 2002; Liu & Saldhana, 2007). Saldhana and Thompson (2002) argue that understanding ideas of sampling distribution and statistical inference require the learner to distinguish between variability among individuals in a sample, variability among individuals in the sampled population, and variability in statistics calculated from samples drawn from the population. A stochastic understanding of sampling variability undergirds understanding sampling distributions and statistical inference.

Probability distributions describe and model the variability of random variables. Graphical representation of probability distributions offer visual descriptions of variability and statistical measures define variation of the random variable being modeled by the distribution. For the purpose of this study, four aspects of variability are envisioned as being important to a stochastic understanding of probability distribution in the post-calculus introductory probability and statistics course. These aspects are: randomness and random variability, the law of large numbers, unit-to-unit variability, sampling variability, and variability of sample statistics and the central limit theorem.

Randomness and random variability. Konold and Pollatsek (2004) use the metaphor of "signal and noise" to illustrate statistical reasoning in relation to random variability and state, "The idea of noisy processes, and the signals that we can detect in them, is at the core of statistical reasoning" (p. 281). The notion of random variability requires an understanding of the nature of random phenomena and random processes. Without an understanding of randomness and probability, the formal study of statistics would have little meaning and any analysis of stochastic situations involving patterns and variability could very likely result in erroneous interpretations.

Randomness is a complex idea whose antithesis is regularity or a deterministic process. It is interconnected to concepts of independence and variation, all of which are foundational to predictability and uncertainty. Defining randomness is particularly difficult because of the number of contexts in which the term is used. In the informal domain, randomness can be thought of as chance; a situation or phenomena which is unpredictable and whose causes we do not know. The definition of randomness in the formal mathematical domain developed over a number of years. The first theoretical development emerged with probability calculus and was related to *equiprobability* in games of chances where the number of possibilities is finite. In the classical view of probability, an object is a random member of a given class if there is the same probability for any other member of its class. In practice, this definition of randomness is restricted since it only applies to a member of a finite class.

Random variability is evidenced in outcomes of a chance experiment; the outcome of a chance experiment is a random phenomenon. A random variable represents random phenomena and exhibits properties associated with randomness. An essential characteristic of random phenomena is that it involves a repeatable process. There is more than one possible outcome resulting from a repetition of the process, and the outcome is unpredictable. When considering random variability, a phenomenon is said to be random if we know what outcomes could result from a repeatable process, but we do not know (i.e., we cannot predetermine) which particular outcome will result for any given repetition of the process. Notions of randomness and random variability are aspects of variational disposition which includes anticipating variability in data and allowing for reasonable variability when using models.

Law of large numbers. Sequences of outcomes resulting from repetition of the process cannot be controlled or predicted. However, amidst this unpredictability and apparent disorder, there are observed global regularities. These global regularities are described by the law of large numbers. The empirical law of large numbers states: “The relative frequency with which event A occurs will fluctuate less and less as time goes on, and will approach a limiting value as the number of repetitions increases without bound" (Tijms, 2007, p. 19). In other words, the empirical law of large numbers means that as the number of repetitions of the process increases, the long-run relative frequency of observing a repeated outcome gets closer and closer to a single value. The theoretical law of large numbers can be derived from probability theory, and this law describes mathematically what is expressed by the empirical law of large numbers. The theoretical law of large numbers states:

> If a certain chance experiment is repeated an unlimited number of times under exactly the same conditions, and it the 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, 2007, p. 20).

The law of large numbers is powerful because it links empirical observations of outcomes of a random process to mathematical theory. This law provides a theoretical basis in probability for statistical analyses. Thus, the law of large numbers provides the link between empirical distributions and theoretical probability distribution models.

The law of large numbers says nothing about the outcome of a single repetition of the process or the outcome of a single experiment. Rather, the law of large numbers describes the long-run behavior of the system that involves an unlimited number of independent repetitions of the experiment/process. A common fallacy is thinking that sequences of random outcomes compensate in the short and a belief in the nonexistent *law of averages* (Kahneman, Slovic, & Tversky, 1982). There is no law of large numbers for short runs. Furthermore, the law of large numbers does not state the relative frequency of repetitions will equal a particular value after a certain number of repetitions and remain at that fixed value. The theoretical law of large numbers refers to convergence. In practicality, this means that in the long-run, observations of relative frequencies get closer to a *single value*, but this is not represented by an asymptotic relationship. As a result of the nature of random variability, the relative frequencies continue to fluctuate in the long-run, both above and below the *single value*. A variational disposition also requires understanding and applying the law of large numbers.

Unit-to-unit variability. Unit-to-unit variability refers to variability within collections of data or variability within a collection of outcomes. Conceiving of unit-to-unit variability includes attending to the clustering and spread of the units (i.e., data scores, or outcomes), as well as coordinating thinking about numerical measure that describe clustering and spread, such as mean, median, standard deviation, and variance. Describing and understanding the nature of unit-to-unit variability is important to understanding probability distribution. Unit-to-unit variability involves describing and measuring variability in data, and this notion corresponds to an element of variation identified by Peters (2011), variability in data for contextual variables. Within a data-centric perspective, this element of variation includes using summary measures for variability. Within a modeling perspective, this element of variation involves the ability to model data to explain variability. Coordinated thinking about statistical measures, as well as the meaning of these measures in a given context, empirical or theoretical, is essential to understanding probability distribution.

Sampling variability. Sampling variability refers to variability between collections of data (i.e., samples taken from a population). Sampling variability is an aspect that requires thinking about random variability and sampling, as well as aggregating features of variability in data to make comparisons by using summary measures to compare groups and examine variability within and among groups. Saldhana and Thompson (2002) describe a multiplicative conception of sample which entails images that support a deep understanding of statistical inference. A multiplicative conception of sample entails conceiving a relationship of quasi-proportionality between a sample and a population. This conception also entails an image of a collection of multiple samples resulting from repeated sampling. An image of repetition of the sampling process along with an image of variability among samples resulting from this process supports reasoning about distribution and sampling distribution.

Variability of sample statistics and the central limit theorem. The effects of sample size on variability is indicated and described by the central limit theorem and is one of four elements of variation related to a robust understanding of variation (Peters, 2011). This element of variation involves thinking about the effects of sample size on random variability and anticipation of the effects of sample size on the variability of sampling distributions. Within a data-centric perspective, these notions correspond to thinking about variability and relationships among data and variables. Within a modeling perspective, this element of variation involves modeling random variability of sample statistics, considering sampling distributions and standard error, and making inferences from data.

Research related to understanding variability. Reasoning about variability is multifaceted and connected to conceptions of distribution, center, sampling, and inference. Research indicates that it is difficult for students to reason about variability

(Garfield & Ben-Zvi, 2008), and we are just beginning to learn how reasoning about variability develops (Garfield & Ben-Zvi, 2005). Wild and Pfannkuch (1999) presented four aspects related to consideration of variation in statistics. These aspects were: (1) noticing and acknowledging variation, (2) measuring or modeling variation, (3) explaining or dealing with variation, and (4) developing investigative strategies in relation to variation. Reading and Shaughnessy (2004) added another aspect, namely describing and representing variation.

Peters (2011) found that a robust understanding of variation is indicated by an integrated reasoning within and across three perspectives for four elements of variation. Peters (2011) identified these four elements of variation as: (1) variational disposition, (2) variability in data for contextual variables, (3) variability in relationships among data and variables, and (4) effects of sample size. Each of these four elements coincides with aspects of variability important to understanding probability distribution. A variational disposition corresponds to anticipation of randomness and random variability and the law of large numbers. Variability in data for contextual variables corresponds to conceptions of unit-to-unit variability. Variability in relationships among data and variables, as well as effects of sample size correspond to conceptions of sampling variability and variability of sample statistics and the central limit theorem.

In addition to identifying four elements of variation, Peters (2011) also described three perspectives of variation that were important to a robust understanding of variation: a design perspective, a data-centric perspective, and a modeling perspective. Only the data-centric and modeling perspectives are relevant to this study. Peters (2011) also found that relational reasoning (Biggs & Collis, 1991) within a data-centric perspective of variation involves reasoning about variability within each distribution and between distributions of data. Relational reasoning within a data-centric perspective also involves anticipating variation and using measures of variation. Relational reasoning within a modeling perspective undergirds inferential reasoning for formal statistical inference and involves fitting models to distributions, which is indicative of a modeling perspective.

Research addressing conceptions of randomness and random phenomena resulting from random processes has consistently revealed that notions of randomness are difficult ideas to grasp (Artigue, et al., 2007; Jones, et al., 2007). Randomness is fundamental to stochastic processes yet it is an elusive concept for children and adults of all ages, and misconceptions about randomness and random phenomena persist in spite of age (Falk & Konold, 1997; Metz, 1999). Research addressing conceptions of randomness has shown that secondary students perceive local variability, lack of patterns, and unpredictability of random processes, but overemphasized unpredictability and luck when justifying attributions of randomness (Batanero & Serrano, 1999; Green, 1997). Rather than exhibiting normative conceptions of randomness, the students in these studies related luck and unknown causes to conceptions of probability. Research which involved high school students, college students, and other adults has documented that understanding randomness is related to individuals' perspectives of probability (Falk & Konold, 1997; Ives, 2007). However, studies also indicate that experiences with particular kinds of tasks, such as generating sequences of random outcomes and observing random phenomena, have the potential to change individuals' perspectives of probability and understanding (Ives, 2009; Pratt, Johnston-Wilder, Ainley, & Mason, 2007).

Research reveals that reasoning about variation is related to perceptions of center and clustering of data (Garfield, delMas, & Chance, 2007; Reading & Shaughnessy, 2004). Research also indicates that many students have merely superficial knowledge of terms, rules, and procedures related to numerical measures which describe data, and that these students lack a conceptual understanding of these measures (Bakker & Gravemeijer, 2004; Batanero, et al., 2004; Garfield, et al., 2007). Although the notion of central tendency is inseparable from variability, students tend to focus on the center of a distribution, only considering the mean or median and neglecting consideration of variation (delMas & Liu, 2005; Konold & Pollatsek, 2002). Peters (2011) found that a robust understanding of variation includes anticipating variability from a data-centric perspective and recognizing the need to characterize data using both center and variation. In a study investigating 12 college-level students' conceptual understanding of standard deviation, delMas and Liu (2005) found that most students used a ruled-based approach to compare variability across distributions. While many students tended to take a rule-based approach and did not indicate a fully coordinated understanding of variation, when these students engaged in tasks designed to promote development of a coordinated conception of variation about the mean, they moved from a one-dimensional understanding that did not consider variation about the mean to conceptions that coordinated the effects of frequency and deviation from the mean.

Research addressing students' conceptions of sampling and sampling variability reveals two different conceptions of sampling, namely additive versus multiplicative (Saldhana & Thompson, 2002). Saldhana and Thompson (2002) found that most students exhibited an additive conception of sample, which meant they tended to view a sample as simply a subset. Students lack experience thinking about samples in terms of a distribution of samples generated from a particular population and lack a notion of sampling variability (Rubin, Bruce, & Tenny, 1990). Other research shows that students tend to focus on individual data values and not perceive data as an aggregate (Hancock, Kaput, & Goldsmith, 1992; Konold & Higgins, 2003), which coincides with Saldhana and Thompson's (2002) findings related to an additive image of sample. Saldhana and Thompson (2002) found that a few students indicated a multiplicative conception of sample and sampling, and they argue this conception enabled these students to consider a sampling outcome compared to similar outcomes, which oriented them towards consideration of the relative unusualness of a sampling outcome. Research shows the importance of conceptual connections between understanding of variability with regards to samples and understanding of sample space in a probability experiment (Artigue, et al., 2007; Reading & Shaughnessy, 2004). A multiplication conception of sampling variability undergirds conceptual connections which support a deep understanding of statistical inference.

The concept of sampling distribution is an important aspect related to an understanding of probability distribution, but this concept is difficult for college-level students. Research addressing understanding of sampling distributions shows students demonstrate factual knowledge of sampling distribution models and procedures which use sampling distributions, but lack relational knowledge of concepts and conceptual connections inherent to sampling distributions (Chance, et al., 2004; Lunsford, et al., 2006). Students demonstrate inconsistent reasoning about sampling distributions. They can follow rules to predict the behavior of sample statistics, but are unable to describe the process that produces a sampling distribution (Chance, et al., 2004). Students have difficulty applying the Central Limit Theorem and sampling distribution models in applied contexts. Noll (2011) found that even graduate teaching assistants, who demonstrated considerable knowledge of probability distribution models, experienced difficulty as they attempted to make inferences using empirical sampling distributions. Research shows college-level students and graduate teaching assistants exhibit difficulty coordinating a data-centric perspective with a modeling perspective of variability in the context of sampling distributions (Chance, et al., 2004; Noll, 2011). As a result of a teaching experiment addressing students' understanding of variability, Slauson (2008) concluded that understanding the connection between data distributions and measures of variability, and understanding the connection between probability concepts and variability, was very important for students' success in understanding of standard error.

Peters (2011) identified specific indicators of reasoning about variation within a data-centric perspective: anticipation of reasonable variation in data by considering the context of data, recognizing that data descriptions should include variability and center, and recognizing unreasonable variability in data. Other studies confirm the importance of integrated reasoning within and across a data-centric perspective and modeling perspective of variation. Pfannkuch (2005) found that normative perceptions of randomness were central to integrating reasoning of a data-centric perspective with a modeling perspective of variation. When college-level students were given opportunities to reason informally about empirical distributions, they were subsequently able to integrate their reasoning about empirical distribution with theoretical probability distribution models when they also attended to the nature of randomness. In a study involving 33 students in an algebra-based probability and statistics course, results showed that instruction which emphasized the omnipresence of variation and the complementarity of theoretical models and empirical distributions in different contexts promoted understanding of the stochastic nature of statistical concepts and helped students build bridges between their intuitions and statistical inference (Meletiou-Mavrotheris & Lee, 2002).

There is little research addressing the impact of instruction which promotes integration of a data-centric perspective of variation with a modeling perspective of variation aimed to support stochastic conceptions of probability distribution. This study seeks to address this gap in knowledge and to investigate the impact of instruction designed to support development of stochastic conceptions of probability in a calculus-based probability and statistics course.

Understanding Distribution

The concept of distribution is complex and related to big ideas such as variation and sampling (Reading & Shaughnessy, 2004; Watson, 2004) and underlies all statistical ways of reasoning about variation (Wild, 2006). For the purpose of this study, four aspects of distribution are envisioned as being important to an understanding of probability distribution in the post-calculus introductory probability and statistics course. These aspects are: distribution of a random variable, parameterization of a distribution model, distribution of sample versus population distribution, and sampling distribution. The sole word, *distribution,* may be used when referring to any one of these specific distributions. Furthermore, the word, *distribution*, may be singularly used when referring to either an empirical distribution of theoretical distribution. Quite often distinctions between these distributions and use of the word, *distribution*, are not salient to students (Chance, et al., 2004; Shaughnessy, 2007). However, an understanding of probability distribution requires understanding these distinctions and coordinating thinking about empirical and theoretical distributions (Garfield & Ben-Zvi, 2008). .

Distribution of random variable. Probability distributions are a result of random processes. When probabilities are assigned to outcomes in the sample space of a chance experiment, these assignments also determine the probabilities associated with the values of the random variable. A probability distribution for a discrete random variable lists the distinct values of the random variable and the probability associated with each value of the random variable. The random variable (independent variable) and the probability (dependent variable) form a mathematical model, which is the probability distribution. The model of a probability distribution is a complex mathematical model that is comprised of chance outcomes modeled by the random variable at one level, and these chance outcomes, which are the inverse of the random variable, are modeled by the probability at the next level.

A probability distribution may be an empirical distribution that is derived from empirical probabilities; or it may be a theoretical distribution that is derived from theoretical probabilities. Both empirical and theoretical probability distributions are used in statistics. Empirical distributions are constructed from data, or may be simulated, and these distributions can permit insight into what theoretical model might be useful for making inferences. Theoretical distributions provide probabilistic models, which are typically used as the basis for making inferences.

A theoretical probability distribution is a mathematical model which gives the distribution of probabilities for a random variable. A probability distribution may be used to model a random experiment or random process. One aspect involved with understanding probability distribution is an understanding that the model represents a distribution of mass (probability) over values that the random variable can take on. The total density/mass (probability) represented by the distribution (function) is equal to 1, and this is true for all probability distribution models. In a graphical representation of the probability, the total area represented by model is equal to 1, which corresponds to the total density/mass (probability).

Parameterization of distribution model. A probability distribution model is a function with particular parameters. An understanding of probability distribution includes an understanding of the functional relation represented, as well as an understanding of parameterization and the effects that changing values of parameters has on the model. Another important aspect related to parameterization of the probability distribution models is that particular parameterizations of discrete probability distribution models may be approximated by a continuous probability distribution model. Thus, in addition to understanding effects or parameterization, students must also make conceptual connections between particular distributions of discrete random variables and continuous random variables.

Distribution of sample versus population distribution. Another aspect of distribution which is important to understanding probability distribution is distinguishing between distributions of samples versus distributions of populations. When we refer to a distribution of a sample, we are often referring to an empirical distribution, however, at time we may be referring to a theoretical distribution of a hypothetical sample. The same is true for a population distribution. We may actually have data for the entire population and refer to an empirical distribution, which is the distribution of the entire population of interest. However, more often we are referring to a theoretical distribution of a population. Understanding these distinctions and coordinating thinking about empirical and theoretical distributions is an important part of understanding how these aspects are related to probability distribution (Wild, 2006).

Sampling distribution. The fourth aspect of probability distribution is the very important notion of sampling distribution. A sampling distribution is a distribution of a sample statistic. Thus, the sample statistic is the random variable, and a theoretical sampling distribution models the probability of obtaining particular values of the sample statistic. A sampling distribution may also be an empirical distribution obtained from repeated sampling from a population. This repeated sampling may consist of actual samples and statistics determined from those samples, or it may be identified through a simulation of sampling or bootstrapping. Once again, understanding these distinctions and coordinating thinking about empirical and theoretical distributions are important parts of understanding how these aspects are related to probability distribution.

Research related to understanding distribution. Based on research involving college-level students' understanding of distribution, Reading and Reid (2006) proposed a hierarchy of reasoning about distribution. This hierarchy includes two cycles of levels related to what are termed the SOLO levels of cognitive development: prestructural, unistructural, multistructural, and relational (Biggs & Collis, 1991). In the first cycle, the focus of understanding is on the key elements of distribution, center, spread, density, skewness, and outliers, but not on the distribution as a whole. A relational understanding in cycle 1 involves making conceptual links between the key elements of distribution. The focus of the second cycle is understanding distribution as a whole and as a tool for making statistical inferences. Variation has an important role in conceptualizing distribution. Although necessary to conceptualizing distribution, there are other complex processes and conceptualizations needed for students to understand aspects of distribution for making statistical inferences (Hammerman & Rubin, 2004). Wild (2006) contends that the process of statistical thinking depends on the interplay between a stochastic understanding of behavior of data and useful models and states, "… an explicit notion of distribution is not needed until we want to motivate, understand and then use probability models" (p. 22).

Distinguishing between theoretical and empirical distributions requires a high level of abstraction and is difficult for students (Batanero, et al., 2004). Noll and Shaughnessy (2012) found that despite their strong statistical and mathematical knowledge of theoretical probability distributions, graduate teaching assistants experienced difficulty resolving differences between theoretical models and empirical distributions. Batanero et al. (2004) investigated college-level students understanding of normal distribution and found students had a good understanding of properties of the normal distribution, such as mean, median and standard deviation, and were facile with procedures using the normal distribution. Graphical representations were more intuitive for students than strictly using numerical values. Students demonstrated difficulty interpreting areas in frequency histograms and interpreting probabilities under the normal curve. The students demonstrated a propensity to view every variable as having a normal distribution, and a common misconception was thinking a discrete variable with only three possible values was normally distributed. Batanero et al. (2004) reported that a high percentage of students exhibited difficulty discriminating between empirical distributions and theoretical probability distribution models and confused these distributions. Students struggled with a modeling perspective of data and were not able to distinguish the model from real data.

Prodromou (2012) investigated students' coordination of empirical and theoretical distributions. She identified a data-centric perspective of distribution as a focus on data, where data are the starting point with spread across a range of values from which a trend might be discovered. In contrast to a data-centric perspective, a modeling perspective of distribution recognizes probabilities as attributed to a range of possible outcomes in the sample space, and a model gives rise to variation in data. With a modeling perspective, data distributions are perceived as variations from the ideal model. Promodrou and Pratt (2006) argue that these two perspectives on distribution offer different interpretations of the role of data. Stochastic reasoning requires coordination of these two perspectives.

Abrahamson and Wilensky (2007) also considered the problem that individuals have with coordinating empirical and modeling perspectives of distribution. These researchers used bridging tools, which involved representations that bore structural properties of both perspectives, to allow students to compare and contrast both perspectives. Their work showed that bridging tools encouraged individuals to make connections between these two perspectives of distribution. Prodromou (2012) found students gravitated toward causal explanations when they first observed a modeling (theoretical) distribution and subsequently observed it as generating data. However, when they observed underlying randomness, they sought to assimilate this obstacle. When students had the opportunity to start from a distribution of data (empirical), they observed the variation in the data and perceptions of a modeling distribution emerged. An important finding from this study (Prodromou, 2012) is that the students were able to construct connections between empirical and theoretical distributions and come to a better understanding of distribution when their intentions were oriented towards stochastic reasoning and when randomness became part of their interpretation.

Research shows that the concepts of sampling and sampling distribution are often difficult for students to grasp (Heid, et al., 2005; Rubin, et al., 1990; Shaughnessy, 2007; Watson, 2000). Research investigating students' reasoning about sampling distributions indicates that a definition of the Central Limit Theorem coupled with static demonstrations of sampling distributions did not help students develop an integrated understanding of sampling distribution nor did it address persistent misconceptions (Chance, et al., 2004, Lunsford, et al., 2006). Chance et al. (2004) found students had difficulty reasoning about the hypothetical behavior of many samples. These researchers also reported that students indicated confusion when asked to distinguish between the distribution of one sample of data and a distribution of several sample means. Research shows students follow rules and tend to look for patterns, but rarely indicate an understanding of the underlying reasoning of the processes and relationships indicated by a sampling distribution (Chance, et al., 2004; Heid, et al., 2005). Studies also indicate that many students neglect the effect of sample size on sampling variability (Chance, et al., 2004; Stijn et al., 2007). Similarly, Noll (2011) observed that graduate student teaching assistants experienced difficulty when attempting to make inferences using empirical sampling distributions. The teaching assistants did not coordinate multiple perspectives of distribution; they focused on variability in frequencies for different outcomes for each sampling distribution instead of statistical variation. Although the teaching assistants had a strong knowledge of probability distribution models, they did not appear to coordinate their knowledge of theoretical distribution models with empirical sampling distributions (Noll, 2011).

Distribution is the conceptual entity for thinking about variability in data and sample statistics, as well as variability in probability models. Bakker and Gravemeijer (2004) argue that consideration of distribution could be a unifying thread throughout a probability curriculum and that a focus on distribution could bring more coherence to learning probability and statistics. There is a lack of research with college-level learners who encounter distribution and variation in a more formal context in learning probability and statistics, and this issue needs to be addressed (Pfannkuch & Reading, 2006). This study, which investigated post-calculus students' understanding of probability distribution, seeks to address this gap in knowledge and may also inform future curriculum development in probability and statistics.

Inferential Reasoning

Understanding statistical inference and using formal inferential statistical methods is a main learning objective for students enrolled in college-level introductory probability and statistics courses. Statistical inference involves coordination of probabilistic and statistical thinking which includes images of data, distributions of data, randomness and random variability, probability distribution models, sampling and sampling variability, and sampling distribution models, as well as notions of the law of large numbers and the central limit theorem. Researchers who address the learning of statistical inference appreciate the multifaceted and complex thinking involved in understanding statistical inference. Pfannkuch and Wild (2012) state:

> Our journey into laying foundations for statistical inference involved us in exploring and thinking about our own conceptual understanding at a very deep level. Through explicating concepts, which were previously implicit, we have come to appreciate the complex nature of inferential thinking in statistics. It has also made us aware of the many gaps in curricula and textbooks such as no attention to shape of distributions, apart from naming shapes, the imprecise use of language, and the mixing up of descriptive and inferential thoughts (p. 10).

Given the complex nature of thinking required for inferential statistical reasoning, it is not surprising that research shows many students struggle with understanding statistical inference.

Studies reveal that after completing a first course in probability and statistics, many college-level students are able to perform the procedures involved in making a statistical inference, but most students do not demonstrate an understanding of the reasoning involved in making a statistical inference and do not demonstrate an understanding of the reasoning required for interpretation of hypothesis test results (Batanero, et al., 2004; Meletiou-Mavrotheris & Lee, 2002; Reaburn, 2011; Smith, 2008; Thompson, et al., 2007). Other studies confirm earlier research indicating that students do not exhibit realistic anticipations of variability in outcomes of stochastic processes (Kahneman, et al., 1982; Reaburn, 2011). Reaburn (2011) found that students indicated misconceptions related probability and variation and about one-fourth of these students did not expect variability in outcomes resulting from a stochastic process. Kahneman et al. (1982) found study participants did not perceive relationships between sample size and sampling variability.

Notions of sampling, sampling variability, and sampling distributions appear to be crucial to understanding statistical inference (Garfield & Ben-Zvi, 2008; Rubin, et al., 1990; Saldhana & Thompson, 2002; Thompson, 2004; Thompson, et al., 2007). Saldhana and Thompson (2002) describe a multiplicative conception of sample, which entails a rich network of interrelated images supporting understanding of statistical inference, and claim a multiplicative conception of sample undergirds thinking which informs relationships between individual outcomes to distributions of a class of similar outcomes. Smith (2008) found that students who struggle with understanding statistical hypothesis testing lack an appreciation of sampling variability and variability within samples. Smith also found that these students do not necessarily understand statistical hypothesis tests use sampling distributions and probability models to determine the relative unusualness of a sample. Conceptions of probability that are grounded in the conception of distribution are crucial to thinking about distributions of sample statistics and support understanding statistical inferences (Thompson, et al., 2007).

Understanding statistical inference requires coordination of probabilistic and statistical thinking, along with images of probability distribution models, images of distributions of data, and images of distributions of sample statistics. This kind of coordinated thinking involves making conceptual connections between probability and statistics, which are grounded in notions of randomness and stochastic processes. Thompson et al. (2007) found that teachers' difficulties in understanding statistical inference resulted partly from compartmentalization of their knowledge of probability and statistical inference. Thompson et al. (2007) argue that a conception of sampling as a stochastic process is essential to understanding statistical inference. Stochastic understanding of probability distribution involves coordination of empirical distributions of data with theoretical probability distribution models and undergirds conceptual understanding of statistical inference. This study may contribute to knowledge of post-calculus students' stochastic understandings of probability distribution, as well as contribute to research addressing notions foundational to understanding statistical inference.

Theoretical Perspective of Student Learning

Learning is a result of experiences that are shaped by the learners' perceptions of those experiences. A learner's experiences are impacted by the individual's previous knowledge; abilities and motivation; and interactions with peers, curriculum and instructional materials, the instructor and modes of instruction. A reflexive relationship exists between perceptions and experiences, whereby learning experiences shape perceptions and the learners' perceptions shape experiences. This relationship plays a fundamental role in the learning process. Learning results in conceptual understandings, whether instrumental or relational (Skemp, 1976), which are built through perceptions of learning experiences. At every point in the learning process, knowledge shapes subsequent development of understandings.

Stanovich (2011) characterizes learning as a tripartite structure involving two types of processing. The tripartite structure includes three components: the autonomous mind, the algorithmic mind, and the reflective mind. Type 1 learning processes only involve the autonomous mind, which includes systems in the brain that operate autonomously in response to triggering stimuli such as processes involving implicit learning, heuristic processing, automated association, and rules, etc. Execution of process in the autonomous mind is rapid and does not involve high level processing or controls. Stanovich (2011) found that humans tend to be cognitive misers and default to Type 1 processing, which requires less computational expense. Type 2 learning processes involve the algorithmic mind and the reflective mind. The algorithmic mind is involved with cognitive ability and processing and the reflective mind involves thinking dispositions, beliefs, and attitudes. Stanovich (2011) describes Type 2 processing as non-autonomous, relatively slow, and computationally expensive.

The most important aspect of Type 2 thinking is to interrupt and override the autonomous Type 1 thinking. Stanovich (2011) found that the override is initiated by algorithmic-level thinking in concert with appropriate thinking dispositions. Furthermore, in order for the override to be successful override, cognitive resources must be available to substitute for the non-normative Type 1 thinking. If there is a gap in knowledge and the resources are not there, the override fails.

Stanovich (2011) also found that rational thinking encompasses both components of Type 2 processing, thinking dispositions and algorithmic-level cognitive processing, whereas intelligence is primarily restricted to algorithmic-level cognitive processing. West, Toplak, and Stanovich (2008) found that thinking dispositions and cognitive ability explained unique variance in rational thinking. These findings may explain why intelligent individuals resort to heuristics and indicate non-normative thinking and offer insight regarding the impact of thinking dispositions in different problem solving contexts.

The view of learning used for framing this research considers constructivist and situated perspectives of learning as complementary and draws on both of these perspectives for designing instructional intervention and assessment of student understanding. A constructivist perspective of learning provides an analytical frame for conceiving how individual learners develop understanding. Because learning occurs in a particular contextual environment, a situated perspective affords a frame and rationale for affecting the learning environment in ways that promote development of understanding. Both of these perspectives can be reconciled and utilized to inform research about individual development of deep conceptual understanding (Abrahamson, 2009d).

A constructivist view is based on the notion that individuals construct their own reality through actions and reflections on those actions (von Glasersfeld, 1991). Understanding is a result of building mental structures from preexisting knowledge pieces such as ideas, notions, perceptions, and previous understandings. Knowledge is restructured as new understandings are accommodated in the learners' mental structures. Understanding grows as relationships are built and as links between knowledge components are strengthened. This learning process is active and recursive, whereby previously built structures form a basis for subsequent restructuring and development of new understandings.

A situated view of learning is characterized by the enmeshment of learners, practices, and the environment. This research study draws on a situated view that is rooted in Vygotsky's (1978, 1986) ideas of cognitive apprenticeship. This view illuminates ways in which the learner interacts with co-participants, knowledgeable others, and tools within the environment. Prior knowledge frames how the learner interacts with the physical and social natures of the learning environment. The learner will use tools based on his perception of those tools, and these perceptions arise from prior experiences as well as the experiences in the moment of the learning. How a learner uses tools may precipitate affordances or constraints to the learning process, so that tool use does not happen in isolation, but inevitably occurs in some type of context. Vygotsky (1978, 1986) describes tool use as an external activity, which is a result of and an influence of internal activity.

Learning and the Individual Learner

A constructive view frames learning as a process that is individual. This view of learning asserts that knowing is active, it is individual and personal, and it is based on previously constructed knowledge (Ernest, 1996). Understanding in a content domain develops as a result of the learning process. A constructivist view of learning accepts and values naïve probabilistic intuitions; a constructivist view characterizes growth in formal mathematical understandings of probability as necessarily developing from intuitive notions.

The development of mathematical understanding is a complex process. Pirie and Kieren (1999) characterize growth in mathematical understanding "as a dynamic and organizing process" (p. 171) and present a model for describing growth of understanding, which consists of eight potential layers-of-action. Each layer contains all the previous layers and each layer is embedded in all succeeding layers. The layers are named: *Primitive Knowing, Image Making, Image Having, Property Noticing, Formalizing, Observing, Structuring, and Inventising*. This model imbeds intuitive mathematical thinking within more formalized mathematical thinking, and each layer is recursively connected to less sophisticated understanding. The nesting of layers and recursive connection between layers illustrates that growth in understanding in neither linear not mono-directional (Martin, 2008).

Two salient features of the model described by Pirie and Kieren (1994) are *don't need boundaries* and *folding back*. The ability to work in mathematics at an abstract level without the need to reference specific images mentally or physically is captured by the notion of *don't need boundaries*. *Don't need boundaries* allow the model to represent learners' abilities to work with conceptual ideas that are no longer tied to previous forms of understanding. The boundaries do not preclude a learner from returning to an inner level of working, if necessary. *Folding back* is the act of returning to an inner layer to re-visit and re-work understandings and ideas about a mathematical concept (Martin, 2008).

Folding back may occur when a learner becomes aware of limitations of existing understandings and shifts to work mathematics in a less sophisticated or less formal way (Martin, 2008). New understandings are generated and the inner-layer understandings are thickened by what is already understood. The learner's understanding that is more formal assists to inform the less formal actions. In this case, the result of *folding back* is a *thicker understanding* at the level the learner returned to, which informs subsequent learning. A *thicker understanding* means the learner conceives of more conceptual connections within the layer and between layers. *Folding back* to collect at an inner layer may occur when a learner needs to retrieve previous knowledge for a specific purpose (Martin, 2008). This collecting may occur when a learner knows what is needed but does not have expertise that is sufficient for automatic recall.

Folding back may not necessarily result in growth in understanding; however, this process is necessary for growth in mathematical understanding (Martin, 2008). An investigation of a student's understanding of calculus found that an inability to *fold back* to proper images created disconnected understanding and made it impossible for the learner to connect related concepts (Borgen & Manu, 2002). An algorithmic approach to learning calculus may have helped the learner initially formalize her understanding, but this approach appears to have hampered further development of understanding because proper conceptual images were not available for the learner to use in *folding back*.

Effective *folding back* is a powerful facilitator of growth. Martin (2008) found that acts of *folding back* enabled continued growth in mathematical understanding when two key components were part of the learning process: (1) the learner's awareness of limitations of current understanding, and (2) engagement in appropriate and useful mathematical activity at the inner layer.

The Pirie and Kieren model of growth in mathematical understanding provides insight into the process of individual learning, as well as insight into the kinds of instructional interventions that offer the potential to promote development of understanding. Martin (2008) noticed that interventions, which facilitate *folding back*, could be either intentional and explicit, or unintentional and unfocused. This research indicates that interventions, which facilitate effective *folding back*, may help the learner make more conceptual connections in related knowledge and develop deeper conceptual understanding. Thus, an instructional intervention, aimed at stimulating deeper conceptual connections and growth in mathematical understanding, should be designed promote *folding back* to inner layers.

Greeno (1989) characterized growth in mathematical understanding as a function of both formal and informal instruction and highlighted the importance of considering the learner's initial implicit understandings. Greeno proposed that subject-matter knowledge builds on a principled knowledge base and as the base expands, so does principled knowledge in the domain. Greeno (1978) described *principled knowledge* as mathematical knowledge that includes understanding of mathematical skills and procedures integrated with understanding of the ideas and concepts which support mathematical procedures. Knowledge of experts in a given domain is organized in terms of formal principles and related conceptual notions. These ideas related to growth in understanding have implications for curriculum and instruction. Instruction should support learners' development of principled knowledge in the context of an organizing framework.

Learning through Interaction with Instructional Materials

A situated view of learning offers insight into how learners interact with instructional material. One perspective of situated learning maintains that conceptual development is tied to learners' interaction with tools and signs, both of which mediate activity (Vygotsky, 1978, 1986). Signs function as internal mediators where the goal of internal activity is mastering oneself as a learner. On the other hand, tools function as external mediators used to coordinate the learner's influence on the object of activity. For learning in the mathematical domain, signs can be conceived of as mathematical objects, and tools can be representations of those objects, such as symbols or graphs. Tools may also be written or virtual. Mathematical activity can be thought of as a special type of semiotic activity, which encompasses the formation and modification of mathematical representations and the (re)construction of meaning (van Oers, 2000). The development of higher order thinking involves a combination of tool use and signs.

Vygotsky characterizes learning as a process and emphasizes the importance of social interaction that involves more knowledgeable others who scaffold learners' encounters with the environment. Furthermore, the scaffolding of learning should be matched with a learner's developmental level.

Greeno (1991), who depicts knowing and learning as environmental activity, describes another aspect of situated learning. For Greeno, the environment is a knowledge domain and knowing in the domain is characterized as knowing one's way around the environment. Greeno notes that learning in a conceptual environment by oneself is possible because the learner may individually interact with tools such as texts. However, he views this kind of learning as being more difficult than learning with other people. A person's knowing in a conceptual environment involves a shared sense of understanding and includes the ability to reason in the environment and participate in meaningful discourse. Knowing in the domain is learning to reason with mental models, which conform to concepts and principles of the domain, in addition to interacting with these mental models ways that correspond to intellectual practices of the community.

These situated perspectives inform a design of instructional materials, which would meditate interaction between tools and signs in a learning environment that promotes conceptual reasoning aligned with intellectual practices in the domain. Application of these ideas in design-based research, which investigated students' understanding of probability in a technology-based learning environment, found that learners used objects in the environment as semiotic tools to help coordinate intuitive perceptions and more formal mathematical conceptions (Abrahamson, 2007, 2009a, 2009c, 2009d; Abrahamson & Wilensky, 2007). An object in each learning activity context was designed as a bridging tool, which afforded mediation between empirical and theoretical probabilistic and statistical constructs. Although this research took place outside of the classroom, it appears that instructional material could be designed to help students learn to reason stochastically about probabilistic and statistical situations.

Learning through Interaction with Teachers

The assumption that all knowledge is derived from perceptual and conceptual experience does not preclude the influence of teachers on individual cognitive construction. Von Glasersfeld (1996) explains that teachers and others play a major part in an individual's learning experiences. Teachers influence the learning environment by focusing on particular concepts and theories, and by determining which behaviors are validated in the learning processes. For example, instructors who primarily assess procedural knowledge validate student efforts towards development of procedural knowledge. Validation of understanding, whether instrumental or relational (Skemp, 1976), occurs when assessment is aligned with instruction. Teachers play a role in the enculturation of learners into schooling and domain learning by emphasizing particular domain knowledge and activities through their instruction.

Preparing to learn. Large lectures are part of the culture at large universities and a reality of the instructional process. Although large lectures do not afford much individual student-professor interaction, professors and teaching assistants have the opportunity to impact student learning via lecture and design of the learning environment. The common instructional model in college lectures courses is teacher controlled and primarily involves explanatory teaching through telling. Schwartz and Bransford (1998) found that found that teaching by telling can play a significant role in deepening students' understanding if students have had the opportunity to develop appropriate prior knowledge. These researchers sought to determine at which points during the development of knowledge college students are ready to learn by being told something. They found that preparing to learn from a lecture, by analyzing contrasting cases prior to hearing the lecture, helped students differentiate knowledge about the domain and generate appropriate prior knowledge.

An instructional approach, which provides a space for students' to prepare to learn from a lecture, creates a time for telling (Schwartz & Bransford, 1998). This approach is based on theories of perceptual learning (Gibson, 1957, 1994), which illuminate how learners interact with the learning environment. In any situation, perception is relative to certain features of the environment, and some of these features are more salient to learners. Through analysis of contrasting cases, which exemplify key conceptual notions, learners are tuned toward noticing specific features and dimensions of the concept that make the cases distinctive. The resulting well-differentiated information provides a basis for other learning activities (Schwartz & Bransford, 1998).

The teacher's role. Direct teaching through lecture plays a valuable role in students learning, because there is limit to what one can reasonable expect students to discover on their own. Instruction addressing mathematical theories of probability offers a higher level of explanation. This important because it can provide a generative framework through which the learner can extend his understanding based on analysis of contrasting cases (Schwartz & Bransford, 1998). A challenge of direct teaching is to present expert knowledge in ways that are meaningful to the novice. Unpacked knowledge in mathematics looks different than the disciplinary content knowledge taught in university mathematics courses (Adler & Davis, 2006). A teacher's perception of students' mathematical understanding is structured by his understanding of the mathematical content (Simon, 1995). As an expert in the content domain, the instructor's knowledge is often tacit. The challenge is to "unpack" that knowledge so that through instruction learners have opportunities to recognize important distinctions and to develop deep understanding.

Learning through Interaction with Peers

Learning includes both cognitive individual and social components. Vygotsky (1978, 1986) portrayed learning as a mediated activity composed of both internal and external aspects and described conceptual development as driven by the use of tools and interpersonal relations. Culturally developed sign systems and tool use facilitate learning and are critical aspects of activity in the learning environment. Vygotsky also emphasized the importance of social interaction, as well as the social aspect of learning that involves one or more knowledgeable others who scaffold the learner's encounters with the environment (Cobb, 1994). Because learning is a process, the potential for development of "knowing" may occur at any moment in a variety of environments (Cobb, 1999; Cobb & McClain, 2004; Roth, 2001). A situated view of learning recognizes that classroom community and classroom culture can play a role in the development of understanding. The nature of classroom participation is delineated by social norms, sociomathematical norms and mathematical/statistical practices (Cobb, 1999; Lampert, 1990; Yackel & Cobb, 1996). Clearly, the learners' interaction with peers can affect development of understanding and a social component that is present in all learning situations. However, this research study will not intentionally seek to impact or investigate the learners' interactions with peers.

Summary of Learning Perspective

The theoretical perspective for this investigation utilizes the complementarity between constructivist and situated views of learning. This view of learning focuses on the learner, but also acknowledges the impact of the teacher, the instructional material, and peers concerning the building of knowledge. A constructivist view maintains that college students neither passively nor mindlessly assimilate information in their classes. Rather, when engaging with the lecture, students may or may not actively build relevant knowledge that is connected to previous learning and understanding. A student listening to a lecture is constructing knowledge of some sort and developing conceptions, but these conceptions and the knowledge associated with those conceptions may not be aligned with the intended curriculum or aim of the instruction. It is highly likely that the lecturing professor or teaching assistant intends for students to develop deep conceptual connections around formal mathematical probability, but, depending on prior knowledge and previous learning experiences related to probability, these learners may or may not be forming conceptual connections, which facilitate deep understanding of the content.

Research indicates that secondary and tertiary students are compartmentalizing their new knowledge of probability and not associating it with probabilistic experiences outside of the classroom (Rubel, 2007). Traditional classroom experience leads students towards mathematizing probability, for example finding the correct answer or seeking an algorithmic procedure. Thus, it is also likely that students are not making conceptual connections between probability and statistics. As a result, learners are not building a normative conceptual basis in stochastic reasoning that will ground understanding of inferential statistics. A learning perspective that considers the learner as an integral part of the learning environment advocates an instructional design where students have opportunities to engage the content in ways that build on prior knowledge and facilitate differentiation of new knowledge.

Both of these learning perspectives offer insight into how learning develops as a result of interaction with instructional material. It is anticipated that student engagement in activities designed to support development of stochastic reasoning and to promote differentiated knowledge of probability will better prepare students for further learning in probability and statistics. It is also anticipated that these types of activities will support development of conception connections between probability and statistics, resulting in deeper understanding (Abrahamson, 2007, 2009b). This study seeks to scaffold learners' encounters with the learning environment through tool use and to investigate understanding which results from that scaffolding. A research hypothesis for this study is that engagement in appropriate activities, which activate prior knowledge and create fertile ground for learning from lectures, could better prepare students to develop deep conceptual connections between probability and statistics.

Use of Technology in Learning Probability and Statistics

There are two perspectives involving pedagogical use of technology for learning probability and statistics (Beihler, Ben-Zvi, Bakker, & Maker, 2013). One perspective is that technology for learning purposes should mirror technology used in professional practice with a focus on computational tools, i.e., use for learning should mirror the theory and use of professional statistics packages. The other perspective is that technology should offer affordance for learning that focuses on concepts rather than computation, i.e., affords development of probabilistic and statistical reasoning. The affordance perspective places an emphasis on ideas such as offering opportunities to run simulations and experiment with data, making statistics visual, interactive, and dynamic. Beihler (1993, 1997) argued for development of a pedagogical tool which utilized both of these pedagogical perspectives and, would thus support both learning probability and statistics, as well as doing statistical analyses.

The majority of pedagogical tools used as resources to support teaching and learning college-level statistics fall under the perspective of using technology for learning that mirrors use of technology in statistical practice. These tools include: versions of statistical software packages (*SAS*, *SPSS*, *Minitab*, *DataDesk*, and *StatCrunch*), spreadsheets (*Excel*, *Google Spreadsheet*), applets, graphing calculators, and multimedia materials (*ActivStats*). Other software, which falls under the affordance perspective, has been developed exclusively to help students learn probability and statistics. These software tools include *Fathom* and *TinkerPlots*.

Fathom Dynamic Statistics is software developed specifically for learning and doing statistics in high school-level and college-level probability and statistics courses. *Fathom* includes specific features which enable students to explore and analyze data both visually and computationally. These features include functions which enable the learner to drag and drop variables onto graphs to visualize distributions and relationships between variables, functions which enable visualization of the effects of dynamically changing data and parameters represented in real time, functions which link multiple representation of data to informally, and functions which allow the user to effortlessly create simulations and visualize simulation runs in real time.

Virtual environments, which offer visual affordances and simulation tools such as *Fathom*, can support stochastic learning experiences. Research related to instruction in probability and statistics points to the promise of learners' engagement in tasks utilizing a computer-based dynamic statistical environment as a means towards facilitating development of notions of sampling distribution, variability, and inferential reasoning (Meletiou-Mavrotheris, 2003; Sanchez & Inzunsa, 2006). Prodromou (2012) found that when students build models from observations of real situations, the way that they express the relationship between signal and noise (Konoal & Pollatsek, 2002) is important. Prodromou (2012) observed that software tools which enabled students to operationalize variation through creation of simulations helped students coordinate a data-centric perspective of distribution with a modeling perspective of distribution. These instructional supports in a virtual environment helped students connect theoretical probability to simulated phenomena.

Research addressing students' learning probability and statistics in a technological environment demonstrates that use of technology for computational statistics in order to produce graphical representation of distributions and/or to run simulations does not necessarily result in an understanding of key concepts related to inferential statistics (Batanero, et al., 2004; Chance, et al., 2004; Meletiou-Mavrotheris, et al., 2007). A study which compared learning experiences of students in a technology-based, college-level introductory probability and statistics courses, where statistical-practice technology was an integral part of problem-based instruction with learning experiences of students in primarily lecture-based courses characterized by minimal or non-existent incorporation of technology, found no difference between statistical-practice technology-based instruction and non-technology-based instruction on students' understanding of fundamental concepts related to statistical inference (Meletiou-Mavrotheris & Lee, 2002; Meletiou-Mavrotheris, et al., 2007). This study found that regardless of whether they completed a statistical-practice technology-based course or a non-technology-based course, students demonstrated superficial and poorly connected knowledge of many statistical concepts such as sampling, sample statistics versus population parameters, effects of sample size on variability, and sampling distributions. However, another study which used the same kind of problem-based instruction, but in conjunction with *Fathom* software as the technological tool instead of a statistical-practice technological tool, found that students demonstrated more coherent understandings of sampling distribution and other key concepts related to statistical inference (Meletiou-Mavrotheris, 2003). Still other research confirms that learning environments which utilize pedagogical-affordance, technology-based tools such as *Fathom* support development of stochastic reasoning and inferential reasoning (Lee, Angotti, & Tarr, 2010; Maxara & Biehler, 2010, Prodromou, 2012).

Prior research has shown that use of pedagogical-affordance, technology-based tools has the potential to offer particular affordances in learning environments designed to support development of stochastic thinking. This research also shows that use of *Fathom* software can be utilized to create virtual environments which potentially undergird development of stochastic reasoning via connections between empirical distributions and theoretical probability distributions. There is little research addressing use of technology in a calculus-based probability and statistics course, and no research addressing the use of *Fathom* in a calculus-based course. More research is needed to address students' stochastic understanding as a result of use technology tools in learning probability and statistics (Garfield, et al., 2007; Jones, 2005; Jones, et al., 2007). This study, which utilized *Fathom* as part of technology-based instruction in a calculus-based introductory probability and statistics course seeks to add to this gap in knowledge and may also inform future curriculum development and use of technological tools in college-level probability and statistics courses.

Conclusion and Implications for Study

In summary, the analysis of background theory and review of studies presented in this chapter show the need for a study which investigates the impact of instruction designed to support development of a principled knowledge of probability and statistics focused on stochastic understanding of probability distribution. An instructional intervention which utilizes dynamic-statistical-environmental tools to promote development of conceptual connections between data-centric perspectives of distribution and modeling perspectives of distribution along a hypothetical learning trajectory may undergird development of stochastic understanding of probability distribution. Little research has been conducted investigating students' understanding of probability distribution in conjunction with course instruction. Furthermore, there is little research investigating understanding demonstrated by students enrolled in an introductory, calculus-based probability and statistics course. This study aims to utilize a hypothetical learning trajectory posting development of a stochastic understanding of probability distribution and investigate the usefulness of the proposed learning trajectory and provide validity for a framework supporting development of stochastic understanding of probability distribution. This research may inform educators about instruction that builds on students' intuitive understandings and supports development of a principled knowledge of probability distribution which includes stochastic understanding and connections between probability and statistics.

CHAPTER 3: RESEARCH METHODOLOGY AND DESIGN

An analysis of research and literature related to students' understanding of probability suggested the need for a large-scale study investigating understandings developed in the context of learning in college-level probability and statistics courses. The purpose of this study was to investigate the impact of an instructional intervention designed to support college students' stochastic understanding of probability through a control-treatment design. The study employed a mix of both quantitative and qualitative research methods to examine students' understandings of probability that resulted from instructional interventions in a calculus-based introductory probability and statistics course. The design of this study supplemented treatment group instruction with lab assignments aimed at the development of stochastic reasoning. The control group also received supplemental lab assignments, but these addressed calculus review and practice problems. During the latter part of the course, students' understandings were measured with an assessment of their stochastic understandings of probability and an assessment of their understandings of confidence intervals. The study methodology and design are described in this chapter.

Research Methodology

The theoretical framework described in Chapter 1 outlines a view of how an understanding of probability distribution develops in a learning environment; this characterization then informed the research design and choice of research methodology. This study of student learning was based on an assumption that students' understandings are built through learning experiences, which are impacted by the learner, the teacher(s), the instructional material, and peers. This perspective of learning maintains that teaching matters. Educators cannot force an individual to learn, but can affect the learning environment through instruction and curricular design. This research assumed that an instructional approach can play an important role in facilitating learning and a particular instructional focus could affect measurable learning outcomes.

The study employed a mix of both quantitative and qualitative research methods to examine students' understandings of probability that resulted from an instructional intervention in a calculus-based introductory probability and statistics course. The choice of research methodology for this study arises from complexities inherent in studying the effects of an instructional intervention. A quantitative approach can reveal information about difference in outcomes relative to variables and in interaction with other variables and permit generalization of findings. However, this type of approach is directed at confirmation of research questions. By itself, a qualitative approach does not permit generalization, but with purposeful sampling of cases and in-depth qualitative analyses this type of approach affords exploration of questions (Patton, 1990). Teddlie and Tashakkori (2003) state, "A major advantage of mixed methods research is that it enables the researcher to simultaneously answer confirmatory and exploratory questions" (page 15).

Through employment of mixed methods, this study provides analyses between different aspects of students' understanding and types of explanation revealed through quantitative data sources and qualitative data sources (Brannen & Moss, 2012). The mixed methods research design used in this study exploited the power of quantitative methods, permitting generalization of findings, and exploited the power of qualitative methods, providing in-depth analyses of students' thinking and reasoning.

A sequential exploratory design (Tashakkori & Teddlie, 2003) allowed elaboration and enhancement of quantitative-based findings with the incorporation of qualitative data in the context of an integrated interpretation. This design incorporated five phases: (1) collection of quantitative data; (2) cursory analysis of conceptual assessment for purposes of interviewee selection, (3) collection and analysis of qualitative data; (4) use of the qualitative findings to inform coding and analysis of conceptual assessment; (5) final analysis and interpretation of both quantitative and qualitative data. Chapter 4 addresses phase (5) of this research design, final analysis and interpretation of quantitative data sources. Chapter 5 addresses phase (3), analysis of qualitative data sources. Chapter 4, which addresses quantitative data sources, is positioned before Chapter 5, which addresses qualitative data sources, due to the sequencing of data collection and analyses. Collection of all quantitative data occurred before selection of interview participant and collection of quantitative data. The results of a cursory analysis of quantitative conceptual assessment data helped inform purposeful sampling and selection of participants for the qualitative interviews and offered elaboration of the principled knowledge to be further investigated in the qualitative interviews.

During the final week of the course, an email was sent to all study participants soliciting participation in a one-hour, end-of-course interview. Twenty-three students indicated that they were willing to participate in an interview. Selection of interview participants was based on a cursory analysis of the conceptual assessment. The conceptual assessment was designed to assess students' stochastic conceptions in two different problem contexts, and the cursory analysis of students' written responses was accomplished by using Liu and Thompson's (2007) *Theoretical Framework for Probabilistic Understanding* (p. 124). Students' responses to the item 1 on the conceptual quiz (the hospital problem) were mapped onto the Liu and Thompson's (2007) framework and evidence of probabilistic thinking was coded according to this framework. Students' answers to both problems on the conceptual assessment were noted and their responses to the items were coded as stochastic or nonstochastic, based on this cursory mapping.

Selection of interview participants resulted from purposeful sampling from the pool of interview volunteers. A strategy aimed at *maximum variation sampling* (Patton, 1990) was used to select interview participants from each of the two groups (treatment and control) who exhibited diverse probabilistic thinking on the conceptual assessment as indicated by the cursory analysis. Eleven students from the treatment (stochastic reasoning) group volunteered to participate in an interview. Four of the treatment volunteers had a cursory coding of stochastic; four had a cursory coding of nonstochastic; and three had a cursory coding that was questionable (e.g. it was questionable whether their responses to the assessment items were stochastic or nonstochastic). Six of the 11 treatment group volunteers were selected to participate in an end-of-course interview. Of the 6 interview participants from the treatment group, two were cursory coded as stochastic, two were cursory coded as nonstochastic, and two were cursory coded as questionable. Twelve students from the control (calculus review) group volunteered to participate in an interview. Two of the control volunteers had a cursory coding of stochastic; and ten had a cursory coding of nonstochastic. Six of the 12 control group volunteers were selected to participate in an end-of-course interview. Of the 6 interview participants from the control group, two were cursory coded as stochastic, and four were cursory coded as nonstochastic.

Purposeful sampling of interview volunteers resulted in six interview participants selected from the treatment (stochastic reasoning) group and six interview participants selected from the control (calculus review) group. The interview sample included an equal number of individuals from each group in order to facilitate group comparison. As a result of this type of purposeful sampling, any common patterns emerging from the qualitative analysis will be of particular interest and value in describing students' conceptions (Patton, 1990). Individuals within each group were selected based on diverse responses to items on the conceptual assessment as indicated by a cursory analysis. By including individuals within each group who indicated diverse responses, it was possible to more thoroughly describe the variation of conceptions within the groups and better understand these variations in conceptions.

Quantitative research methods enable an examination of understanding and analysis of learning outcomes for large numbers of students and offer the potential for generalization of the findings. These methods provide estimates of the population at large. In this quasi-experimental study, students were assigned to either the treatment or control group for intervention purposes via their registration is a class discussion section. It was assumed this particular sample of students was representative of students who enroll in an introductory, calculus-based, probability and statistics course at the university where the study took place. The design and sampling method used in this study increase confidence in making generalizations to groups of students who enroll in an introductory, calculus-based, probability and statistics course. Research sub-questions 2, 3, and 4 were addressed via collection and analysis of quantitative data:

2. Does instruction designed to support development of stochastic understanding of probability distribution impact students' stochastic conceptions of a probability situation as evidenced on a conceptual assessment?
3. Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding of confidence intervals as measured by the ARTIST assessment?
4. Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding as evidenced on final course examinations administered in an introductory, calculus-based, probability and statistics course?

In this study, the instruments used to assess students' understandings consisted of a conceptual quiz designed to assess students' stochastic understandings of probability, items from the Confidence Interval and Sampling sub-scales of the ARTIST assessment, and the course final examination. Valid measurement of learning outcomes and generalizability of quantitative findings are highly dependent upon the measurement instrument. The ARTIST assessment was used to survey students' understanding of confidence intervals at the end of the course. This tool, which was developed and validated through large-scale testing (delMas, Garfield, Ooms, & Chance, 2007), was utilized to address the third research sub-question. The fourth research sub-question was addressed via analysis of students' course final examinations. In order to address the second research sub-question, students' stochastic conceptions of probability were assessed via two problem contexts in the form of an in-class conceptual quiz. This measurement tool was validated through the use of qualitative research methods via structured interviews with select students.

Qualitative research methods enable exploration of student thinking and complexity of student understanding. In this study, analysis of qualitative data sources enabled the researcher to describe student thinking and to explore contextual meanings that students' ascribed to different problem situations. After a cursory analysis of quantitative results, select students were purposefully sampled to participate in structured interviews with the goal of identifying common patterns, as well as diverse characteristics, of thinking and reasoning (Patton, 1990). Research sub-question 1 was addressed via collection and analysis of qualitative data from structured interviews:

1. What is the nature of students' reasoning when confronted with a probability situation?
    a) How do students characterize a probability situation in terms of an image of a repeatable process?
    b) How do students characterize a probability situation in terms of specification of conditions of a repeatable process?
    c) How do students characterize a probability situation in terms of an image of a distribution of outcomes?

Analysis of the qualitative data facilitated descriptions of students' thinking in and across different problem contexts and provided insight into students' conceptions of probability. Results of the qualitative analyses also informed development of a scoring rubric for students' stochastic conceptions. Students' answers to the conceptual quiz problems were scored using this rubric, and these scores were subsequently used for quantitative analyses of the larger, full data set. Furthermore, qualitative data sources illuminated interpretation of quantitative results.

The utilization of mixed methods afforded an integrated analysis of the main research question: What is the impact of an instructional intervention designed to support development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory calculus-based probability and statistics class? Using mixed methods in this study provided contextual understandings garnered through qualitative analyses to support quantitative findings. Quantitative results were elaborated through insights of students' conceptions gleaned from qualitative analysis of students' interview responses.

The mixed methods used in this research study also provided validation of the assessment of stochastic conceptions and enhanced credibility of the findings through validation of this instrument. Furthermore, employing a mixed methods design facilitated exploration of the quantitative findings. The qualitative results were used to develop a rubric for scoring students' written answers on the conceptual assessment. The qualitative results also helped to explain and interpret the quantitative findings. In other words, the qualitative findings elaborated the quantitative findings and enhanced the final analysis. Employing mixed methods enabled the researcher to bring together a more comprehensive analysis of the inquiry into student understanding of probability distribution.

Research Design

The research design for this study was an instructional intervention in a control-treatment setting. Students were randomly assigned to either the treatment group or control group via their course registration and enrollment in a discussion section associated with one of two large lecture classes. Students in the treatment group received supplemental lab assignments aimed at the development of stochastic reasoning in the context of probability distribution. Students in the control group received supplemental lab assignments consisting of a review of calculus content used in this introductory probability and statistics course. Near the end of the course, all students completed the same stochastic reasoning assessment and the same assessment of their understanding of confidence intervals in order to measure their learning outcomes.

Setting

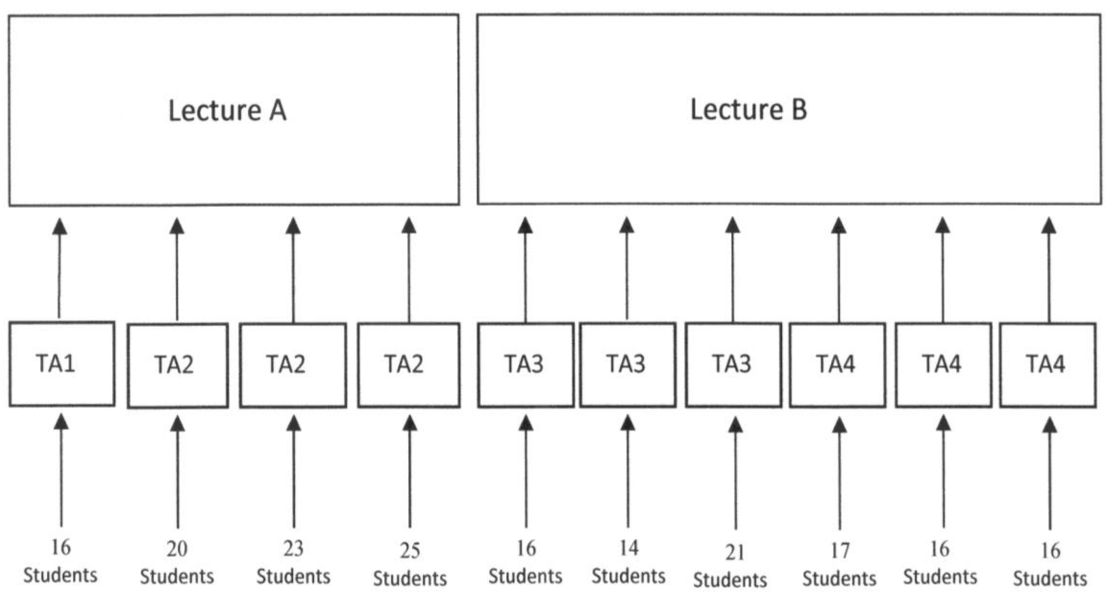
The study was conducted at a large, public university located in an urban area. The undergraduate student population is approximately 25,000 and there are about 10,000 graduate students at this university. This study aimed to improve learning in an introductory calculus-based probability and statistics course. This course is the first of a two-semester sequence in probability and statistics; the prerequisite for the course is two semesters of calculus for science, engineering, and mathematics majors. Most of the students enrolled in the course were juniors and seniors. The course is designed for majors in the applied sciences. It is a required course for computer science majors and economics majors. This particular introductory probability and statistics course is optional for mathematics majors, but many dual mathematics and mathematics education majors take the course.

The textbook used in the course was a custom text that contained selected chapters from *Probability and Statistics for Engineers and the Sciences (7th Ed.)* by DeVore (2009). This custom text also included a section of material covering the law of large numbers that was written by one of the mathematical statisticians in the department. Much of the course focused on probability and probability distribution, and the course culminated with a study of point estimation and confidence intervals. An official syllabus on file in the undergraduate office of the Department of Mathematics at this university outlined content to be covered in the course. The topics include: probability, discrete random variables and their probability distributions (binomial, hypergeometric, Poisson, negative binomial), continuous random variables and their probability distributions (normal, exponential, Gamma), joint probability distributions, sampling distributions and the central limit theorem, point estimation, and confidence intervals. The relative frequency interpretation of probability was emphasized, and coverage of counting techniques was minimized. The perspective of the Departmental faculty responsible for this offering is that the course is an applied course. The Department also offers a probability course for undergraduate students, which involves a study of probability theory; it is proof oriented.

The introductory calculus-based probability and statistics course that is the focus of this study was taught in a large-lecture format with a total enrollment of 255 students. During the semester that data collection for the study took place, students were registered for one of two large-lecture offerings of the course. The two large-lecture sections were each taught by a different professor, and each lecture section had a number of discussion sections attached to it. Both professors presented the content outlined in the official syllabus on file in the mathematics department. While the professors developed and taught their lectures independent of each other as determined by their own syllabus, the instructional format was the same for both classes. In addition to the topics listed in the official syllabus on file in the Department of Mathematics, one of the professors also addressed material on hypothesis tests and test procedures in the lectures. Although each professor customized the syllabus for their offering of this course, students in both lecture sections studied the same core content, worked similar homework problems, and took mid-term examinations and a final examination. These examinations were developed independently by each professor. One professor gave two midterm examinations and the other professor gave three midterm examinations. Both professors gave a comprehensive final examination.

All enrolled students were expected to attend large-lecture classes twice each week for 75 minutes. In addition, a discussion section associated with the lecture-class met one day each week for 75 minutes. The professor-led, large lectures were taught in a traditional lecture format. Teaching in the discussion sections was also primarily lecture, but depending on the teaching assistant, there may or may not have been considerable interaction with the students during class time. The instructional approach to probabilistic and statistical content in this course was traditional and mathematically oriented, aligned with the manner in which content was presented in the textbook for the course. Grades were assigned based on students' performance on midterm examinations, homework, and a comprehensive final examination. The professors allowed use of calculators on examinations. Formula sheets were not provided for examinations.

The participants in this study were the students enrolled in a discussion section that was associated with one of the two lecture sections. Teaching assistants (TA), who were doctoral students in either pure mathematics or applied mathematics, taught the discussion sections. One of the lecture sections had four discussion sections, and the other lecture section had six discussion sections. There were approximately 25 students in each discussion section. Figure 3.1 illustrates the lecture-discussion class distribution and lists the number of study participants enrolled in each discussion section. The research design used in this study assumed that although teaching in the lectures and discussion sections was not identical, they were homogenous in most respects. Because there were at most 25 students enrolled in each discussion section, it would be more likely for a discussion section to have its own classroom culture than the large lecture section.



Organization of the Lecture-Discussion Class Distribution

Figure 3.1. The two large-lecture sections are denoted by large rectangles. Smaller rectangles show the distribution of discussion sections associated with each large-lecture section and accompanying teaching assistants.

Course lecturers. The professors who led the large-lecture sections were mathematicians with many years of teaching experience; both had previously taught this particular course more than once. Each professor held a Ph.D. in mathematics. The professor of Lecture A had 15 years of teaching experience. This professor completed a dissertation in analysis and probability theory and at the time of the study was doing research in the area of probability. The professor of Lecture B had over 30 years of teaching experience and completed a dissertation in functional analysis. In addition, Professor B held a minor in probability and statistics. Both professors indicated a keen interest in teaching and helping students learn.

Professor A indicated that she addressed the notions of randomness and variability in her lectures by using examples of coin flips and dice. Through these examples, students considered what the possible outcomes for a probability experiment could be, and she emphasized that one could not predetermine what particular outcome would occur. Professor A contrasted determinism with randomness. She connected the notion of variability in outcomes and sampling to the notions of random variable and variance. Professor A indicated that she frequently used lots of examples in her lectures and that she did address the law of large numbers in her lectures.

Professor B reported that she addressed the notion of randomness when teaching about probability. In particular, Professor B indicated that she talked about continuity and the difference between discrete and continuous random variables. She also indicated that it was important for students to have the "big picture" in mind. When teaching about sampling variability, Professor B used *Fathom* to demonstrate variability in sample means. She indicated that she taught the law of large numbers in her lecture classes.

Course teaching assistants. Each large-lecture section was supported by two teaching assistants. Three of the teaching assistants were doctoral students in the applied mathematics program, and one was a doctoral student in pure mathematics. One of the two teaching assistants for Lecture A held a bachelor's degree in mathematics and the other held a bachelor's degree in business with a master's degree in applied mathematics. One of the two teaching assistants for Lecture B held a bachelor's degree in mathematics, and the other held both a bachelor's degree and a master's degree in statistics. All of the teaching assistants were experienced and had previously assisted with this particular course. Also, all of the teaching assistants had previous experience teaching discussion sections for calculus classes. One teaching assistant (TA2) had taught this course during the previous summer session as a stand-alone instructor.

All of the teaching assistants indicated that in their opinion this course was challenging to teach in comparison to calculus or other mathematics courses they had been involved with. Some said the course was challenging because probability was a new topic for students. All indicated thinking that students had difficulty reasoning about probability. In particular, the teaching assistants mentioned students' difficulties with material in the latter part of the course when the topics moved from mathematical models to reasoning about statistics and point estimation. All of the teaching assistants indicated that a good number of students in the course struggled with integration and applying integral calculus.

The teaching assistants expressed a desire to help students learn. They appreciated students asking questions and participating in class, but said that in some discussion classes only a few students talked. Each of the teaching assistants indicated that it was rewarding to help students understand what they had learned in the lecture and to help students solve problems.

Participants

Participants in this study were students enrolled in a calculus-based, one-semester, introductory course in probability and statistics at a large, public, university. Each participant was enrolled in a discussion section associated with one of two large-lecture sections. A total of 184 students participated in the study. Eighty-four of the participants were enrolled in the offering deemed Lecture A, and 100 participants were enrolled in Lecture B. At the beginning of the course, participants in the study were contacted via *Blackboard* course email and asked to complete an online survey (Appendix A); demographic information about study participants was obtained from this survey. Study participants were asked to self-report information about academic major, college mathematics credits earned, grade point average, Math SAT score, prior statistics credits earned, and prior calculus credits. Data indicating gender and race was not collected for this study.

Of the 184 students who agreed to participate in the study, 157 completed the online background survey. The typical student enrolled in this calculus-based probability and statistics course was majoring in mathematics, science, or other quantitative field of study. Table 3.1 shows the distribution of majors in this course.

Table 3.1

Distribution of Participants' Major Field of Study

Major	Frequency	Percent
Computer Science	62	33.7
Business, Economics, or Other Social Sciences	33	17.9
Engineering and Physical Sciences	24	13.0
Mathematics	20	10.9
Biology and Chemistry	18	9.8
Major not reported	27	14.7
Total	186	100

This course fulfilled a requirement for computer science majors and economics majors. Approximately half of the students were majoring in either one of these two fields of study, or in a business or other related social science major. The class standing of most of the students enrolled in this course was either a Junior or Senior. Of the study participants, 71 were Juniors (38.6%), and 52 were Seniors (28.3%). Only two of the study participants were Freshman (1.1%), and 32 were Sophomores (17.4%). The majority of participants reported grade point averages between 3.0 and 4.0. Table 3.2 shows the distribution of self-reported grade point average for students in the study at the beginning of the semester.

Students in the study were also asked to report their SAT Math score and the number of college mathematics credits they had earned prior to enrollment in the course. The median SAT Math score for study participants was 710. Students majoring in mathematics reported the highest SAT Math scores with a median of 740. Those majoring in engineering or a physical science field reported SAT Math scores with a median of 730. The median SAT Math scores for the other majors was approximately 700. Mathematics majors entered the course with the highest number of prior earned credits in mathematics with an interquartile range between 9 and 29 credits. The median number of prior college mathematics credits earned for mathematics majors was 19. For engineering and physical science majors, the median number of prior mathematics credits earned was 15. For computer science, business, economics, and other social science majors the median number of mathematics credits was 12; the median was 10 for the biology and chemistry majors.

Table 3.2

Distribution of Grade Point Averages for Participants

GPA range	Frequency	Percent
3.5 – 4.0	61	33.2
3.0 – 3.4	50	27.2
2.5 – 2.9	32	17.4
2.0 – 2.4	14	7.6
GPA not reported	27	14.7
Total	184	100

The prerequisite for this course was two semesters of calculus. Thirty-seven percent of the students reported earning credits for their first semester of calculus via AP Calculus examination, and 48.4% of the students reported earning credits for their first semester of calculus by completing a calculus course in college. While 15.8% of the students reported earning credits for their second semester of calculus via AP Calculus examination, 69.6% reported earning credits for their second semester of calculus while in college. The remaining 27 students (14.7%) did not complete the survey and this information is missing.

Although the course in which the study took place did not require a prerequisite statistics course, a number of students had completed a college-level statistics course prior to enrollment in the introductory calculus-based probability and statistics course in which the study took place. Of the study participants, 34 (18.5%) had earned credit for

AP Statistics prior to enrollment in this course. There were 31 (16.8%) study participants who entered this course with credits earned for completion of a prior statistics course in college, and 90 (48.9%) participants had not taken a college-level statistics course prior to enrollment in the course in which the study tool place. The remaining 29 students (15.8%) did not provide this information.

The study participants were representative of groups of undergraduate students who enroll in an introductory calculus-based probability and statistics courses across the country. Because the course prerequisite is two semesters of calculus, these students can be considered to be representative of the population of undergraduate students enrolled at a four-year, Research I university, with background in calculus. Although these students had taken more mathematics courses than the general undergraduate population, for many, this was their first course in probability and statistics. Based on prior research, it was assumed that these students entered the course holding probabilistic conceptions aligned with thinking exhibited by novice learners in an algebra-based introductory statistics classes and by high school students (Abrahamson, 2007; Abrahamson & Wilensky, 2007; Barragues, et al., 2007; Hernandez, et al., 2006; Ives, 2007; Lunsford, et al., 2006; Wilensky, 1997).

Methods

This study involved implementation of an instructional intervention and subsequent assessment of students' understanding. This research assumed that instructional interventions aimed at development of stochastic reasoning can be used to elicit and foster students' understandings of probability distribution. This study of the impact of such instructional interventions was based on the following assumptions about teaching and learning: (1) teaching impacts learning and can facilitate learning with understanding; (2) effective teaching elicits students' pre-existing understandings and builds on that understanding; (3) effective teaching helps students develop deep knowledge connections in the context of a conceptual frame for the content domain (Bransford, et al., 2000).

Treatment-control design. In the probability and statistics class that is the focus of this study, students were grouped in discussion sections via their course registration, and the discussion sections were grouped under one of two lecture sections. The two large-lecture classes meet on Tuesdays and Thursdays in back-to-back time slots. The discussion sections for both lectures meet on either Mondays or Wednesdays. On either of these days, discussion sections from both lectures could be meeting at the same time of day. It is assumed that students were randomly assigned to a discussion sections via their course registration.

Assignment of each discussion section to either the treatment or control condition was done by random selection. As a result, students in a particular discussion section were assigned to either the treatment or control group for purposes of the instructional intervention. Three of the teaching assistants each led three discussion sections, and one teaching assistant led a single discussion section. As represented in Figure 3.2, one teaching assistant associated with lecture A taught two stochastic reasoning (treatment) discussion sections and one calculus review (control) discussion section. The other teaching assistant associated with Lecture A taught a calculus review (control) discussion section. One teaching assistant associated with Lecture B taught two stochastic reasoning (treatment) discussion sections and one calculus review (control) discussion section, while the other teaching assistant associated with Lecture B taught two calculus review (control) discussion sections and one stochastic reasoning (treatment) discussion section. This assignment balanced the treatment and control across the lectures and discussion sites for the data collection. Thus, this research design sought to mitigate confounding variables due to differences in teaching between the lecturers and between the teaching assistants.

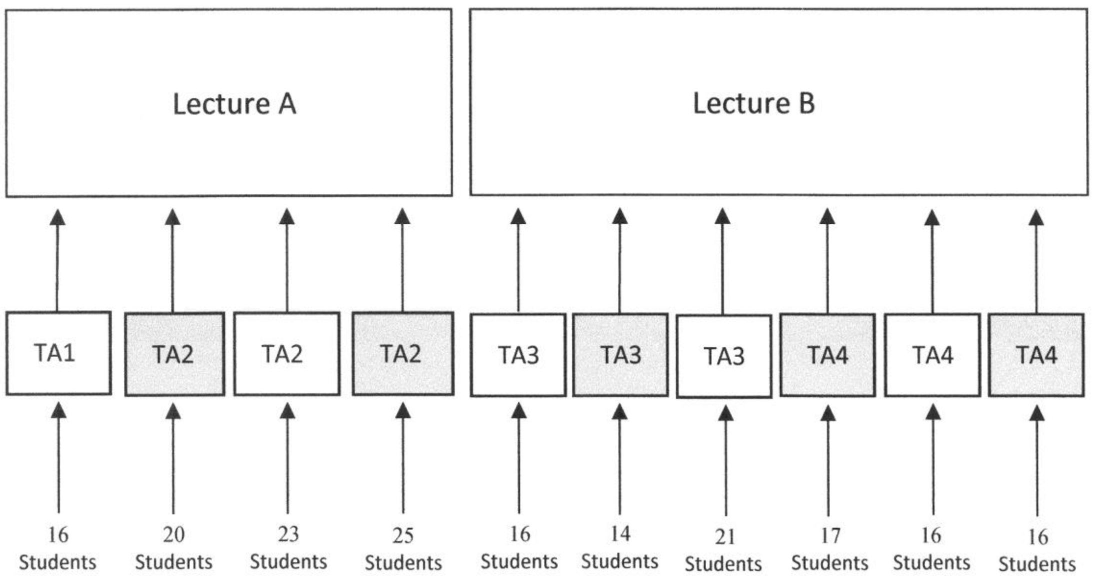


Diagram of the Treatment-Control Design

Figure 3.2: The shaded "TA" boxes represent treatment (stochastic reasoning) discussion sections, and the "TA" boxes with a white background represent control (calculus review) discussion sections.

Instructional intervention. The instructional intervention in this study consisted of two aspects. One aspect of the intervention supported student learning via supplementary instructional lab material designed for the control and treatment contexts, and the other instructional aspect developed and supported student learning via direct instruction. For the purposes of this study, students were assigned to either the treatment or control group via the random assignment of registration in a discussion section. All students in the course received supplemental lab assignments to be completed outside of class, in addition to their regular homework assignments. These lab assignments were required as part of the course and counted towards students' final course grades as a completion grade.

The instructional intervention was designed to prepare students to learn from the lectures and therefore provide greater opportunity for students to make deeper conceptual connections. Schwartz and Bransford (1998) found that "teaching by telling" can play a significant role in deepening students' understanding if students have had the opportunity to develop appropriate prior knowledge. The instructional approach of the intervention for this study provided a space for students to prepare to learn from a lecture, theoretically preparing students to build deeper connections when engaged in learning in the course of the lecture or discussion section.

Students in the treatment group engaged in technology-supported lab assignments aimed at development of stochastic reasoning. Students in the control group engaged in lab assignments that reviewed calculus topics used in the course. Engagement with the lab assignments occurred before receiving a lecture on the relevant content. After submitting the lab assignment, students in the treatment group received direct instruction linking the stochastic concepts explored in the lab assignment with specific content in the course. Students in the control group received direct instruction linking the calculus review problems to types of problems encountered in the probability and statistics course curriculum.

The direct instruction for both the treatment and control group was conducted via online videos. Instructional videos for each group were posted on *Blackboard* the day after each lab assignment was due. Students in each group watched the video instruction and responded to an online closure question relevant to the lab material. In order to receive a completion grade for each lab, students were required to do four things: (1) engage in the lab and complete the written work required for the lab assignment, (2) hand in a paper copy of the assignment to their teaching assistant in the discussion section, (3) watch the direct instructional video connecting the lab to course content, and (4) submit an online answer to the closure question for that lab assignment.

Engagement in the supplemental lab assignments occurred before students received a lecture from their professor on the applicable content. Also, direct instruction connecting the assignments to the course curriculum was set up to occur before students received a lecture from their professor on the applicable content. The majority of participants watched the instructional videos prior to receiving a lecture on the content; however, a few students waited until the latter part of the course to watch the instructional video and answer the closure question associated with each lab. After each lab assignment and accompanying instructional video were posted on *Blackboard;* these remained available to students for the remainder of the semester.

All lab assignments were posted on the course *Blackboard* website for the study. As such, both treatment and control group lab assignments were handled via *Blackboard* where all postings could be controlled according to the discussion section students were enrolled in. Hence, the treatment group only "saw" the stochastic-reasoning assignment material, and the control group only "saw" the calculus-review assignment material.

Students accessed the lab assignments directly from *Blackboard* and submitted their completed lab assignments to their teaching assistant at their discussion section meeting. The direct instruction linking the lab assignment material to course content took place via *Blackboard*. The researcher used *Camtasia* software to produce the video lessons, which students accessed and downloaded from the *Blackboard* website.

The researcher wrote scripts for each of the direct instructional videos that accompanied the stochastic reasoning lab material. She utilized *PowerPoint* and *Fathom* software to accomplish the instruction in a manner that could be electronically recorded and produced using *Camtasia* software. In order to write and produce the direct instructional videos that accompanied the calculus-review lab material, the researcher worked with a mathematician who regularly taught calculus at the university where the study took place and also had considerable experience teaching the course in which the study took place. This mathematician was not teaching this course at the time of the study. The researcher wrote a draft script for each of the calculus-review instructional videos, and then consulted with the mathematician prior to teaching and recording the lesson. After the lesson was finalized, it was taught using *PowerPoint*. Video and audio of all direct instruction was recorded using *Camtasia* software, and this software was also used to edit and produce the video lessons. The mathematician approved each video lesson prior to its posting on *Blackboard.*

These direct instructional videos ranged from 4 to 10 minutes in length. After viewing the video lesson, each student was required to answer a closure question about the video content in order to receive a completion grade for the lab assignment. Students' answers to the closure questions were submitted electronically via *Blackboard*. The questions served three purposes. First, in the case of the stochastic lab group, the closure questions provided a formative assessment of students' learning related to the lab assignment and course content. Second, the closure questions provided a way of tracking student engagement with the study materials. Third, submission of the lab assignment, watching the video, and answering the closure question were required to receive the points for a completion grade. Closure questions for each lab group can be found in Appendix B.

Description of instructional intervention: Treatment. The overall theme of the instructional intervention for the treatment group was development of stochastic reasoning in the context of different aspects of probability distribution. Students in the treatment group engaged in a series of tasks which comprised six lab assignments. Each lab assignment was aimed at developing students' thinking and reasoning about probability distribution as a stochastic model with an emphasis on stochastic conceptions of probability, variability, and distribution (Ireland & Watson, 2009). The tasks, which comprised these lab assignments, were designed to elicit students' probabilistic intuitions and then to build on those intuitions with the intention of coordinating students' intuitive understanding with more formal understandings through engagement in computational experiments (Kuhn, Hoppe, Lingnau, & Wichmann, 2006; Kyriazis, Psycharis, & Korres, 2006). It was anticipated that student engagement in these activities would promote differentiated knowledge of probability, which would better prepare students for further learning about probability and would support that learning in ways that would allow for development of deeper understanding (Abrahamson, 2007, 2009).

Each task directly connected to topics that students would be learning in the course and required students to use *Fathom* software to simulate and model probability experiments related to the topics. A hypothetical learning trajectory for probability distribution adapted from Liu and Thompson's framework (2007) informed the sequence of tasks for the instructional intervention. This hypothetical learning trajectory consisted of five phases: (1) conceive of a stochastic process, (2) understand conditions of that process, (3) understand implementations of the process, (4) conceive of a distribution of outcomes, and (5) conceive of a probability distribution model. Conceptual notions comprising a deep, integrated understanding of probability distribution were addressed is each phase of the learning trajectory and informed the development of each laboratory assignment.

The laboratory assignment tasks were carefully structured to connect probabilistic learning outcomes within and across each phase and were aimed at development of stochastic conceptions of probability (see Table 3.3). The tasks were designed to engage students in coordinating their thinking about complementary notions related to probability distribution such as randomness and determinism, experimental and theoretical probabilities, and independence versus dependence (Abrahamson & Wilensky, 2007; Ireland & Watson, 2009; Lane Peres, 2006). An aim of these tasks was to enable students to couple the "real world" with "the real math" by connecting their intuitive and experimental knowledge with their formal mathematical knowledge (Abrahamson, 2007). Furthermore, by engaging in these tasks, students had the opportunity to develop appropriate prior knowledge, which helped set the stage for deeper understanding of content and explanations offered by their professor and discussion teachers (Schwartz & Bransford, 1998; Schwartz & Martin, 2004).

Table 3.3

Stochastic-Reasoning Lab Assignment Tasks

Learning axes	Bridging tool	Context	Probability Distribution Construct
Lab 1: Random versus deterministic	Law of large numbers; random walk	Flipping coins	Randomness and variability
Lab 2: Independence versus dependence	Contingency table; marginal probabilities; $P[A \cap B]$; $P[A \mid B]$	Hands and eyes; flipping 2 coins	Role of independence
Lab 3: Theoretical probability versus empirical probability	Binomial random variables with $p = 0.5$ and $n = 4$	Flipping 4 coins	Comparison of theoretical with experimental distribution; sample space; random variable
Lab 4: Theoretical probability versus empirical probability	Binomial random variables with varying p and n: $p = .5, .1, .8$ with $n = 20$; and $p = .5$ with $n = 20, 50, 100$ and 200	Bottled water versus tap water	Comparison of theoretical with experimental distribution; parameterization of binomial distribution
Lab 5: Distribution of data versus probability distribution model; discrete distribution model versus continuous distribution model	Binomial random variables with varying parameters, Poisson distribution, samples of different sizes from continuous distributions (uniform; normal, and exponential)	Hard drive failure; bus arrival times; weights of packages; time between arrivals	Binomial distribution relationship to Poisson distribution; probability distribution as a model; sampling from continuous distributions
Lab 6: Distribution of sample versus distribution of population versus sampling distribution of statistic	Law of large numbers and the central limit theorem; sampling and distribution	Sampling pennies	Comparison of distributions of a sample, a population, and sampling distribution for a statistic

Note. Learning axes and bridging tools (Abrahamson & Wilensky, 2007) for these tasks were designed to engage students in thinking about complementary notions in probability and statistics.

Description of instructional intervention: Control. Faculty in the Department of Mathematics expressed concern that many students in this introductory probability and statistic course do not remember how to do many types of calculus problems and need more practice with calculus ideas that are utilized in this introductory probability and statistics course. The overall theme of the instructional intervention for the control group was a review of calculus concepts and calculus tools used in this course. Students in the control group engaged in six lab assignments; each calculus-review lab consisted of problems which reviewed calculus topics students had previously encountered in their two-semester calculus course sequence. The aim of the calculus review lab assignments was to review content students should have already learned, but they may not have used for a while.

The calculus-review lab assignments focused on five major content areas in calculus: sequences and series, functions and their analytic properties, implications of the Fundamental Theorem of Calculus, differentiation, and integration. Each lab required students to work calculus problems and provided engagement with calculus concepts that students encountered in this probability and statistics course. Table 3.4 gives a list of the topics covered in each of the six calculus-review lab assignments.

Table 3.4

Calculus-Review Lab Assignment Topics

Lab	Calculus Content
1	Sequences and series:
	Limits of sequences and sums of infinite series
	Convergence and divergence
	Arithmetic and geometric series
	Taylor's Theorem and Taylor series
	Expansions of power, exponential, logarithmic, and trigonometric functions
2	Analytic properties of functions and differentiation:
	Continuity and differentiability
	Functions defined piecewise
	Limiting values and asymptotes
	Derivatives of elementary functions
	Product, quotient and chain rules for derivatives
	Sketching the graph of a function
3	Implications of the Fundamental Theorem of Calculus:
	Sketching the graph of a function
	Integral as area of a region under a curve
	The Fundamental Theorem of Calculus
	Definite and indefinite integrals
4	Integration part I:
	Integrals of elementary functions
	Evaluation of definite integrals
	Integrals of piecewise functions
5	Integration part II:
	Evaluation of definite integrals
	Substitution techniques
	Improper integrals
6	Techniques of integration:
	Substitution techniques
	Integration by parts
	Improper integrals
	Infinite limits and integrals of unbounded functions

Expectations for teaching faculty and teaching assistants. The lectures given by the course professors and the teaching conducted by teaching assistants in the discussion sections proceeded as was typical for this course. Student engagement with the lab material took place outside of class lecture and discussion periods. No additional time or preparation was requested of the professors or teaching assistants, other than incorporation of the additional completion grades for the lab assignments onto their grade-sheets. The professors and teaching assistants received copies and solution keys to all of the supplemental lab assignments for both the treatment and control group. Although the professors were aware of the lab material content, they did not address connections of lab assignment material to course content during their lectures. Furthermore, the teaching assistants did not connect lab material to the course content in their discussion sections. However, teaching assistants indicated that they addressed some students' questions about the calculus review material. All student questions or concerns about the stochastic lab material were forwarded to the researcher. At the beginning of the course, a few students complained that the calculus-review problems were too difficult. After consultation with the professors, the researcher simplified questions on the calculus-review lab assignments.

Time Line

The researcher met with the course professors and teaching assistants in late August 2010, prior to the beginning of the semester, and explained the study and solicited their help. Prior to the beginning of fall classes, the researcher consulted with instructional technology staff at the university regarding implementation of technologies for purposes of the study, and the researcher set up *Blackboard* space for facilitation of the study.

On the first day of classes, the researcher attended both lecture classes and was given time to explain the study to students and to solicit their participation in the study. The researcher described the study during the large lecture class, explained the procedures impacting students, and handed out student consent forms. Those who agreed to participate in the study were asked sign the consent form, and those who did not agree to participate in the study were asked to leave the form blank. All students folded the form in half and returned one copy to the researcher. The researcher collected the consent forms and compiled the list of volunteers for the study. The volunteer participants were sent an individual email via *Blackboard* within a few days of filling out the consent form. The email thanked the student for his/her willingness to participate in the study and directed the student to go the *Blackboard* website and complete the online student survey questionnaire. An overview of the study and a copy of the consent form were posted on the *Blackboard* website devoted to the study.

As an incentive to participate in the study, students had an opportunity to win an *iPad*. For every two lab assignments completed, a student received one entry into a raffle for the *iPad.* Students who completed the conceptual quiz and the confidence interval assessment received one additional ticket for the raffle. Thus, a student who was highly engaged with the study, who completed all of the lab assignments, and who took the conceptual quiz and the confidence interval assessment, received five entries in the raffle. The raffle for awarding the *iPad* took place in a public location within the building housing the Department of Mathematics during the final week of the course.

All students in the course had the opportunity to complete the supplemental lab assignments and earn extra credit for completion of the lab assignments. Each lab assignment was posted one week before it was due, and the instructional videos associated with each lab assignment were posted within one week of the lab due date. The teaching assistants collected the completed lab assignments and delivered them to a prearranged location in the mathematics department for the researcher to pick up. The researcher graded all lab assignments and recorded a completion grade. At the end of the semester, completion grades for each student were compiled by the researcher and given to the teaching assistants. The lab assignments were designed to prepare students to learn from the lecture. Hence, the labs were scheduled for students to engage lab material before related course content was covered during a course lecture. As such, students completed the lab assignments during the first 10 weeks of the class. Table 3.5 shows the lab assignment schedule over the course of the semester.

Table 3.5

Lab Assignment Schedule

Assignment	Due at end of:
Lab 1	Week 2
Lab 2	Week 3
Lab 3	Week 4
Lab 4	Week 6
Lab 5	Week 8
Lab 6	Week 10

On a prearranged date near the end of the semester, the teaching assistants administered the conceptual assessment in the form of a quiz in each of their discussion section meetings. The researcher completed a preliminary analysis of students' answers to the conceptual quiz problems and sent out an email to all study participants seeking students who would be willing to participate in a one-hour interview. Students were offered a stipend of $30 for participation in the interview.

Eleven students from the stochastic reasoning lab group and 12 students from the calculus review lab group responded to the email and indicated they were willing to participate in an interview. From these 23 respondents, 6 students from each group were selected for an interview based on their written answers to the conceptual quiz. The interviews took place in a conference room in the building housing the Department of Mathematics during the week of final examinations, after students had completed their final examination for the course in which this study took place.

For the additional workload imposed by the study, each teaching assistant received a stipend of $80. The professors for the course were not asked to do any additional work in conjunction with this study other than meeting with the researcher before the course began. The professors agreed to give permission for the researcher to access their students, to work with their teaching assistants, and to obtain the final examination score and final course grade for students participating in the study. The professors also ensured that the participants' final examination scores and final course grade were provided to the researcher.

Design of Lab Assignments

Design of stochastic lab assignments. The stochastic reasoning lab assignments are designed to move students through a learning trajectory aimed at the development of stochastic conceptions in the context of probability distributions. Each lab focuses on characteristics of probability distribution, which are identified in the literature as important aspects of understanding probability distribution.

Framework for understanding probability distribution. The first step in designing the stochastic reasoning lab assignments was to develop a conceptual framework that described and delineated the important aspects of probability distribution (see Table 3.6). Three overarching constructs frame an understanding of probability distribution that was developed for this study: probability, variability, and distribution. Probabilistic notions comprising each of these three constructs are included in the framework for understanding probability distribution. A principled knowledge of probability distribution means that one understands the ideas and concepts which undergird mathematical procedures used in conjunction with probability distribution (Spillane, 2000). This kind of knowledge also includes an understanding of connections between the constructs of probability, variability, and distribution, as well as understanding connections among notions within each construct and across the constructs.

Table 3.6

Framework for Understanding of Probability Distribution

Probability Distribution		
Probability	Variability	Distribution
Coordination of empirical and theoretical probability	Randomness and random variability	Distribution of random variable
Random variable	Law of large numbers	Parameterization of distribution model
Sample space	Unit-to-unit variability	Distribution of sample versus population distribution
Independence versus dependence	Sampling variability	Sampling distribution
Model for inference	Variability of sample statistics and the central limit theorem	

In an introductory calculus-based probability and statistics course, learning outcomes for a principled knowledge of probability distribution include the following objectives. The learner will:

A. Interpret the probability distribution as a mathematical model for describing random phenomena.

B. Understand the role of independence in probability distribution models.

C. Coordinate understandings of experimental distributions and theoretical distributions:

   - Compare theoretical distributions with experimental distributions derived from chance experiments;
   - Coordinate notions of empirical outcomes with sample space and combinations and permutation of theoretical outcomes.

D. Describe what a probability distribution would look like for different parameters:

   - Describe how changes in parameters affect the shape, expected value, and variance of a probability distribution.

E. Interpret graphical representations of probability distributions

   - Describe likelihood corresponding to particular values of random variables and particular outcomes;
   - Interpret notions of area as probability: represented as histogram bars in discrete distributions, and area under the probability density function curve for continuous distributions.

F. Compare theoretical distributions and describe the relationships among theoretical distributions:

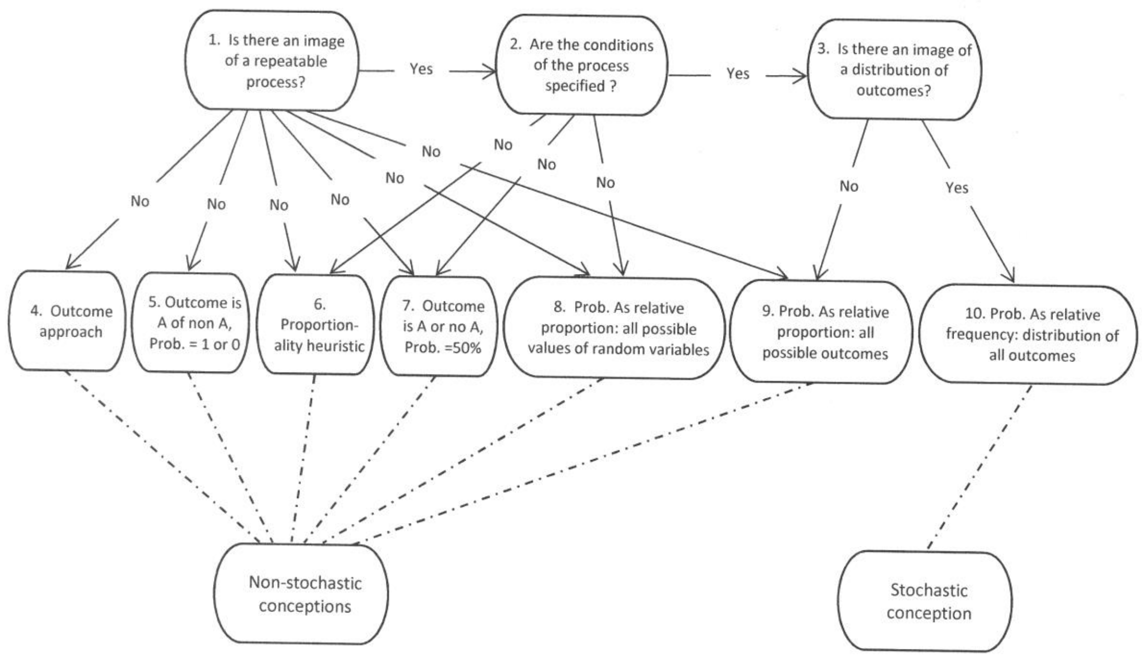
   - Describe the role of parameters for the model;
   - Apply the model in appropriate contexts.

G. Interpret probability distribution as a model:

- Compare samples from different population distributions;
- Use probability distribution models to make appropriate inferences.

These learning outcomes were used to develop stochastic reasoning assignments around constructs that frame probability distribution.

Stochastics conceptions of probability. After framing an understanding of probability distribution, the constructs within this framework were mapped onto a framework that describes the development of stochastic conceptions of probability (Liu & Thompson, 2007). Liu and Thompson (2007) found that statistical inference is grounded in a stochastic conception of probability and argue that "a powerful conception of probability that supports reasoning in statistical inference builds heavily on the conception of distribution of outcomes" (pg. 156). These researchers claim that developing a stochastic conception of probability involves a series of particular ways of thinking: (1) conceiving an underlying repeatable process, (2) understanding the conditions and implementations of this process in such a way that it produces a collection of variable outcomes, and (3) imagining a distribution of outcomes that are developed from repeating this process. Liu and Thompson's theoretical framework for probabilistic understanding names ten ways of conceiving of probability. These researchers describe a number of paths that lead to a nonstochastic conception of probability, but only one path leads to a stochastic conception of probability which culminates in conceiving of probability as a relative frequency distribution of all outcomes (see Figure 3.3). This model provides a tool for thinking about how learners might develop stochastic conceptions and a framework to begin thinking about how stochastic understandings of probability distribution might develop.



Theoretical Framework for Probabilistic Understanding

Figure 3.3. The theoretical framework for probabilistic understanding shows ten ways of thinking about probability. Path 1→2→3→10 is the only way to conceive of probability as a stochastic process (Liu & Thompson, 2007, pg. 124).

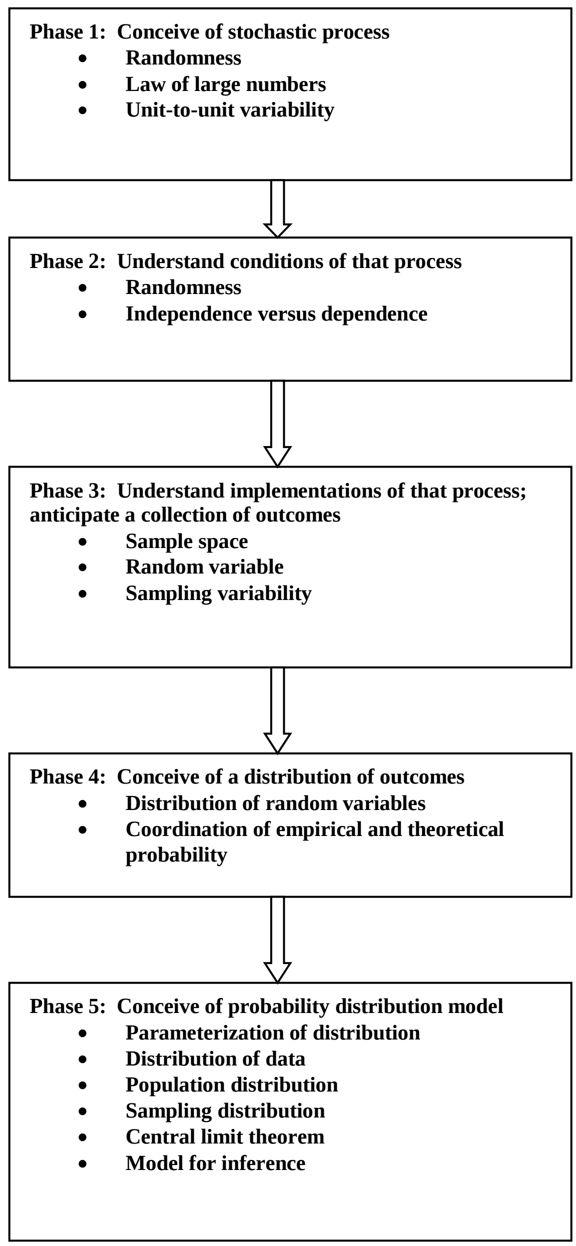
Using the framework for understanding probability distribution which was developed by the researcher (see Table 3.6) and Liu and Thompson's theoretical framework for probabilistic understanding, the researcher created a combined framework for examining development of stochastic conceptions of probability distribution. This combined framework is comprised of five phases and describes development of stochastic conceptions in the context of probability distributions.

1. Conceiving of a probability situation as stochastic process that has an underlying in repeatable process: developed through notions of randomness, the law of large numbers, unit-to-unit variability of random phenomena;

2. Understanding the conditions of a stochastic process: developed through notions of randomness, independence versus dependence;
3. Understanding implementations of a stochastic process and anticipating that repeating a stochastic process would produce a collection of outcomes: developed through notions of sample space, random variable, sampling variability;
4. Conceiving of a distribution of outcomes that are developed from repeating a stochastic process: developed through notions of distribution of random variable, coordination of empirical and theoretical distribution;
5. Conceiving of probability distribution as a model: developed through parameterization of distribution, distribution of data, population distribution, sampling distribution; central limit theorem, model for inference

The combined framework was used to construct a hypothetical learning trajectory for a stochastic understanding of probability distribution.

Hypothetical learning trajectory. The hypothetical learning trajectory for development of stochastic conceptions of probability begins with engaging learners' notions of randomness and culminates in a modeling perspective of probability distribution. Each phase in the learning trajectory is connected to a research-based framework for the development of stochastic conceptions in the context of probability distribution, which is adapted from the chains of reasoning in the framework Liu and Thompson (2007) developed for probabilistic understanding. This learning trajectory is also connected to learning outcomes for understanding probability distribution. Therefore, each phase in the learning trajectory is designed to support specific learning outcomes related to stochastic conceptions of probability distribution. The hypothetical learning trajectory is used to inform the design of stochastic reasoning tasks and sequence these tasks within the course content (see Figure 3.4).



Hypothetical Learning Trajectory for Probability Distribution

Figure 3.4. Probabilistic notions comprising an understanding of probability distribution mapped onto five phases of a hypothetical learning trajectory.

Phase one. A hypothetical learning trajectory for probability distribution begins with conceiving of a probability situation as the expression of a stochastic process in contexts that support thinking about randomness and variability in random phenomena. The goal of learning would be to conceive of repeatable processes that underlay a stochastic process. Learning would focus on thinking about the randomness, unit-to-unit variability, and the empirical law of large numbers; learning would ground formal understandings of a theoretical development of the law of large numbers.

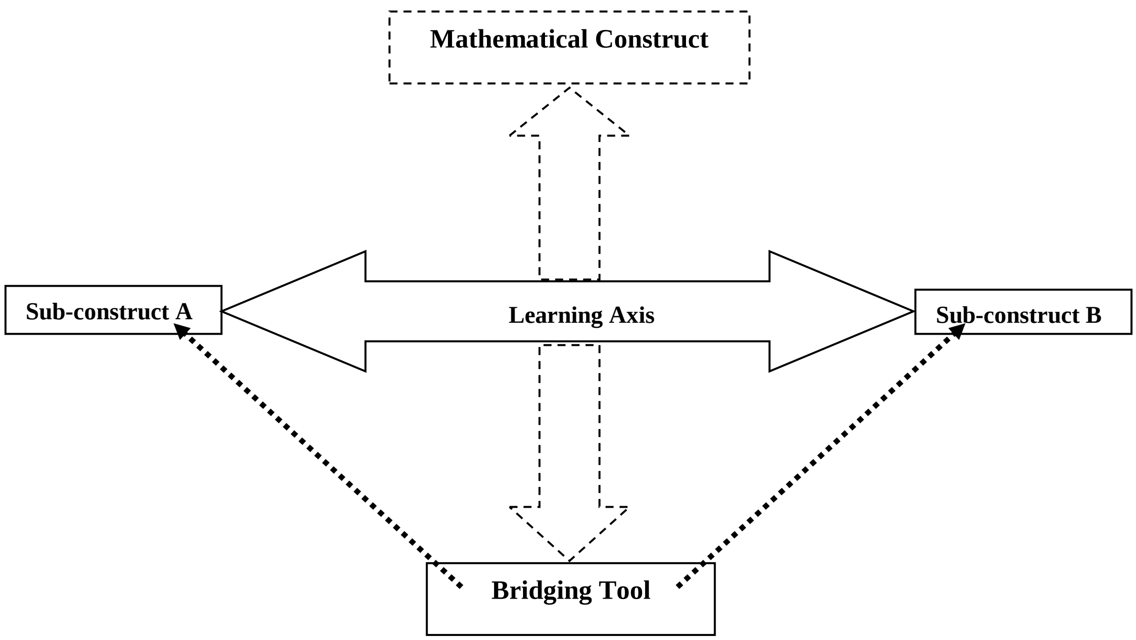
Phase two. The second phase in a hypothetical learning trajectory for probability distribution facilitates understanding that the processes could be repeated under essentially the same conditions. In this phase, the goal of learning would be to understand the conditions and implementations of the process. Learning would support development of understanding of independence versus dependence, and connecting these to notions to randomness and variability.

Phase three. The third phase in a hypothetical learning trajectory for probability distribution promotes understanding about the implementations of a stochastic process, so that the process produces a collection of variable outcomes. Learning is this phase would be directed towards notions of random variable and sample space. This phase would prompt the learner to coordinate empirically observed outcomes with notions of sample space and combination and permutations of theoretical outcomes. Notions of sampling variability would also be developed as learners repeatedly implement a stochastic process and analyze the outcomes of each repetition.

*Phase four***.** The fourth phase in this trajectory moves learners toward envisioning a distribution of outcomes. Implementation of the stochastic process produces a collection of outcomes that can be aggregated and graphically represented. Learning would concentrate on notions of distribution of random variable and on coordination of empirical and theoretical distributions. Learning would support interpretation of graphical representations of distributions and interpretations of area as probability. This phase would develop learners' ability to conceive of how changes in distribution parameters affect the shape, expectation, and variance of the distribution.

Phase five. The fifth phase in the learning trajectory extends learners' understanding of probability as a distribution toward envisioning probability distribution as a mathematical model. In this phase of the trajectory, learning encapsulates understandings of probability distribution developed in phases one through four and moves the learner towards thinking of probability distribution as a probabilistic (mathematical) object. In this phase the following notions of probability distribution are developed: parameterization, distributions of data, population distribution, sampling distribution, the central limit theorem, and model for inference. Learning would be directed towards comparisons of theoretical distributions and describing relationships among distributions, as well as parameterization and modeling, and applying modeling in appropriate contexts. Learning would also prompt thinking about a distribution of data as a sample at one level of a model and repeated sampling resulting in distributions of sample statistics at another level of a model. In addition, learning would support thinking about probability distribution as a model used to make inferences.

Learning axes and bridging tools. Learning axes and bridging tools inform how the learning environment could support development of stochastic conceptions of probability distribution. In the treatment group lab assignments, stochastic notions are presented as learning axes (Abrahamson & Wilensky, 2007), which involve comparing and contrasting stochastic ideas. These learning axes are conceptual composites of various aspects of probability, variability, and distribution that inform development of stochastic conceptions of probability distribution. Within each lab assignment, complementary ideas of the target concept are used to support students' construction of the concept as a coordination of these elements. Cognitive conflict is mediated through problem-solving activities with a bridging tool. In this study, bridging tools are designed to tap and stimulate students' previous understanding and link these notions to representations, which facilitate development of formal mathematical, probabilistic and statistical understandings (see Figure 3.5). The bridging tools mediate students' mathematical understandings.



Learning Axis and Bridging Tools

Figure 3.5. Diagram of relationships between components consisting of mathematical construct, learning axis, and bridging tool in the learning environment (adapted from Abrahamson & Wilensky, 2007, pg. 31)

Sequence of stochastic lab assignments. The hypothetical learning trajectory informs the sequencing of stochastic reasoning tasks and the content focus of each task. Each phase in the hypothetical learning trajectory for development of stochastic conceptions of probability distribution is associated with one or more lab assignments:

1. Conceiving of a probability situation as stochastic process that has an underlying in repeatable component

   Lab 1: Randomness and variability

2. Understanding the conditions of a stochastic process

   Lab 2: Independence

3. Understanding implementations of a stochastic process and anticipating that repeating a stochastic process would produce a collection of outcomes

   Lab 3: Sample space and binomial distribution parameters

4. Conceiving of a distribution of outcomes that are developed from repeating a stochastic process

   Lab 3: Sample space and binomial distribution parameters

   Lab 4: Binomial distribution relationship to Poisson; and normal approximation to binomial distribution

5. Conceiving of probability distribution as a model: developed through parameterization of distribution, distribution of data, population distribution, sampling distribution; central limit theorem, model for inference

   Lab 4: Binomial distribution relationship to Poisson; and normal approximation to binomial distribution

Lab 5: Probability distribution as a model; sampling from continuous distributions

Lab 6: Sampling distribution

The phase of the hypothetical learning trajectory associated with each stochastic lab assignment is displayed in Table 3.7. The table also indicates the mapping of each of the learning outcomes associated with probability distribution onto the stochastic lab assignment aimed at development of those particular aspects of learning. The first column of this table gives the number of the stochastic reasoning lab assignment and a lab code. This code maps each lab assignment lab to learning outcomes for probability distribution (see pages 121, 122) and to phases of the hypothetical learning trajectory for development of stochastic conceptions of probability distribution (see Figure 3.6). The first number in the left-hand column represents the lab assignment number. The numbers in parenthesis represent phases in the hypothetical learning trajectory. The letters refer to learning outcomes for probability distribution. For example, “4-CD (4, 5)” indicates the fourth stochastic lab assignment was aimed at development of two stochastic conceptions: (4) conceiving of a distribution of outcomes that are developed from repeating a stochastic process and stochastic conception, and (5) conceiving of probability distribution as a model. These conceptions were developed through comparing and contrasting notions of empirical probability and theoretical probability and were associated with learning outcome C (coordinate understandings of experimental distributions and theoretical distributions) and learning outcome D (describe what a probability distribution would look like for different parameters).

Table 3.7

Stochastic Reasoning Labs Associated with Hypothetical Learning Trajectory

Lab Code	Learning axes	Bridging tool	Context	Probability Construct
1-A(1)	Random versus deterministic	Law of large numbers; random walk	Flipping coins	Randomness and variability
2-B(2)	Independence versus dependence	Contingency table; marginal probabilities; P[A∩B] ; P[A | B]	Hands and eyes; flipping 2 coins	Role of independence
3-C(3, 4)	Theoretical probability versus empirical probability	Binomial random variables with $p = 0.5$ and $n = 4$	Flipping 4 coins	Comparison of theoretical with experimental distribution; sample space; random variable
4-CD(4, 5)	Theoretical probability versus empirical probability	Binomial random variables with varying p and n: $p = .5, .1, .8$ with $n = 20$; and $p = .5$ with $n = 20, 50, 100$ and 200	Bottled water versus tap water	Comparison of theoretical with experimental distribution; parameterization of binomial distribution
5-EF(4, 5)	Distribution of data versus probability distribution model; discrete distribution model versus continuous distribution model	Binomial random variables with varying parameters, Poisson distribution, samples of different sizes from continuous distributions (uniform; normal, and exponential)	Hard drive failure; bus arrival times; weights of packages; time between arrivals	Binomial distribution relationship to Poisson; probability distribution as model; sampling from continuous distributions
6-FG(5)	Distribution of sample versus distribution of population versus sampling distribution of statistic	Law of large numbers and central limit theorem; sampling and sampling distribution	Sampling pennies	Comparison of distributions of a sample, a population, and sampling distribution for a statistic

Note. The code in the first column maps each lab assignment lab to learning outcomes for probability distribution and to phases of the hypothetical learning trajectory for development of stochastic conceptions of probability distribution.

Instructional approach of stochastic lab assignments. First, the students were asked to make a conjecture about a stochastic situation. Next, the student constructed a simulation of the problem, conducted the virtual simulation, and then virtually organized data obtained the simulation. Depending on the problem situation, the student was first instructed to conduct a small number of "hands-on" trials of a simulation before constructing and conducting the virtual simulation using *Fathom*. After conducting the simulation, the student investigated contrasts between a mathematical probabilistic model for the situation and experimental results. Then, the student was required to analyze his experimental (simulated) results in light of the mathematical model. An exemplar of the stochastic-reasoning lab assignments can be found in Appendix C.

Design of calculus-review lab assignments. The calculus-review lab assignments addressed five major content areas in calculus, which were connected to content in this introductory probability and statistics course. In this course, developing an understanding of probabilistic models is built on an understanding of these five content areas in calculus and on the ability to use calculus to solve problems. The five content areas addressed in the calculus-review lab assignments were: sequences and series, functions and their analytic properties, implications of the Fundamental Theorem of Calculus, differentiation, and integration. Each lab required students to work calculus problems and provided practice of calculus concepts that students encounter in this introductory probability and statistics course.

For each calculus-review lab assignment, review problems were selected that addressed specific topics that were included in that lab. The review problems written for each lab assignment were adapted from homework exercises found in the textbook used in the two-semester calculus sequence taught at the university where the study took place, *Calculus with Analytic Geometry (6th Ed.)* by Ellis and Gulick (2002). Because many of the students in this course completed calculus at this university, they would have encountered similar problems when taking their calculus course. An exemplar of the calculus-review lab assignments can be found in Appendix D.

Quantitative Phase

Data collected from all study participants for purposes of quantitative analyses consisted of: (1) a student background survey, (2) a conceptual quiz assessing stochastic understanding, (3) a confidence interval assessment consisting of ARTIST assessment items, (4) data indicating lab assignment completion, (5) students' final examinations, and (6) students' final grades for the course.

Student Background Survey

A background survey, in the form of an online questionnaire (see Appendix A), was administered to study participants. After the researcher collected the signed consent forms and compiled a list of volunteers for the study, the participants were sent an email via *Blackboard* that instructed students to go the *Blackboard* website and complete the online student questionnaire. This questionnaire was posted on *Blackboard*, and all background information was collected via *Blackboard.* Hence, no class time as taken for filling out the questionnaire. The survey asked students to report information about their major area of study, grade point average, class, SAT Math score, number of mathematics credits earned, and previous coursework in mathematics and statistics. These data were used to describe the sample and to secure control variables for the quantitative analysis.

Conceptual Assessment

A conceptual assessment of students' stochastic understanding of probability was given at the end of the semester in the form of a quiz (see Appendix E). This assessment was designed to measure students' stochastic understanding of probability in two different contexts and was administered during a discussion section's class-time. The assessment contained two problems. Liu and Thompson (2007) stress that there are different interpretations of probability statements and situations: "A situation is not stochastic in and of itself. It is how one conceives of a situation that makes it stochastic or nonstochastic" (p. 126). The two problems on the conceptual assessment were selected not only because they involved different contexts, but also because each problem represents a probabilistic situation that can be approached stochastically or nonstochastically. Thus, it was anticipated that students' responses to these two conceptual quiz items could reveal differing interpretations of probability.

The hospital problem. The first question posed to students on the conceptual quiz was a problem that had been used is previous studies which investigated undergraduate students' understanding of probability (Fischbein & Schnarch, 1997; Reaburn, 2008; Tversy & Kahneman, 1972; Watson, 2000). Researchers in these studies used adaptations of a problem with a scenario involving two hospitals. In this research study, the problem is referred to as the "hospital problem."

> A town has two hospitals. On the average, there are 45 babies delivered each day in the larger hospital. The smaller hospital has about 15 births each day. Fifty percent of all babies born in the town are boys. In one year each hospital recorded those days in which the number of boys born was 60% or more of the total deliveries for that day in that hospital. Is it more likely that the larger hospital recorded more such days, that the smaller hospital did, or that the two hospitals roughly recorded the same number of such days? Explain your reasoning.

The hospital problem can be conceived of stochastically or nonstochastically. A stochastic conception of the hospital problem would include an image of a repeatable process, specification of conditions of the process, and an image of a distribution of outcomes (Liu & Thompson, 2007).

In the hospital problem context, an image of a repeatable process would include conceiving of gender associated with each birth as an outcome of a random phenomenon. In addition, an image of a repeatable process requires considering that the number of births each day is the result of a random process and each day represents a repetition of the process.

Specification of conditions of a repeatable process in the hospital problem context would include an assumption that the births are independent and the outcome of a baby boy is equally likely for each repetition of the process. Specification of conditions of the process would also require understanding that the number of boys born each day at each hospital would vary. Thus, each hospital can be conceived of as a sample of births and each day the number of boys born at the hospital will vary, as will the proportion of boys born at that hospital. Furthermore, repetition of the process will produce a collection of births at each hospital.

An image of a distribution of outcomes for the hospital problem would include conceiving of a random variable that represents the number of boys born at the hospital on a given day. Repetition of the process would produce a collection of values of the random variable and a frequency distribution of values of the random variable representing the number of boys born at a hospital on a given day. Liu and Thompson (2007) define having an image of a distribution of outcomes as anticipating that the relative frequency of outcomes will have a stable distribution. In the context of the hospital problem, there are two distributions pertaining to the two hospitals. For the smaller hospital, the random variable representing the number of boys born each day can take on the values {0, 1, 2, …, 15}. For the larger hospital, the random variable representing the number of boys born each day can take on the values {0, 1, 2, …, 45}. A stochastic conception of the hospital problem entails an image of each of these distributions as well as the distinctive features of each distribution. The important distinction is that the variability in the number of boys born each day is greater for the smaller hospital than for the larger hospital. Furthermore, a stochastic conception of the hospital problem entails understanding that over the course of a year, there will be stabilization in the frequencies of the random variable representing the number of boys born each day.

The theater problem. The second problem posed to students on the conceptual quiz was adapted from a question Liu and Thompson (2007) posed to secondary mathematics teachers in a study investigating stochastic conceptions of probability. This problem is referred to as the "theater problem."

> Anthony works at a theater, taking tickets for one movie per night at a theater that holds 250 people. The town has 30,000 people. He estimates that he knows 300 of them by name. Anthony noticed that he often saw at least two people he knew. Assume that people are not coming to the theater because they know Anthony and there is nothing special about the type of movie. Is it in fact unusual that at least two people Anthony knows attend the movie? Give an explanation of your reasoning and be sure to address issues of randomness and distribution.

The theater problem can also be conceived of stochastically or nonstochastically. A stochastic conception of the theater problem would include an image of a repeatable process, specification of conditions of the process, and an image of a distribution of outcomes (Liu & Thompson, 2007).

In the context of the theater problem, an image of a repeatable process includes conceiving that an individual attending a movie at the theater where Anthony works is the result of a random process. This means that an image of a repeatable process for the theater problem considers that each day Anthony works at the theater represents a repetition of the process, and this repetition of the process results in someone going to the movie theater where Anthony works.

Specification of conditions of the process in the theater problem context requires considering that an outcome of the process is either Anthony knows the person who attends the movie or he does not know the person. The problem states that the reader is to assume that people are not coming to the theater because they know Anthony and that there is nothing special about the type of movie. This assumption indicates that it is equally likely for any individual in the town to go to the movie theater. A stochastic conception of this problem requires consideration of the assumption of independence. Because people attending the movie are drawn from a large population, the independence assumption is met. Thus, each day that Anthony works at the theater can be conceived of as a random sample from a large population. This sampling process is repeated each day that Anthony works. The sample size is 250 based on an assumption that all theater seats are filled. The proportion of people Anthony knows on any given day is a sample proportion. Conceiving of the conditions of the stochastic process requires understanding that the number of people that Anthony knows at the theater each day will vary, thus the proportion of people that Anthony knows will vary. Repeating the process of filling the theater with people on nights that Anthony works produces a collection of outcomes, which are the number of people Anthony knows on any given night.

Image of distribution in the theater problem context requires conceiving that repetition of the process of filling the theater each night produces a collection of outcomes, where the outcomes of this process are the number of people Anthony knows on any given night. This number varies and plotting the distribution of these values yields a frequency distribution. Each night the theater is filled represents a random sample from the population of the town. A stochastic conception requires understanding that there will be stabilization in the frequencies of the random variable representing the number of people Anthony knows each night. The proportion of people Anthony knows each night is a sample proportion, and there will also be a stabilization of relative frequencies for the random variable representing the proportion of people Anthony knows each night. Students were asked, “Is it in fact unusual that at least two people Anthony knows attend the movie?” A stochastic conception requires an interpretation of unusualness that involves an underlying repeatable process. It also involves an understanding that unusualness is quantifiable and can be conceived of as a statistical value (Liu & Thompson, 2007).

Stochastic conception scoring. Data from the conceptual assessment was used to analyze students’ stochastic understandings and investigate the relationship between these understandings and their performance in the course. In addition this assessment data was used to investigate the relationship between students’ stochastic understandings and their understanding of confidence intervals. In order to conduct quantitative analyses, student work on the conceptual assessment needed to be scored. Scores were assigned to each study participant’s written work on the conceptual assessment using a rubric based on evidence gleaned from qualitative analysis of interviews conducted with twelve students. The qualitative analysis, which is presented in Chapter 5, validated the conceptual assessment as an instrument used to measure evidence of stochastic reasoning and informed development of a scoring rubric for this instrument.

The qualitative analysis presented in Chapter 5 indicated evidence of stochastic reasoning across three hierarchical categories: image of a repeatable process, image of specification of conditions, and image of distribution. Interview participants who presented the highest level of conceptualization, an image of distribution, also presented an image of specification of conditions and an image of a repeatable process. The interview participants who presented an image of distribution in both problem contexts were characterized as holding a stochastic conception. Table 3.8 summarizes findings of the qualitative analysis which describe evidence of a stochastic conception.

Table 3.8

Summary of Interview-Based Evidence for Stochastic Conception

Stochastic Conception
Image for Repeatable Process
Indicates understanding that the process is repeated under essentially the same conditions
Indicates the repeatable process yields outcomes and *describes outcomes of the process*
Indicates thinking about a repeated experiment
Indicates that repetition of process results in repeated sampling
Connects thinking about the process to a model
May connect thinking about the process to running a simulation
Image for Specification of Conditions
Indicates that repetition of the process yields a collection of variable outcomes
Assumes outcomes are independent
Describes a sampling process where each selection is equally likely
Indicates the sampling process produces samples that are representative of population
Indicates that variability in outcomes is related to sample size
Indicates conceiving of conditions of the process in relation to an underlying model
Images for Distribution of Outcomes
Indicates thinking about a distribution of outcomes
Connects thinking about expectation to variability
Attends to variability when thinking about distribution of outcomes; connects notions of distribution variance, shape, and sample size
Indicates thinking about the law of large numbers in relation to stabilization of frequencies over a large number of repetitions of the process
Indicates thinking about an underlying distribution model
Quantifies "unusual" *as deviation from expectation in terms of a distribution modal*

Note. Italicized images were only evidenced as a result of the interview. Images not italicized were evidenced as a result of the interview and via student's individual written answers for the conceptual quiz.

Evidence of stochastic reasoning presented in Table 3.8 was gleaned from students' written work on the conceptual quiz, as well as students' responses to interview probes. In order to develop a rubric for scoring study participants' written work on the conceptual assessment, only images of stochastic reasoning evidenced both as a result of students' interviews and students' written answers for the conceptual quiz were considered. Stochastic images that were only evidenced as a result of the interview were not used to develop the scoring rubric. Table 3.9 gives the rubric used to score all study participants' written work on the conceptual assessment.

Table 3.9

Stochastic Conception Rubric

Stochastic Conception	Possible Points Hospital	Possible Points Theater
Image for Repeatable Process		
Indicates the repeatable process yields outcomes	1	1
Indicates that repetition of process results in repeated sampling	1	1
Indicates connecting process to a model	1	–
Image for Specification of Conditions		
Indicates that repetition of the process yields a collection of variable outcomes	1	1
Indicates the sampling process produces samples that are representative of population	–	1
Indicates that variability in outcomes is related to sample size	1	–
Images for Distribution of Outcomes		
Connects expectation to variability	1	1
Indicates underlying distribution model	1	1
Quantifies "unusual" *as deviation from expectation in terms of a distribution modal*	–	1

Not all aspects for each image were evidenced in both problem contexts. For example, an indication that variability in outcomes is related to sample size was only evidenced in the hospital problem context, and an indication that the sampling process produces samples that are representative of the population was only evidenced in the theater problem context. Using the rubric shown in Table 3.9, the maximum number of points possible for the stochastic conception quiz was 14.

Confidence Interval Assessment

Assessment of students' understanding of confidence intervals consisted of 12 items drawn from Assessment Resource Tools for Improving Statistical Thinking (ARTIST) topic scales (Garfield, delMas, & Chance, 2006). The confidence interval assessment for this study, which consisted of 12 ARTIST items, was used to measure students' understandings of confidence intervals (see Appendix F). ARTIST assessment items were developed by statistics education researchers to measure student-learning outcomes for a first course in statistics. This instrument is designed to assess aspects of the knowledge which many educators in the statistics education community believe is a key to statistical thinking and reasoning (delMas, Garfield, Ooms, & Chance, 2007). The focus of the instrument is assessment of students' conceptual understanding rather than merely knowledge of statistical procedures. In this study, students' scores on the confidence interval assessment were used to investigate the relationship between students' stochastic understanding about probability distribution and their understanding of confidence intervals.

Confidence interval assessment scoring. The confidence interval assessment that was administered to students in this study consisted of 12 multiple choice items. Ten of the 12 items were the items which comprise the ARTIST Confidence Interval Scale. The two additional items were drawn from the ARTIST Sampling Distribution Scale. This assessment was administered online via *Blackboard* in the form of an extra-credit quiz. Students were instructed that the quiz was closed book and closed notes. After submitting the quiz, students were asked to affirm that they had not received any unauthorized assistance on the quiz in adherence with the university honor pledge.

Students were given 30 minutes to complete the quiz and a timer appeared in the upper-right corner of their computer screen during the quiz. Students were not awarded any extra-credit points if they exceeded the 30-minute time limit for the quiz.

Answers to each question were scored as correct or incorrect; students received one point for each correct answer. The distribution of scores on the confidence interval assessment was fairly symmetric with mean 7.91 and standard deviation 2.066. The median score was 8 and 50% of the students scored between 6 and 10 on the assessment.

Final Course Examination

Scores from the comprehensive final course examination were used to investigate the traditionally assessed understandings of students in this introductory, calculus-based probability and statistics course. These measurements provided an indication of students' procedural knowledge and instrumental understandings of probability and statistical concepts, as well as applications of calculus. Comprehensive final examinations were developed independently by professors for each lecture section. Thus questions on the final examination for the two lecture sections differed as well as the total number of points possible on the examination for each lecture section. The maximum score possible on the final examination for Lecture A was 100 (M = 78.3, SD = 13.6). The maximum score possible on the final examination for Lecture B was 150 (M = 98.1, SD = 24.6). The distribution of final examination scores for each lecture section was fairly symmetric. Final examination scores were standardized independently for each lecture section and the standardized final examination scores were used for analytical purposes.

Qualitative Phase

Twelve students participated in interviews at the end of the course. Six students were selected from the treatment group and six from the control group. The purpose of the interviews was to probe students' reasoning about probability in the two contexts of the conceptual quiz, as well as to probe their generalized reasoning about probability. Prior research indicates the importance of context and cues in students' responses to probability tasks. An important reason for conducting the interviews was to ascertain students' interpretation of the conceptual quiz items, which presented students with two differing contexts for thinking about probability. The qualitative data components of the study provided insight into students' thinking and reasoning about their written answers on the conceptual assessment, which subsequently informed the development of a scoring rubric that was used to score all student participant responses to the conceptual quiz items. Probing students' thinking about their answers to the conceptual assessment provided validity to the conceptual quiz items, informed interpretation of the quantitative data, and strengthened the overall interpretation of results.

Interview Design

The end-of-course interview was designed to provide insight into student thinking about stochastic conceptions of probability distribution and conceptual connections between probability and statistics. The purpose of the interview was to gain deeper insight into students' stochastic conceptions of probability and to describe and illuminate the quantitative results. An important objective of the interview was to validate interpretation of students' written answers on the conceptual assessment and to use results of this analysis to develop a rubric for scoring the conceptual assessment.

Another objective was to add depth to the investigation and analysis of the research sub-questions. Based on these objectives, an interview protocol was developed to probe student thinking related to the conceptual assessment.

Interview Protocol

The interview consisted of distinct, predetermined questions designed to probe the individual responses that the students had provided on the written conceptual assessment. Before the interview began, the purpose of the interview was explained to each student. The interviewees were each told that they would be asked questions about their thinking and that it would be very helpful for them to think out loud. The interviewees were also told that during the interview they would be asked to clarify or explain their thinking. After the introduction, the flow of the interview followed a general outline that included four phases. As the interview progressed, questions were adapted based on the interviewee's responses. During the first phase, the interviewee was presented with his written work on the first problem in the conceptual assessment, the hospital problem. During this phase, the interviewee was asked specific questions about his written work on the hospital problem and about his thinking and reasoning about that work. In the second phase of the interview, the interviewee was asked about his thinking regarding particular notions of probability distribution in relation to the hospital problem. At this point in the interview, the hospital problem was put away, and the interviewee was presented with his work on the second problem in the conceptual assessment, the theater problem. During the third phase of the interview, the interviewee was asked specific questions about his written work on the theater problem and about his thinking and reasoning about that work. The fourth phase of the interview consisted of general questions about the student's conceptions of notions across constructs which comprise probability distribution, with a particular focus on the construct of distribution.

Phase 1 of the interview. At the beginning of phase 1, the interviewee was shown his work on the hospital problem. The first question that all interviewees were asked was: How do you conceive of probability in this problem? Next, the interviewee was asked to explain how he used his understanding of probability to solve the problem. This question was followed by specific, individualized questions which were based on the each student's original written response to the problem and were aimed at probing the students' thinking about the probability in the context of the problem. Students were asked to explain what they meant when they wrote particular words such as average, expected percentage, skewed, bias, normal, probability of error, proportional variance, and "statistically susceptible." The general format of these questions was:

- You mentioned _____________. Tell me how you were thinking about __________.
- Please explain what you mean by _____________.
- Why are you thinking that ___________is important?
- How is _______________ relevant to this situation?

Students were also asked questions addressing the notions of sample size, sampling, and expectation and variability.

- You mentioned sample size (or, another student mentioned that sample size was important). How might thinking about sample size be important to this situation?
- You mentioned sampling (or, another student mentioned that sampling was important). How might thinking about sampling be important to this situation?

- You mentioned expectation (or, another student talked about expectation when she answered this problem). Why might thinking about expectation be important in this situation?

Next, the interviewees were asked to articulate their thinking about any processes that they thought were relevant to the problem and to talk about any assumptions or conditions that they were thinking about when solving the problem.

- Explain any processes that might be relevant to this problem.
    - You talk about______________. How are you thinking of this process?
    - What are the results of this process?
- Please tell me about any underlying assumptions or conditions you were thinking about when solving this problem.

Some of the interviewed students appeared to indicate thinking about distribution on their written responses to the problem, while others did not. All interviewees were asked about the notion of distribution and why it might be important to think about distribution in relation to the problem.

- You mentioned distribution (or, another student mentioned that thinking about distribution was helpful). Why is distribution important to your thinking? (Or, why might it be helpful to think about distribution for this problem?)

Possible follow-up questions were:

- Explain what distribution means to you in this problem.
- Draw a picture of the distribution you are thinking of.
- What aspects of distribution could be related to thinking about this problem?

Phase 2 of the interview. After answering questions aimed at probing stochastic conceptions of probability specifically related to their written work on the hospital problem, students' understanding about probability distribution was probed in phase 2. Each interviewee was asked to articulate thinking about specific ideas across constructs defined in the framework for understanding probability distribution: probability, variability, and distribution (see figure 5.x). The students were asked to explain how thinking about relative frequencies, independence, sample space, random variable, and sampling related to the hospital problem. In order to probe thinking about coordination of empirical and theoretical probability, each interviewee was asked what he remembered learning about the law of large numbers and then asked how the law of large numbers was related to his thinking about the hospital problem. Questions asked in phase 2 of the interview were:

- How do you think the following probabilistic notions might relate to this problem: relative frequencies, independence, sample space, random variable, and sampling?
- You talked about the law of large numbers (or, tell me what you recall about the law of large numbers). How is the law of large numbers related to your thinking about this problem?

Phase 3 of the interview. The third phase of the interview began when the interviewee was presented with his written work on the theater problem. First, students were asked to explain how they figured out the answer to this problem. The word "unusual" was used in the problem stem. Student were asked what unusual meant to them, and what the word meant in the context of this problem.

- How did you figure out your answer to this problem?

- What does unusual mean to you?
- Give an example of unusualness.
- How are you thinking about the notion of unusual for this problem?

Next, students were asked individualized questions that were based on their original written response to the problem. This questioning followed a general format.

- You mentioned __________. Tell me how you were thinking about __________.

This format also included follow-up questions such as:

- Tell me how you were thinking __________relates to the problem.
- Please explain what you mean by _____________.
- Why are you thinking that _________is important?
- How is _______________ relevant to this situation?

Students were asked how they conceived of probability in the context of this problem.

They were also asked about their thinking related to any relevant processes, in addition to any assumptions and conditions they made when answering the problem.

- How do you conceive of probability in this problem?
- Explain any processes that might be relevant to this problem.
- Please tell me about any underlying assumptions or conditions you were thinking about when solving this problem.

In addition, all of the interviewees were asked to articulate their thinking about distribution in the context of the theater problem. They were also asked to explain their thinking about outcomes in this problem.

- You mentioned distribution (or, another student mentioned that thinking about distribution was helpful). Why is distribution important to your thinking?

- Explain how you are thinking about "outcomes" in this situation. What are the outcomes in this situation?

Phase 3 of the interview closed with a question for students who had completed the stochastic reasoning lab assignments. This question was only posed to students who had completed the stochastic labs and who had not spontaneously talked about the influence of the labs earlier in the interview. This question was: How did your work on the labs influence your thinking about the problems on the extra-credit quiz? The purpose of this question was to prompt the stochastic lab group students to think about the two conceptual quiz problems in relation to the lab assignments. Because calculus content was not addressed in the conceptual quiz items, it did not make sense to ask the calculus lab students about the influence of the calculus labs on their thinking about the quiz problems.

Phase 4 of the interview. During the final phase of the interview, students were asked more general questions about their understanding of major concepts related to probability distribution. The first two questions asked students to articulate their thinking about randomness and the law of large numbers.

- How is probability related to randomness?
- What is the meaning of the law of large numbers?

These questions were followed by questions that probed student thinking about the construct of distribution in relation to probability distribution and sampling distribution.

- What comes to mind when you think of distribution?
- What are the important aspects of probability distribution?

- What comes to mind when you think of sampling distribution?
- What are the important aspects of sampling distribution?
- Explain any relationships you conceive between sampling distribution and probability distribution.

Notions about sampling and sampling distribution were probed by posing a problem from the ARTIST Sampling Variability Scale. Students were presented with a diagram representing a hypothetical distribution for a population along with for graphs representing possible distributions of sample means for random samples drawn from the population (see Appendix F). The interviewees were asked to respond to the following questions and explain their reasoning.

- Which graph best represents a distribution of sample means for 1000 samples of size 4? Explain.
- What do you expect for the shape of the sampling distribution for n = 4? Explain.
- Which graph best represents a distribution of sample means for 1000 samples of size 50? Explain.
- What do you expect for the shape of the sampling distribution for n = 50? Explain.

Next, the interviewees were asked an overarching question: What do you think is the relationship between probability and statistics? The interview closed with a question that asked the interviewee to offer an opinion of his learning experience with the labs.

Interview Participants

Participant selection. A two-problem assessment investigating the conceptions of probability held by students after a semester of instruction was administered in the form of an extra-credit quiz during a regularly scheduled discussion session associated with the course. All students had the opportunity to take the assessment. Interview participants were selected based on their written responses to a conceptual assessment. A description of the selection criteria follows.

Students' written answers to the two problems on the conceptual assessment were analyzed using Liu and Thompson's (2007) framework for probabilistic understanding. This analytic frame characterizes thinking which follows chains of reasoning along paths that may result in a stochastic conception. Liu and Thompson (2007) describe the path leading to a stochastic conception of probability as having an image of a repeatable process, specifying conditions of the process, and having an image of distribution of outcomes. Using this framework, student responses to the two quiz problems were coded as either indicating a stochastic conception of probability or not indicating a stochastic conception of probability. Thus, for the purpose of a preliminary analysis, stochastic conception was coded as a binary variable.

Two weeks prior to the end of the course, an email message was sent to all student participants asking for volunteers to participate in interviews associated with the study. Students were told the interview would last approximately 1 hour and take place at the close of the semester, after the course's final examination. There were 23 students who indicated willingness to participate in an interview, and 12 students were selected. Criteria used to select the 12 interviewees included balanced representation in terms of lab group identification and, within each of the lab groups, distribution of performance on the written conceptual assessment.

Participant description. Six interview participants were selected from the calculus review lab group and 6 were selected from the stochastic reasoning lab group. Table 3.10 gives a summary of characteristics of the interview participants. This information was compiled from survey data submitted by each student at the beginning of the course.

The student number in the first column of the table denotes the study identification number assigned each student. All students in the study were assigned to one of two groups for the purpose of receiving particular lab materials. Students' lab group assignment is provided under the column heading "Group." The calculus review lab group is denoted by "CR" and the stochastic reasoning lab group is denoted by "SR."

The course in which the study took place was designed for students in science, engineering, and mathematics, and course of study of the interview participants represented a variety of majors. Students' majors along with their class standing are provided in the third and fourth columns of the table. Students self-reported their cumulative grade point average within categories of ranges, and these are given in the column labeled "GPA." The self-reported number of college mathematics earned prior to enrollment in the course is given under the heading "Math Credits." Enrollment in the course did not require a prerequisite course in statistics, however the background survey asked students to report whether or not they had taken a college-level statistics course to enrollment in the course in which the study took place. Under the column labeled "Prior

Table 3.10

Characteristics of Interview Participants

Student Number	Group	Major	Class	GPA	Math Credits	Prior Statistics	Calculus	SAT Math
214	CR	Computer Science	Junior	2.5 – 2.9	12	No	2	770
428	CR	Mathematics	Junior	3.5 – 4.0	34	AP Statistics	2	770
518	CR	Computer Science	Junior	3.0 – 3.4	12	No	1	730
711	CR	Social Science	Sophomore	3.0 – 3.4	12	AP Statistics	2	740
715	CR	Biology	Senior	2.5 – 2.9	17	Yes	1	NR
1016	CR	Economics	Senior	2.5 – 2.9	18	Yes	1	600
113	SR	Finance	Junior	3.5 – 4.0	15	Yes	1	640
313	SR	Engineering	Junior	2.5 – 2.9	12	No	2	730
325	SR	Biology	Junior	2.5 – 2.9	11	No	1	550
329	SR	Engineering	Junior	3.0 – 3.4	15	AP Statistics	1	680
811	SR	Engineering	Sophomore	3.0 – 3.4	21	AP Statistics	3	790
817	SR	Engineering	Sophomore	3.5 – 4.0	7	No	3	800

Statistics," "Yes" denotes that the student reported earning credits for a prior college-level statistics course, "AP Statistics" denotes that the student reported earning college credits through the AP Statistics examination, and "No" denotes that the student reported he had not taken a prior college-level statistics course. A prerequisite for the course was completion of two semesters of calculus. Some students reported that they had earned calculus credits via AP Calculus examinations, and some students completed calculus courses while enrolled in college. Interview participants' background in calculus is denoted by the code 1, 2, or 3: students who completed two semesters of calculus in college are coded as "1," students who completed one semester of calculus via AP Calculus examinations and completed one semester of calculus in college are coded as "2," and students who completed both semesters of calculus via AP Calculus examinations are coded as "3." The final column contains students' self-reported SAT Math score. One of the interview participants did not report her SAT Math score, and this is denoted as "NR," not reported.

Three sophomores, seven juniors, and two seniors participated in the interviews. The interview participants reported overall cumulative grade point averages that ranged from 2.5 to 4.0. These students entered the course in which the study took place having earned between 7 and 34 college credits in mathematics. The student who entered the course with 34 credits of mathematics coursework was a mathematics major. Four of the interviewed students were engineering majors, two were biology majors, and two were computer science majors. The remaining three students included a social science major, a finance major, and an economics major. Seven interview participants had earned credit for a prior college-level statistics course, and four of these seven students had earned college credit via the AP Statistics examination. Five of the interview participants had not taken a prior statistics course at either the college or high school level.

Analytic Framework for Interview

The purpose of the interviews was to investigate students' stochastic conceptions of probability and to describe students' conceptions of probability distribution. Liu and Thompson (2007) stress that there are different interpretations of probability statements and situations: "A situation is not stochastic in and of itself. It is how one conceives of a situation that makes it stochastic or nonstochastic" (p. 126). The two problems on the conceptual assessment were selected not only because they involved different contexts, but also because each problem represents a probabilistic situation that can be approached stochastically or nonstochastically. Thus, it was anticipated that students' responses to these two conceptual quiz items could reveal differing interpretations of probability. When developing the analytic framework with which to examine students' responses, the focus was to characterize students' thinking in terms of their stochastic conceptions of probability and probability distribution. The analytic framework developed for analysis of the interview data is comprised of three categories which were described by Liu and Thompson (2007) as three constructs that lead to a stochastic conception of probability: image of a repeatable process, specification of conditions of the process, and image of a distribution of outcomes (see Table 3.11). The categories in the framework represent a standard that was only applied to the interview data as the researcher's questioning during the interview permitting probing to ascertain whether a student's understanding was both correct and complete. This evaluation standard was not applied to the written conceptual assessment ("quiz") responses submitted by the entire sample because the absence of a component in an open-ended response means no evidence of understanding, but not an absolute determination of a lack of understanding.

Table 3.11.

Categories of Analytic Framework for Analysis of Interviews

Category	Stochastic Conception
Image of a Repeatable Process	• Conceiving of probability situation as the expression of a stochastic process. • Taking for granted that the process could be repeated under similar conditions
Specifies Conditions of the Process	• Taking for granted that the conditions of and implementation of the process would differ among repetitions in small yet important ways. • Anticipating that repeating the process would produce a collection of outcomes
Image of a Distribution of Outcomes	• Anticipating the relative frequency of outcomes will have a stable distribution

Note: Analytic framework adapted from "Teachers Understandings of Probability," by Y. Liu and P. Thompson, 2007, *Cognition and Instruction*, 25, p. 122.

The first category in the analytic framework addresses the second research sub-question: When confronted with a probability situation do students present an image of a repeatable process? An image of a repeatable process is the first construct that Liu and Thompson (2007) describe along the path towards a stochastic conception of probability. Having an image of repeatable process is defined as "…conceiving of a probability situation as the expression of a stochastic process … taking for granted that the process could be repeated under essentially the same conditions" (Liu & Thompson, 2007, p. 22). Thus, a key notion related to stochastic processes is repetition. A stochastic conception of the process entails conceiving of a single outcome as but one expression of the underlying process.

Understanding the conditions that produced the outcome is an important aspect of a stochastic understanding of probability. This second category in the analytic framework addresses the third research sub-question: When confronted with a probability situation, are students able to specify the conditions of the process? Specifying the conditions of the process is the second construct in Liu and Thompson's (2007) chains of reasoning leading toward a stochastic conception. Specification of conditions of the stochastic process implies recognition that "the conditions and implementations of the process would differ among repetitions in small, yet important ways," as well as "anticipating that repeating the process would produce a collection of outcomes" (Liu & Thompson, 2007, p. 22).

Key notions related to specific conditions of a stochastic process involve understanding the nature of random phenomena. Each repetition of the process will result in an outcome and that outcome is one expression of the underlying process. Assuming a finite number of repetitions, one condition of the random process is that the outcomes of the process are equally likely. Another condition is that the outcomes are independent. Furthermore, the result of the process is not necessarily the same each time the process is repeated, so the outcomes produced by the process will vary. However, this variability is not haphazard. In the short run, observations of outcomes produced by the random process will exhibit much more variability than in the long run. In other words, for a large number of repetitions less variability in the outcomes is expected. This is the notion described by the law of large numbers.

A stochastic image of a distribution of outcomes means "anticipating that the relative frequencies of outcomes of the stochastic process will have a stable distribution in the long run" (Liu & Thompson, 2007, p. 22). Associating a numerical value with some characteristic pertaining to the outcome of a stochastic process gives a random variable. Repetition of the random process produces a collection of outcomes, which results in a collection of values of a random variable. A list of all possible values of the random variable and associated relative frequencies produces a distribution. Thus, repetition of the stochastic process gives relative frequencies associated with each value of the random variable, and these relative frequencies are empirical probabilities associated with each value of the random variable. In the long run, the relative frequencies will stabilize as the stochastic process is repeated many times.

A stochastic image of distribution involves an application of the law of large numbers in the context of a particular situation. In addition to understanding that the stochastic process will produce random variability in outcomes, this image of distribution requires thinking about the nature of the variability in outcomes and understanding that the variability is not haphazard. Although there may be considerable variability in the short run, the relative frequencies of the outcomes (empirical probabilities) will settle down over a large number of repetitions to a value that is very close to the theoretical probability. Thus, a stochastic image of distribution of outcomes supports thinking about an empirical probability distribution in relation to a theoretical probability distribution model.

Summary

The study outlined in this chapter was an investigation of an instructional intervention designed to support mathematically advanced college students' stochastic understanding of probability. The control-treatment design was balanced across lectures and discussion sections, and the study employed a mix of quantitative and qualitative research methods to examine students' stochastic understandings of probability distribution resulting from an instructional intervention. Each phase of this study, the quantitative phase and the qualitative phase, was used to address the research question and sub-questions and offer additional insight into students' understandings of probability.

CHAPTER 4: QUANTITATIVE RESULTS AND ANALYSIS

The results for the quantitative portion of the study are presented and analyzed in this chapter. Quantitative data were collected from 184 students enrolled in a calculus-based introductory probability and statistics course at a large, public university. The number of students included in the sample used for analyses presented in Chapter 4 was 145. The analytic sample included 71 students in the treatment group and 74 students in the control group. Students in the treatment group engaged in supplementary lab assignments aimed at supporting development of stochastic reasoning (SR). Students in the control group engaged in supplementary lab assignments consisting of review of calculus content required for the course (CR). The data collected consisted of a student background survey, a conceptual assessment of stochastic reasoning, an assessment measuring understanding of confidence intervals, and final course examinations. Analysis of these data yielded information that addressed the following research question: What is the impact of an instructional intervention designed to support the development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory, calculus-based, probability and statistics course? The study also addressed four research sub-questions. The qualitative analysis presented in Chapter 5 addresses research sub-question 1. Analyses of quantitative data presented in Chapter 4 addresses research sub-questions 2, 3, and 4.

- Research sub-question 2: Does instruction designed to support development of stochastic understanding of probability distribution impact students' stochastic conceptions of a probability situation as evidenced on a conceptual assessment?

- Research sub-question 3: Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding of confidence intervals as measured by the ARTIST assessment?
- Research sub-question 4: Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding as evidenced on final course examinations administered in an introductory calculus-based probability and statistics course?

This chapter is organized into three sections. The first section gives a description of the analytic sample used for analyses presented in this chapter. The second section discusses regression modeling and analysis of data collected during administration of the stochastic reasoning assessment, the understanding of confidence interval assessment, and the final course examination. The last section is a summary of quantitative findings addressing the main research question and three research sub-questions.

Description of Analytic Sample

The participants in the study were 184 students enrolled in one of two large-lecture classes for an introductory, calculus-based, probability and statistics course. This course was configured with weekly discussion sessions as well as weekly lecture sessions. A description of study participants is found in Chapter 3. The analytic sample used for the analyses presented in Chapter 4 is a subset of study participants consisting of 145 students. There were 155 students who completed the student survey, and 168 students who completed a conceptual assessment of stochastic reasoning in the form of a quiz. There were 145 students who completed the stochastic reasoning assessment, the student survey, and the final course examination. These are the students who comprise the analytic sample. Of the 145 students in the analytic sample, 124 students completed the confidence interval assessment. Study participants not included in the analytic sample were those students who did not complete each of the survey questions pertaining to variables used in the quantitative analyses or those who were absent the day the stochastic reasoning quiz was administered, yielding missing data.

Variables describing students' characteristics include students' majors, GPA range, and class standing were garnered from data self-reported by students in the student survey administered at the beginning of the course. Two semesters of college calculus was a prerequisite for this course, and the variable labeled CalcI summarizes what students reported as the manner in which they earned credits for their first semester of college calculus. The student survey also asked students to report if they had taken a statistics course prior to enrollment in the course in which the study took place, and this information is indicated by the variable PriorStat which labeled the prior statistics background of students. The variable Lecture indicates the lecture class that the student was enrolled in.

Table 4.1 gives a comparison of characteristics for all study participants with those students who comprised the analytic sample. This table shows that the distributions for major, GPA range, class standing, first-semester-college calculus, and prior statistics course for all study participants are similar to the distributions for the analytic sample. The majority of the students in the analytic sample were computer science majors (39.3%). A large majority of the students reported a cumulative GPA at or above 3.0 (73.1%), with 40% of students reporting a cumulative GPA of at least 3.5. Most of the students were either juniors or seniors (78.6%). A little more than half of the students completed their first semester of calculus while enrolled in college (56.6%), while 43.4% of students earned credits for their first semester of college calculus via the AP Calculus examination. Although a number of students had earned credits for a previous college-level statistics course, the course in which the study took place was the first college-level course in statistics for the majority of students in the analytic sample (58.6%). Distributions of the characteristics noted in Table 4.1 indicate that students comprising the analytic sample were representative of all study participants.

Table 4.1

Comparison of Characteristics for All Study Participants and Analytic Sample

	Study Participants			
	All Participants		Analytic Sample	
Characteristic	*n*	%	*n*	%
Major				
Computer science	62	39.5	57	39.3
Business; economics; other	33	21.0	32	22.1
Engineering; physical sciences	24	15.3	19	13.1
Mathematics	20	12.7	20	13.8
Biology; chemistry	18	11.5	17	11.7
Total reported	157	100	145	100
GPA range				
3.5 – 4.0	61	38.9	58	40.0
3.0 – 3.4	50	31.8	48	33.1
2.5 – 2.9	32	20.4	30	20.7
2.0 – 2.4	14	8.9	9	6.2
Total reported	157	100	145	100
Class standing				
Senior	52	33.1	46	31.7
Junior	71	45.2	68	46.9
Sophomore	32	20.4	30	20.7
Freshman	2	1.3	1	0.7
Total reported	157	100	145	100
First-semester, college calculus				
College course	89	56.7	82	56.6
AP Calculus	68	43.3	63	43.4
Total reported	157	100	145	100
Prior Statistics				
College course	31	20.0	29	20.0
AP Statistics	34	21.9	31	21.4
No prior college-level course	90	58.1	85	58.6
Total reported	155	100	145	100
Lecture				
Lecture A	84	45.7	68	46.9
Lecture B	100	54.3	77	53.1
Total	184	100	145	100

Characteristics of Treatment/Control Groups for Analytic Sample

The study employed a treatment/control design whereby students were assigned to either a treatment or control group for the purposes of an instructional intervention. The instructional intervention consisted of supplementary lab assignments, which students completed outside of class. Students in the treatment group engaged in supplemental lab assignments aimed at development of stochastic reasoning (SR), while students in the control group engaged in lab assignments that reviewed calculus topics used in the course (CR). Table 4.2 provides a comparison of characteristics for the analytic sample across the treatment/control groups. Counts within subgroups for each of the characteristics were examined as a percentage of the number of students within the SR group and as a percentage of the number of students within the CR group. This examination showed similar subgroup percentages for the SR group and the CR group for the characteristics class standing, prior statistics, and lecture class. The differences in proportion of students in each major between the SR group and CR group were not statistically significant, χ^2, (4, N = 145) = 8.92, p = .063, and neither were the differences in the proportion of students who had completed their first-semester of calculus via AP Calculus versus a college course, χ^2, (1, N = 145) = 2.98, p = .084. The GPA categories 2.5 – 2.9 and 2.0 – 2.4 were combined and a Chi-square test for homogeneity was performed to examined the differences in GPA between the SR group and CR group. The test revealed that these differences were not statistically significant, χ^2, (2, N = 145) = 2.38, p = .305.

Table 4.2

Characteristics of Analytic Sample by Treatment (SR) and Control (CR)

	Lab Assignment Group			
	SR		CR	
Characteristic	*n*	%	*n*	%
Major				
Computer science	20	28.2	37	50.0
Business; economics; other	16	22.5	16	21.6
Engineering; physical sciences	13	18.3	6	8.1
Mathematics	12	16.9	8	10.8
Biology; chemistry	10	14.1	7	9.5
GPA range				
3.5 – 4.0	31	43.7	27	36.5
3.0 – 3.4	25	35.2	23	31.1
2.5 – 2.9	12	16.9	18	24.3
2.0 – 2.4	3	4.2	6	8.1
Class standing				
Senior	22	31.0	24	32.4
Junior	35	49.3	33	44.6
Sophomore	14	19.7	16	21.6
Freshman	0	0.0	1	1.4
First-semester, college calculus				
College course	35	49.3	47	63.5
AP Calculus	36	50.7	27	36.5
Prior Statistics				
College course	12	16.9	17	23.0
AP Statistics	15	21.1	16	21.6
No prior college-level course	44	62.0	41	55.4
Lecture				
Lecture A	37	52.1	31	41.9
Lecture B	34	47.9	43	58.1
Total	71	100	74	100

The instructional intervention for both the stochastic reasoning group and the calculus review group required students to submit written lab assignment work. After submitting the lab assignments, students in the SR group received explicit instruction linking stochastic concepts explored in the lab assignment with specific content in the course. Similarly, after submitting CR lab assignments, the students in the CR group received explicit instruction linking the calculus review problems to the types of problems encountered in the curriculum for this probability and statistics course. The instructional interventions for both groups required comparable time to complete assignments.

Lab assignments for the treatment group (SR) were specifically designed to undergird development of stochastic reasoning. The SR labs were carefully structured to connect probabilistic learning outcomes within and across each phase of the hypothetical learning trajectory described in Chapter 3. Because the lab assignment tasks were anticipatory tasks (Simon, 2013) designed to engage students in coordinating their thinking about complementary probabilistic notions along the hypothetical learning trajectory and to set the stage for deeper understanding of content and explanations offered by their professors and discussion teachers, SR students needed to engage with all lab assignments tasks in order to complete the *treatment* comprising the instructional intervention.

Students who completed and submitted all written work associated with all lab assignment tasks were coded as having completed all labs which is indicated by the variable labeled AllLabs. Thus students in the SR group who completed all of the assigned tasks along the hypothetical learning trajectory and submitted their written work for these labs were coded as having completed all labs (AllLabs = 1). Similarly, students in the CR group who completed all of the calculus review problems and submitted their written work were coded as having completed all labs (AllLabs = 1). Students in the SR group, as well as students in the CR group, who did not submit complete written work for all lab assignments were coded as not having completed all labs (AllLabs = 0). Table 4.3 shows the distribution of lab assignment completion for the treatment group (SR) and control group (CR). A slightly higher percentage of students in the SR group completed all of the lab assignments tasks than students in the CR group.

Table 4.3

Lab Assignment Completion for Analytic Sample

	Lab Assignment Group			
	SR		CR	
Lab Completion	*n*	%	*n*	%
All labs completed	55	77.5	48	64.9
Not all labs completed	16	22.5	26	35.1
Total	71	100	74	100

Measurements for Analytic Sample

Data for this study were acquired via three measurements: a conceptual assessment of stochastic reasoning, an assessment measuring students' understanding of confidence intervals, and the final course examination. The conceptual assessment of stochastic reasoning consisted of two items administered in the form of an in-class quiz. The assessment of students' understanding of confidence intervals consisted of 12 items drawn from ARTIST topic scales (Garfield, et al., 2006). Comparison of summary statistics for the three assessments across treatment/control groups is provided in Table 4.4.

Table 4.4

Mean and Standard Deviations from Measures of Student Understanding by Treatment (SR) and Control (CR) Groups

Data Source	*n*	M	SD
Stochastic reasoning quiz			
CR	74	1.65	2.29
SR	71	3.87	3.72
Confidence interval assessment			
CR	64	7.98	2.09
SR	60	7.95	2.14
Final examination[a]			
CR	74	0.18	0.94
SR	71	0.10	0.89

[a]Final examination scores were standardized independently for each lecture section across the available sample: Lecture A (M = 78.3, SD = 13.6), Lecture B (M = 98.1, SD = 24.6). Standardized final examination scores were used for analytical purposes.

Stochastic reasoning assessment scores. A description of the two problems which comprise the stochastic reasoning conceptual assessment is provided in Chapter 3. Students' written answers to the conceptual quiz problems were analyzed and scores were assigned to each student's written work using a rubric that was developed using evidence gleaned from analysis of the participant interviews presented in Chapter 5. The rubric shown in Table 4.5 was used to assign points for various aspects of images of stochastic reasoning that were evidenced in students' written answers on the conceptual quiz. Development of this rubric is described in Chapter 3. Not all aspects for each image were evidenced in both problem contexts. For example, an indication that variability in outcomes is related to sample size was only evidenced in the hospital problem context, and an indication that the sampling process produces samples that are representative of the population was only evidenced in the theater problem context. Students received one point for each image aspect revealed by their written work on the conceptual quiz. Up to seven points were awarded for stochastic image aspects evidenced in the hospital problem context and up to seven points were awarded for stochastic image aspects evidenced in the theater problem context. Thus, the maximum number of points possible for the stochastic conception assessment was 14.

Table 4.5

Stochastic Conception Rubric

	Possible Points	
Stochastic Conception	Hospital	Theater
Image for Repeatable Process		
Indicates the repeatable process yields outcomes	1	1
Indicates that repetition of process results in repeated sampling	1	1
Indicates connecting process to a model	1	–
Image for Specification of Conditions		
Indicates that repetition of the process yields a collection of variable outcomes	1	1
Indicates the sampling process produces samples that are representative of population	–	1
Indicates that variability in outcomes is related to sample size	1	–
Images for Distribution of Outcomes		
Connects expectation to variability	1	1
Indicates underlying distribution model	1	1
Quantifies "unusual" *as deviation from expectation in terms of a distribution modal*	–	1

The overall distribution of scores on the stochastic reasoning conceptual assessment ranged between 0 and 13 with a median score of 2. This distribution was highly skewed (M = 2.74, SD = 3.26). The distribution of scores for each of the treatment and control groups was also skewed. Students in the SR group had a median stochastic score of 3, while students in the CR group had a median stochastic score of 0.5. The mean stochastic score for the SR group (*M* = 3.87, *SD* = 3.718) was significantly higher than the mean stochastic score for the CR group (*M* = 1.65, *SD* = 2.290); [$t(143) = 4.36$, p <.000].

Confidence interval assessment scores. The confidence interval assessment consisted of 12 multiple choice items and was administered online via *Blackboard*. Items used as the confidence interval assessment for this study were drawn from ARTIST topic scales (Garfield, et al., 2006). Ten of the 12 items were the items which comprise the ARTIST Confidence Interval Scale. The two additional items were drawn from the ARTIST Sampling Distribution Scale. Students received one point for each item answered correctly on the confidence interval assessment for a total of 12 possible points. The distribution of scores for students in the analytic sample ranged between 8 and 12 and was fairly symmetric (*M* = 7.97, *SD* = 2.103). The average confidence interval assessment score was approximately the same for the SR group (*M* = 7.95, *SD* = 2.135) and the CR group (*M* = 7.98, *SD* = 2.089); [$t(122) = 0.09$, $p = .808$].

Final examination scores. Comprehensive final examinations were developed independently by the professors for each lecture class. Questions on the final course examinations differed, as well as the total number of points possible on the final examination for each lecture class. Final examination scores were standardized independently for each lecture section across the available sample: Lecture A (*M* = 78.3, *SD* = 13.6), Lecture B (*M* = 98.1, *SD* = 24.6). Standardized final examination scores were used for analytical purposes. Investigation of standardized scores for treatment/control groups revealed that the distributions were approximately normal. The distribution of standardized final examination scores for the SR group (*M* = 0.096, *SD* = 0.886) and the CR group (*M* = 0.181, *SD* = 0.938) were similar; [$t(143) = 0.56$, $p = .906$].

Regression Analyses

Statistical analysis of stochastic conceptions was completed using logistic regression modeling techniques. The analysis resulted in two models with StochImage as the outcome variable. Controls and predictor variables included lab assignment group, mathematical background, prior statistics course, lecture class, and lab assignment completion. Statistical analysis of students' understanding of confidence intervals and students' understanding as evidenced on the final course examination was accomplished using multiple regression modeling techniques. The analyses resulted in two models: one model with the confidence interval score as the outcome variable and the other model with the standardized the final exam score as the outcome variable. The explanatory variables for multiple regression analyses were the same explanatory variables used for the logistic regression analysis modeling StochImage.

Descriptions of Variables

Scores for the variable StochImage were determined through analysis and scoring of students' written work on the stochastic reasoning conceptual assessment. In addition to receiving a stochastic score for work the on stochastic reasoning conceptual assessment, students' responses were categorized according to evidence of stochastic images they presented on the written work. The images which comprise the stochastic categories are hierarchical, thus the stochastic image categories are hierarchical. As described in Chapter 3, evidence of stochastic reasoning required an image of a repeatable process, specification of condition of the process, and an image of distribution of outcomes. For students' responses to be characterized as *stochastic*, the student had to present evidence of stochastic reasoning for both problem contexts on the conceptual assessment. Students who presented evidence of stochastic reasoning on their written work for exactly one of the two conceptual assessment contexts were characterized as *situational.* A *nonstochastic* characterization indicated that the student may have presented an image of a repeatable process and may have also specified conditions of a repeatable process for one or both problem contexts on the conceptual assessment, but did not present evidence of all images that comprise stochastic reasoning for either problem context. Students who did not present any images related to a stochastic conception for either problem context were characterized as having *no image*. A comparison of frequency distributions for stochastic image category for the treatment and control groups is presented in Table 4.6. This table shows that more students in the stochastic reasoning lab group presented evidence of stochastic reasoning for one or both problem contexts on the conceptual assessment than did students in the calculus review lab group.

Table 4.6

Frequency Distribution of Stochastic Reasoning Category for Treatment and Control Groups

	Lab Assignment Group			
	SR		CR	
Stochastic Reasoning Category	*n*	%	*n*	%
No Image	21	29.6	37	50.0
Nonstochastic	22	31.0	27	36.5
Situational	16	22.5	7	9.5
Stochastic	12	19.9	3	4.1
Total	71	100	74	100

The dependent variable which measures stochastic reasoning is StochImage. StochImage is a binary variable indicating whether or not a student presented evidence of stochastic reasoning on their written work for the stochastic reasoning conceptual assessment, i.e. conceptual quiz. Students who were categorized as stochastic or situational presented evidence of stochastic reasoning for one or both problem contexts and were coded as having a stochastic image. Students who were categorized as nonstochastic or no image were coded as not having a stochastic image.

Information on variables used in the regression analyses for this study is presented in Table 4.7. The table is organized by outcome variables and explanatory variables. The table lists the variable labels, the variable names, descriptions of each variable, and the type of data represented. The type of data clarifies the coding for the dichotomous variables. The outcome variable used for logistic regression analyses was StochImage. The outcome variables used for multiple regression analyses were ConfidInt and

FinalExam. Confidence interval score represents the total points scored on the confidence interval assessment. Comprehensive final examinations were developed independently by the professors for each lecture class. Scores on the final examination for each lecture class were standardized for use in this analysis and represented by FinalExam. Predictor variables for all models were treated as dichotomous indicators.

Table 4.7

Variable Descriptions for Regression Models

Variable Label	Variable Name	Description	Coding
Outcome Variables			
Stochastic image	StochImage	Variable indicating evidence of stochastic reasoning on stochastic conception assessment	0 = Stochastic image 1 = No stochastic image
Confidence interval score	ConfidInt	Total score for confidence interval assessment	Continuous
Final exam Score	FinalExam	Standardized score for final course examination	Standardized and continuous
Explanatory Variables			
Lab group	Group	Variable indicating treatment/control group	0 = Calculus review 1 = Stochastic reasoning
First-semester, college calculus	CalcI	A student demographic variable: dummy variable indicating source of credits earned for first-semester calculus	0 = College course 1 = AP Calculus
Prior statistics	PriorStat	A student demographic variable: dummy variable indicating whether the student had earned credits for a prior college-level statistic course	0 = No prior college-level statistics course 1 = Prior college-level statistics course
Lecture	Lecture	Variable indicating enrollment in large-lecture class	0 = Lecture B 1 = Lecture A
All labs completed	AllLabs	Variable indicating completion of all written lab assignments	0 = Not all labs completed 1 = All labs completed

The variable Group is a binary indicator of whether the student was assigned to the treatment group and received stochastic-reasoning lab assignments or was assigned to the control group and received calculus-review lab assignments. The variable lecture is also binary and indicates which of the two lecture classes the student was enrolled in.

A variable indicating the source of credits students' earned for first-semester, college calculus was chosen as a proxy for mathematical background. Analysis of student demographic data revealed that students who earned their first-semester college calculus credits via AP Calculus reported higher college grade-point averages. Furthermore, students who earned first-semester college calculus credits via AP Calculus had statistically significant higher SAT Math scores. Students with higher SAT Math scores also reported higher college grade-point averages. Studies indicate that AP course taking is related to achievement and success in college (College Board, 2007; Klopfenstein & Thomas, 2006). Thus, the variable indicating first-semester college calculus was used to control for prior mathematics knowledge and college achievement. The binary variable CalcI indicates the source of credits earned for calculus. Students who earned credit for first-semester, college calculus via AP Calculus were coded 1; otherwise students were coded 0 indicating they earned credit for first-semester, college calculus by taking a course in college.

PriorStat was treated as a dichotomous variable and used to control for prior statistics knowledge. Students who reported they had either taken AP Statistics or enrolled in a statistics course while in college and earned credits for the course, were coded as 1, meaning they had received credits for a prior college-level statistics course. Students who had not earned credits for a prior college-level statistics course were coded as 0.

The variable AllLabs is an indicator of students who completed and submitted all written work associated with all lab assignment tasks. Students received completion points for submitting written work for each lab assignments. Those who submitted complete written work for all six lab assignments, regardless of treatment/control group, were coded as 1, meaning they had completed all of the lab assignments.

Analysis of Stochastic Conception Models and Results

Statistical analysis of stochastic conceptions was completed using logistic regression modeling techniques. The analysis resulted in two models with stochastic image as the outcome variable. Controls and predictor variables included lab assignment group, mathematical background, prior statistics course, lecture class, and lab assignment completion.

Comparisons of frequencies for the dependent variable stochastic image versus no stochastic image are shown in Table 4.8 for each of the explanatory variables. The two right-hand columns provide Chi-square statistics for each comparison. An investigation of each of the predictor variables indicated that students in the stochastic reasoning lab group were significantly more likely to present evidence of having a stochastic image than students who were in the calculus review lab group, $\chi^2(1, n = 145) = 12.59, p < .001$. Also, students who had earned first-semester college calculus credits via AP Calculus were significantly more likely to have a stochastic image, $\chi^2(1, n = 145) = 4.37, p = .036$. Furthermore, there was a significant relationship between stochastic image and lecture class enrollment. Students enrolled in Lecture A were significantly more likely to have a stochastic image, $\chi^2(1, n = 145) = 17.90, p < .001$.

Table 4.8

Frequencies for Explanatory Variables

Variable	Stochastic Image (*n* = 38)	No Stochastic Image (*n* = 107)	$\chi^2(1)$	*p*
Group				
1 = SR	28	43	12.591	<.001
0 = CR	10	64		
CalcI				
1 = AP	22	66	4.374	.036
0 = College	16	41		
Prior statistics				
1 = Yes	15	45	.077	.781
0 = No	23	62		
Lecture				
1 = Lecture A	29	39	17.896	<.001
0 = Lecture B	9	68		
AllLabs				
1 = Yes	35	72	2.783	.095
0 = No	7	31		

Intercorrelations between the variables were also investigated and the results are presented in Table 4.9. Analysis of relationships between the variables indicated a significant correlation between stochastic image and lab group. There was also a significant correlation between stochastic image and source of first-semester, college calculus, and between stochastic image and lecture class enrollment. Intercorrelations also revealed a significant relationship between lecture class enrollment and lab completion. There was not a significant relationship between having a stochastic image and completing all of the lab assignments. Furthermore, having earned credits for a prior college-level statistics class was not related to having a stochastic image.

Table 4.9

Intercorrelations for Stochastic Image and Predictor Variables

Variable	1	2	3	4	5	6
1. StochImage	—					
2. Group	$.295^{***}$	—				
3. CalcI	$.174^{*}$	.143	—			
4. PriorStat	.783	−.067	.055	—		
5. Lecture	$.351^{***}$	.102	−.015	−.088	—	
6. AllLabs	.139	.139	.100	.012	$.204^{*}$	—

Note. Stochastic Image coded as 1 = Stochastic or Situational category, 0 = Nonstochastic or No Image category.
$^{*}p < .05$. $^{***}p < .001$.

A logistic regression model was fitted to the data to predict StochImage and to investigate the relationship between lab group (treatment/control intervention) and the likelihood a student presented evidence of a stochastic conception. The explanatory variables used in the model identified lab group, first-semester college calculus, prior statistics course, lecture class enrollment, and lab completion. The resulting model was:

$$\hat{y}_{StochImage} = \hat{\beta}_0 + \hat{\beta}_1 Group + \hat{\beta}_2 CalcI + \hat{\beta}_3 PriorStat + \hat{\beta}_4 Lecture + \hat{\beta}_5 AllLabs$$

The logistic regression model using the binary outcome variable StochImage is presented in Table 4.10.

Table 4.10

Summary of Logistic Regression Analysis Predicting Stochastic Image

Variable	*B*	*SE*	*OR*	95% CI	Wald Statistic	*p*
Group	1.341	.452	3.798	[1.565, 9.217]	8.704	.003
CalcI	.788	.437	2.199	[0.943, 5.179]	3.252	.071
PriorStat	.105	.442	1.111	[0.467, 2.640]	0.056	.812
Lecture	1.773	.463	5.890	[2.377, 14.595]	14.668	<.001
AllLabs	.354	.534	1.424	[0.501, 4.053]	0.439	.507
Constant	−3.510	.703	0.030		24.913	<.001

The overall logistic regression model was significantly more effective than the null model in predicting stochastic image, $\chi^2(5, n = 145) = 33.88$, $p < .001$. The Hosmer-Lemeshow goodness-of-fit test yielded $\chi^2(8)$ of 9.17 and was not statistically significant ($p = .328$) suggesting that the model fit the data well.

According to the model, the log of the odds of both lab group assignment (treatment/control) and lecture class enrollment were significant predictors of stochastic image. The odds of a student in the stochastic reasoning lab group presenting a stochastic image were 3.80 times greater than the odds for a student in the calculus review lab group ($p = .003$). The odds of a student enrolled in Lecture A presenting a stochastic image were 5.89 times greater than the odds for a student in Lecture B ($p < .001$). The variables used as controls for prior achievement and mathematical and statistical background were not significantly related to stochastic image.

Lab completion was not a significant predictor of stochastic image for this model (odds ratio = 1.42, $p = .507$). However, the variable AllLabs represented lab completion for either group, i.e. completion of the stochastic reasoning labs or the calculus review labs. An interaction term was introduced to investigate whether lab completion by lab group assignment impacted the model. The resulting model was:

$$\hat{y}_{\text{StochImage}} = \hat{\beta}_0 + \hat{\beta}_1\text{Group} + \hat{\beta}_2\text{CalcI} + \hat{\beta}_3\text{PriorStat} + \hat{\beta}_4\text{Lecture} + \hat{\beta}_5\text{AllLabs} + \hat{\beta}_6\text{AllLabsbyGroup}$$

The overall logistic regression model with the interaction all labs completed by group was significantly more effective than the null model in predicting stochastic image, $\chi^2(6, n = 145) = 39.01$, $p < .001$. The Hosmer-Lemeshow goodness-of-fit test also suggested the model fit the data well as the statistic was not statistically significant, $\chi^2(7, n = 145) = 3.41$, $p = .845$.

The logistic regression model using the same explanatory variables as the first logistic regression model with an added interaction, all labs completed by group, is presented in Table 4.11.

Table 4.11

Summary of Logistic Regression Analysis Predicting Stochastic Image with Interaction All Labs Completed by Lab Group

Variable	*B*	*SE*	*OR*	95% CI	Wald Statistic	*p*
Group	−0.523	.940	0.593	[0.094, 3.745]	0.309	.578
CalcI	.775	.449	2.170	[0.900, 5.326]	2.975	.085
PriorStat	.002	.455	1.002	[0.410, 2.446]	0.000	.997
Lecture	2.002	.490	7.401	[2.832, 19.339]	16.683	<.001
AllLabs	−1.038	.774	0.354	[0.708, 1.615]	1.799	.180
AllLabsbyGroup	2.475	1.115	11.882	[1.336, 105.670]	4.927	.026
Constant	−2.633	.710	0.072		13.736	<.001

According to this model, the log of the odds of the interaction was significantly related to stochastic image. The odds of a student who was in the stochastic reasoning group and completed all of the lab assignments having a stochastic image were 11.88 times greater than the odds for a student in the calculus review group who completed all the lab assignments, assuming the odds for the non-significant variables in the model are 1:1 (p = .026).

Table 4.12 illustrates the interaction by displaying the mean stochastic image for the SR group and the CR group by completion of all labs. The mean stochastic image for the entire sample was 0.262. Table 4.12 gives a comparison of the mean stochastic image for the SR group and the CR group by completion of all labs and proportion of lab group assignment. This investigation of the interaction between all labs completed and group identification revealed that the mean stochastic image for students in the SR group who completed all the lab assignments was higher than the mean stochastic image for students in the CR group who completed all the lab assignments. Furthermore, the mean stochastic image for students in the SR group who did not complete all the lab assignments was similar to the mean stochastic image for students in the CR group who completed all of their lab assignments, as well as the mean for students in the CR group who did not complete all of the lab assignments. Therefore, the statistically significantly higher interaction of AllLabsbyGroup in predicting stochastic image means that the increase in odds was due to completion of all SR lab assignments. This evidence indicates that completion of the stochastic reasoning lab assignments had a significant impact on students' stochastic conceptions. These findings also show that completion of the stochastic reasoning lab assignments facilitated movement along a hypothetical learning trajectory towards a stochastic understanding of probability distribution.

Table 4.12

Mean Stochastic Image by Completion of All Labs and Proportion of Treatment Group

	Lab Assignment Group	
	SR	CR
All labs completed		
Mean stochastic image	0.455	0.128
n	55	48
Not all labs completed		
Mean stochastic image	0.188	0.154
n	16	26

Note. Stochastic Image is dichotomous with 1 = Stochastic or Situational and 0 = Nonstochastic or No Image.

The overall logistic regression model which included the interaction AllLabsbyGroup also showed a significant relationship between lecture class enrollment and stochastic image. The odds of a student enrolled in Lecture A presenting a stochastic image were 7.40 times greater than the odds for a student in Lecture B ($p < .001$). The same variables used in the first reported logistic regression model were used as control variables for prior achievement as well as mathematical and statistical background. Neither prior statistics nor route for completion of first-semester, college calculus was significantly related to stochastic image.

Analysis of Student Understanding Confidence Intervals

Statistical analysis of students' understanding of confidence intervals was accomplished using multiple regression modeling techniques. The analyses resulted in a model with confidence interval score as the outcome variable. Explanatory variables for the multiple regression analyses were the same predictor variables used for logistic regression analyses modeling stochastic image. Controls and predictor variables included lab assignment group, mathematical background, prior statistics course, lecture class, and lab assignment completion.

A multiple regression model was fitted to the data to predict confidence interval score and to investigate the relationship between lab group identification (treatment/control intervention) and score on the confidence interval assessment. The analysis was designed to determine the impact of five explanatory variables: lab group, first-semester college calculus, prior statistics course, lecture class enrollment, and lab completion. The resulting model was:

$$\hat{y}_{\text{ConfidInt}} = \hat{\beta}_0 + \hat{\beta}_1\text{Group} + \hat{\beta}_2\text{CalcI} + \hat{\beta}_3\text{PriorStat} + \hat{\beta}_4\text{Lecture} + \hat{\beta}_5\text{AllLabs}$$

Intercorrelations between the variables were investigated and the results are presented in Table 4.13. Analysis of relationships between the variables indicated a significant correlation between confidence interval score and source of first-semester, college calculus credits. There was also a significant correlation between confidence interval score and prior college-level statistics. Intercorrelations also revealed a significant relationship between lecture class enrollment and lab completion. There was not a significant relationship between confidence interval scores and completing all of the lab assignments. Furthermore, the confidence interval score was not related to lab group assignment (treatment/control) or lecture class enrollment.

Table 4.13

Intercorrelations for Confidence Interval Score and Predictor Variables

Variable	1	2	3	4	5	6
1. ConfidInt	—					
2. Group	−.008	—				
3. CalcI	.347**	.143	—			
4. PriorStat	.271**	−.067	.055	—		
5. Lecture	−.165	.102	−.015	−.088	—	
6. AllLabs	.117	.139	.100	.012	.204*	—

Note. $^{*}p < .05$. $^{**}p < .01$.

The multiple regression model with confidence interval score as the outcome variable is presented in Table 5.14. The overall regression was statistically significant, $R^2 = 0.217$, $F_{5,\,118} = 6.54$, $p < .001$. Taken together, the predictor variables accounted for 21.7% of the variation in confidence interval score. Two of the independent variables had a statistically significant effect on the confidence interval score: first-semester college calculus and prior statistics. Lab group assignment, lecture class enrollment, and lab completion did not have a statistically significant effect on confidence interval score. Students who earned their first-semester calculus credits via AP Calculus scored significantly higher on the confidence interval assessment when controlling for lab group, prior statistics, lecture class enrollment, and lab completion, $B = 1.39$, $t(118) = \mathbf{3.97}$, $p < .001$; $d = .737$. The means of the group of students who earned their first-semester calculus credits via AP Calculus and the group of students who earned their first-semester calculus credits via a college course differed by approximately seven-tenths of a standard deviation, which is a medium effect size. Students who had earned credits for a prior college-level statistics course scored significantly higher on the confidence interval assessment when controlling for lab group, first-semester calculus course, lecture class enrollment, and lab completion, *B* = 0.95, *t*(118) =2.68, *p* = .008; *d* = .571. The means of the group of students who had earned credits for a prior college-level statistics course and the group of students who had not earned credits for a prior college-level statistics course differed by approximately six-tenths of a standard deviation, which is a medium effect size. An investigation into a possible interaction between first-semester college calculus and prior statistics revealed no significant interaction between those predictors, $\beta^* = -0.093,\ t(118) = -0.651,\ p = .517$. The results of this investigation suggest the effect of first-semester, college calculus was not moderated by prior statistics, nor was the effect of prior statistics was not moderated by first-semester, college calculus.

Table 4.14

Summary of Multiple Regression Analysis Predicting Confidence Interval Score

Variable	*B*	*SE B*	*β*	*t*	*p*
Group	−0.186	.349	−.044	−0.532	.596
CalcI	1.386	.349	.329	3.974	<.001
PriorStat	.945	.353	.222	2.679	.008
Lecture	−.670	.358	−.159	−1.873	.064
AllLabs	.664	.412	.136	1.612	.110
Constant	6.831	.429		15.935	<.001

Analysis of Student Understanding of Probability and Statistics as Evidenced on Final Examination

Statistical analysis of students' understanding evidenced on the final course examination was accomplished using multiple regression modeling techniques. The analyses resulted in a model with final exam score as the outcome variable. Explanatory variables for multiple regression analyses were the same explanatory variables used for the logistic regression analysis modeling stochastic image and the multiple regression analysis modeling confidence interval score. Explanatory variables included lab assignment group, mathematical background, prior statistics course, lecture class, and lab assignment completion.

A multiple regression model was fitted to the data to predict final examination score and investigate the relationship between lab group (treatment/control intervention) and score on the final course examination. The analysis was designed to determine the impact of five explanatory variables: lab group, first-semester college calculus, prior statistics course, lecture class enrollment, and lab completion. The resulting model was:

$$\hat{y}_{\text{FinalExam}} = \hat{\beta}_0 + \hat{\beta}_1\text{Group} + \hat{\beta}_2\text{CalcI} + \hat{\beta}_3\text{PriorStat} + \hat{\beta}_4\text{Lecture} + \hat{\beta}_5\text{AllLabs}$$

Intercorrelations between the variables were investigated and results are presented in Table 4.15. Analysis of relationships between the variables indicated a significant correlation between final examination score and source of first-semester college calculus credits. There was also a significant correlation between final examination score and lab completion. Intercorrelations also revealed a significant relationship between lecture class enrollment and lab completion. There was not a significant relationship between final examination scores and lecture class enrollment. Furthermore, final examination score was not related to lab group assignment (treatment/control) or having earned credits for a prior college-level statistics course.

Table 4.15

Intercorrelations for Final Exam Score and Predictor Variables

Variable	1	2	3	4	5	6
1. FinalExam	—					
2. Group	−.047	—				
3. CalcI	.313**	.143	—			
4. PriorStat	.148	−.067	.055	—		
5. Lecture	.045	.102	−.015	−.088	—	
6. AllLabs	.239**	.139	.100	.012	.204*	—

Note. $^{*}p < .05$. $^{**}p < .01$.

The multiple regression model with final exam score as the outcome variable is presented in Table 4.16. The overall regression model was statistically significant, R^2 = 0.414, $F_{5,\,139}$ = 5.74, p < .001. Taken together, the predictor variables accounted for 41.4% of the variation in final examination score. Two of the independent variables had a statistically significant effect on final examination: first-semester college calculus and lab completion. Lab group assignment, prior statistics course, and lecture class enrollment did not have a statistically significant effect on final examination score. Students who earned their first-semester calculus credits via AP Calculus scored significantly higher on the final course examination when controlling for lab group, prior statistics, lecture class enrollment, and lab completion, B = 0.55, t(139) = 3.83 , p < .001; d = .653. The means of the group of students who earned their first-semester calculus credits via AP Calculus and the group of students who earned their first-semester calculus credits via a college course differed by approximately two-thirds of a standard deviation, which is a medium effect size. Students who completed all the lab assignments scored significantly higher on the final course examination when controlling for lab group, first-semester calculus course, prior statistics, and lecture class enrollment $B = 0.44$, $t(139) = 2.73$, $p = .007$; $d = .539$. The means of the group of students who completed all the lab assignments and the group of students who did not complete all the lab assignments differed by approximately one-half of a standard deviation, which is a medium effect size. An investigation into a possible interaction between lab group (treatment/control) and lab completion showed there was not a significant interaction between those predictors, $\beta^{*} = -0.004,\ t(138) = -0.03,\ p = .979$. The investigation revealed that the lab completion effect was not moderated by lab group assignment.

Table 4.16

Summary of Multiple Regression Analysis Predicting Final Exam Score

Variable	*B*	*SE B*	*β*	*t*	*p*
Lab Group	−.208	.144	−.115	−1.447	.150
Calculus I Course	.551	.144	.301	3.832	<.001
Prior Statistics	.229	.143	.124	1.596	.113
Lecture Class	.051	.145	.028	0.353	.725
All Labs Completed	.436	.160	.218	2.725	.007
Constant	−.426	.166		−2.569	.011

Findings by Research Question

The main research question was addressed through investigation of three research sub-questions. The findings are presented for each research sub-question, followed by a synthesis of the findings which addresses the main research question. The research questions were analyzed using two logistic regression models and two multiple regression models. The two logistic regression models evaluated the impact of the instructional intervention on the same binary outcome variable StochImage, which represented whether or not a student indicated a stochastic conception on a conceptual assessment. One of the logistic regression models included an interaction term in addition to the predictors used on the other logistic regression model. In order to investigate the impact of the instructional intervention on students' understanding of confidence intervals, one of the multiple regression models used the student's score on an assessment measuring understanding of confidence intervals as the outcome variable. The other regression model investigated the impact of the instructional intervention using the final course examination score as the outcome variable.

Research Sub-question 2

Does instruction designed to support development of stochastic understanding of probability distribution impact students' stochastic conceptions of a probability situation as evidenced on a conceptual assessment? The results indicate that implementation of the instructional intervention had a significant impact on students' stochastic understanding. Students in the stochastic reasoning lab group were significantly more likely to indicate evidence of a stochastic conception for one or both problems on the conceptual quiz than students in the calculus review lab group (p = .003). The lab assignments were specifically designed to support development of stochastic reasoning along a research-based, hypothetical learning trajectory. Thus, students needed to complete all of the stochastic reasoning lab assignments in order to move along the hypothetical learning trajectory toward development of stochastic reasoning. The interaction between lab assignment group and completions of all lab assignments was significant (p = .026). An investigation of this interaction revealed that students in the SR lab group who completed all of the lab assignments had a higher mean stochastic image than students in the CR lab group who completed all of the lab assignments. Furthermore, the mean stochastic image for students in the SR group who did not complete all of the lab assignments was similar to the mean for students in the CR group who completed all of the lab assignments, as well as the mean for students in the CR group who did not complete all of the lab assignments. This evidence strongly suggests that the stochastic reasoning labs impacted students' thinking about probabilistic notions related to probability distribution.

Lecture class enrollment was also significantly related to students' stochastic understanding. Students in Lecture A were significantly more likely to exhibit evidence of a stochastic conception ($p < .001$). Although this study did not specifically investigate the instruction that took place during the lecture and discussion classes, the statistical evidence suggests that this instruction mattered. It appears that the instruction students in Lecture A received impacted their thinking in ways that promoted development of stochastic understanding.

Research Sub-question 3

Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding of confidence intervals as measured by the ARTIST assessment? There was no evidence that the instructional intervention impacted students' understanding of confidence intervals, as measured by scores on the assessment of students' understanding of confidence intervals. Confidence interval scores did not differ for students in the treatment (SR) and control (CR) groups. However, there was a significant relationship between students' mathematical background and their score on the confidence interval assessment. Students who earned their first-semester, college calculus credits via AP Calculus scored significantly higher on the confidence interval assessment ($p < .001$). Also, students who had earned credits for a prior statistics course scored significantly higher on the confidence interval assessment ($p = .008$). An investigation into a possible interaction effect revealed the effect of first-semester college calculus was not moderated by prior statistics, nor was the effect of prior statistics was not moderated by first-semester college calculus. These results imply that students with stronger backgrounds in mathematics or statistics had higher achievement on the confidence interval assessment.

Research Sub-question 4

Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding as evidenced on final course examinations administered in an introductory calculus-based probability and statistics course? There is no evidence that the instructional intervention impacted students' scores on the final examination. Final examination scores for students in the SR group did not differ from final examination scores for students in the CR group. However, there was a significant relationship between students' mathematical background and final examination score. The students who earned credits for their first semester of college calculus via AP Calculus scored significantly higher on the final course examination than those who took their first semester of calculus in college ($p < .001$). As a proxy for student achievement and mathematical background, the significant relationship between first-semester calculus and final course examination implies that higher achieving students with stronger backgrounds in mathematics performed better on the final course examination. There was also a significant relationship between lab completion and final examination score. Students who completed all of the lab assignments scored significantly higher on the final examination (p = .007). There was not a significant interaction between completion of all labs and lab assignment group. Hence, it may be that lab completion could be considered a proxy for student diligence. The significant relationship between final examination score and lab completion could imply that students who diligently completed the required work for the course did better on the final course examination.

Main Research Question

What is the impact of an instructional intervention designed to support development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory, calculus-based probability and statistics course? The stochastic reasoning labs were designed to support development of stochastic understanding of probability, and these labs had a significant impact on students' conception of probability distribution as measured by the conceptual assessment. Students who completed all of the stochastic reasoning labs were more likely to present evidence of stochastic reasoning. The instructional intervention did not impact students' understanding of confidence intervals, as measured by the confidence interval assessment. Furthermore, the instructional intervention did not impact students' understanding of probability and statistics, as measured by the final course examination. Students needed to use calculus to successfully complete some final examination problems, but the calculus review labs did not significantly impact students' scores on the final course examination. Although there is no evidence that the instructional intervention impacted students' scores on the confidence interval assessment or the final course examination, there is evidence that the instructional intervention impacted students' thinking in ways that promoted a deeper understanding of probability distribution and statistical inference.

CHAPTER 5: QUALITATIVE RESULTS AND ANALSYIS

Chapter 5 presents results and analysis of the qualitative phase of the study. Twelve students participated in interviews at the end of the course. Six students of these students were selected from the treatment group and six from the control group. The purpose of the interviews was to probe students' reasoning about probability in two problem contexts represented on the conceptual quiz, as well as probe their generalized reasoning about probability. Analysis of the interview data provided insight into students' thinking and reasoning about their written answers on the conceptual assessment, which subsequently informed the development of a scoring rubric that was used to score all student participant responses to the conceptual quiz problems. Probing students' thinking about their answers to the conceptual assessment provided validity to the conceptual quiz, informed interpretation of the quantitative data, and strengthened the overall interpretation of findings.

The twelve interview participants held a variety of conceptions of probability and their responses provided insight into students' conceptions of probability. The analysis presented in Chapter 5 addresses research sub-question 1 and three components of this research sub-question related to students' conceptions of probability distribution:

- Research sub-question 1: What is the nature of students' reasoning when confronted with a probability situation?
    a) How do students characterize a probability situation in terms of an image of a repeatable process?
    b) How do students characterize a probability situation in terms of specification of conditions of a repeatable process?

c) How do students characterize a probability situation in terms of an image of a distribution of outcomes?

This chapter is organized into four sections. The first section gives a description of images comprising a stochastic conception in each problem context and examples of students' conceptions gleaned from interview data. The second section describes a stochastic conception rubric and evidence of stochastic reasoning across categories of the analytic framework. The third section offers an analysis of interview participants' stochastic conceptions and characterization of students' thinking. The final section summarizes the analysis which framed four ways of reasoning about probabilistic situations in the problem contexts: stochastic, situational, nonstochastic, and no image.

Description of Images Comprising a Stochastic Conception

The purpose of the interviews was to investigate students' stochastic conceptions of probability and probability distribution. The two problems on the conceptual assessment involved different contexts and each problem represented a situation that could be approached stochastically or nonstochastically (Liu & Thompson, 2007). The analytic framework developed for analysis of the interview data is comprised of three categories which were described by Liu and Thompson (2007) as constructs that lead to a stochastic conception of probability: image of a repeatable process, specification of conditions of the process, and image of a distribution of outcomes. Interview data provided insight into student thinking within each category of the analytic framework for each of the two problem contexts in the conceptual quiz. The following are descriptions, which were expressed by students during the interviews, of images held for each category of the analytic framework.

Image of a Repeatable Process

This section describes evidence of students' conceptions of an image of a repeatable process for each of the two problem contexts. First, an image of a repeatable process is depicted for the hospital problem context. Following this depiction are descriptions of an image of a repeatable process garnered from students who presented an overall stochastic conception, situational conception, nonstochastic conception, or no image for the hospital problem context. Next, an image of a repeatable process is depicted for the theater problem context. Following this depiction are descriptions of an image of a repeatable process garnered from students who presented an overall stochastic conception, situational conception, nonstochastic conception, or no image for the theater problem.

Hospital problem image of a repeatable process. In the context of the hospital problem, the gender outcome of a birth is a random phenomenon, and the number of births per day can be thought of as the result of a random process or a stochastic process. Each day represents a repetition of this process. What follows are exemplars garnered from the interviews that illustrate understanding held by students who were characterized as having either a stochastic conception, a situational stochastic conception, a nonstochastic conception, or no image of a repeatable process for the hospital problem.

Hospital problem image of repeatable process: Stochastic conception. Five students indicated a stochastic conception of the hospital problem, and two of these students indicated a stochastic conception of both the theater and the hospital problems. Students who held a stochastic conception of the hospital problem indicated thinking about numbers of days and also indicated thinking about possible outcomes for each day, several days, or a particular day. For example, one of these students interpreted the problem as "asking me how many times, how many days will there be, which ones will have more days. I was quickly trying to see, okay, what do you expect each day the boys will be." (Student 113). Another student used phrases such as "at early stages" and "as time goes on" to describe thinking about the process. While these students talked about births each day, they also talked about variability in the numbers of births. Thus, they did not interpret the number of births per day as a static birth rate.

The students who held a stochastic conception also talked about sample size and equated sample size to the number of births in each hospital. Student 214 said "When you're talking about sample sizes, you're talking about the number of births." The notion of repetition was evidenced by talk about repeated sampling along with the importance of attending to sample sizes. Student 329 mentioned that he "just remembered, like, wondering about the sample sizes and how the smaller ones vary a lot more n the larger ones." He went on to say "if you keep doing the same sample over and over and over again and the same, like, experiment, that you would get some type of consistency in your numbers." Other students who held a stochastic conception of the hospital problem also thought about sampling of the total births in the town and related that to repetitions of experiments.

In addition to thinking about outcomes and numbers of days in terms of a repeated process, four of the students who held a stochastic conception of the hospital problem also mentioned thinking about repeated trials. These students connected their thinking about repetition to a model. Three students connected thinking about underlying processes in the hospital problem to the model of flipping a fair coin, and the other student related her thinking about the hospital problem to running repeated experiments in chemistry. Students 113 and 329 were in the stochastic lab group and connected their thinking about "trials" in the hospital problem to simulations that were run in the stochastic lab assignments. For example, Student 113 stated, "You have a certain number of trials, a fixed probability, so I was thinking similar to flipping a coin. The way I thought about it was: If I flipped a coin 45 times and another experiment I flip a coin 15 times, which one will more likely to get over 60% heads? So I was thinking very quickly, you'll see that it would be easier, through the things we were doing all semester, running these experiments through *Fathom*, that's what I was doing." Student 113 recalled what he observed in the lab assignment simulations and talked about a "vast number of trials" and what might happen "if you have a lot and a lot of trials" in relation to the hospital problem. Student 329 explained he thought about the *Fathom* lab and flipping a coin: "...if you keep doing the sample over and over and over again and the same, like, experiment then you would get some consistency in your numbers."

Although not in the stochastic lab group, Student 214 also connected thinking about the hospital problem to repeatedly flipping a fair coin. He explained:

> I ended up coming up with my answer not seemingly through any math here but through pre-conceived knowledge, I guess that I had, that in larger samples sizes, things tend to fall closer to average than you would expect. Like, if you flip a coin 100 times you're more likely to get closer to an even 50/50 than if you flipped it 4 times." He also mentioned, "Here I was actually visualizing in my head this situation happening.

Hospital problem image of repeatable process: Situational stochastic conception. One student held a stochastic conception of the theater problem, but did not hold a stochastic conception of the hospital problem. Liu and Thompson (2007) characterize individuals who hold stochastic conceptions in particular situations and not others as having a situational conception. When interpreting the hospital problem, the student with a situational conception demonstrated an image of a repeatable process by talking about the number of babies born in each hospital over a period of time. Student 313 explained:

> When I recognized that hospital one has 45 babies a day as opposed to hospital two, which has 15 babies a day, I know that hospital one is going to, (pause) if I'm given a certain amount of time, it's more likely that more babies will be born in that amount of time than at hospital two.

Student 313 also mentioned repetition when explaining what he meant when he wrote the words "probability of error" on his answer to the quiz problem: "I said it was more likely the smaller hospital will record more of these days because if, I was considering if one more boy was born than girls. …Because of error we're saying that." Furthermore, Student 313 talked about sample size and equated sample size to the number of births in each hospital:

> I'm just using in regards to this problem, 'cause if they're seeing 45 babies a day, 15 babies a day. With 45 babies a day I recognize that there's going to be a greater sample size that has more, (pause) now I'm playing with variability and error in my head (pause). There's the probability (pause) and the statistics concepts. … I'm thinking that the greater sample size at 45 babies a day.

Hospital problem image of repeatable process: Nonstochastic conception. Three students held an image of a repeatable process but did not have a stochastic conception of the hospital problem. These students indicated thinking about numbers of days as well as thinking about which hospital would have more days with a greater number of boys born. For example, Student 711 mentioned the importance of thinking about sample size in conjunction with repetition. She also equated the number of births in each hospital to sample size and stated:

> It's more likely the smaller hospital will record more such days because in terms of variance, an outlier would affect the variance of like, sample size. That's important because sample size is, just as I was saying earlier, the sample you're taking out (pause) but a typical sample is not random at all, it's just like each day … . Yes, sample size is important because one has 15 (pause) as I was explaining before.

The other students who held an image of a repeatable process for the hospital problem did not mention sampling but talked about percentages. One student indicated that thinking about probability in terms of percentages was important. This student stated, "You would need sample sizes, and you would need how many days you would be doing the study over. You'd be finding the mean… . Then the expected value would give you a more precise definition of the probability that you've already given." (Student 715). The other student thought about how the variation in the number of boys born versus girls would have a greater effect on the percent of boys born in the smaller hospital. Student 1016 explained:

> There's more births here and there's less births in the smaller hospital. So the smaller hospital will have more days from which there's a greater number of boys born, because of a greater percentage, like a ratio. Because just one birth can alter the ratio, and that's why I said the smaller hospital. … When you look at the boys, and you change it to 9 boys and 6 girls, just one difference, then you get 60%.

Both of these students appeared to be thinking in terms of repeating the process of births and observing the numbers of births over a numbers of days. In addition, Student 715 seemed to be thinking about expectation and probability in terms of percentages in relation to expectation, while Student 1016 focused on the how changes in the number of boys born in each hospital would affect the percent of boys born in that hospital.

Hospital problem no image of repeatable process. The three students who did not indicate an image of a repeatable process focused on the ratio of births per day and mentioned that one additional boy born in the smaller hospital would have a greater effect on the proportion of boys born in that hospital as compared to the larger hospital. In contrast to students who indicated an image of a repeatable process, these students focused on a static birth rate and probability as a ratio. For example, Student 518 explained:

> So I didn't even try to assume anything about the chance of getting a boy or girl. … So I just looked at the problem that I was stuck with a birth rate per hospital. And, accounting for days that are 60% or more and given that to those two pieces of information, that came down to [I] should simply be determining how many births per day would need to take place instead of trying to worry about percentages.

Student 811 also thought about probability as a ration and said, "I still visualize probabilities as fractions. Because it's 60%, it said it was 60% number of boys born, so it was on my mind." Student 811 also explained, "I look at probabilities on the basis of, I look at is as a fraction of something, that it's a set ratio. …That ratio is going to hold for both cases, small and large because that was the population." In addition to holding a perception of probability as a set ratio, these students appeared to wonder why they were asked about their thinking of sample size in relation to this problem. They thought that sample size was neither relevant nor important to thinking about the hospital problem. One student admitted, "It's not very clear how, what the sample size would be representing."

Theater problem image of a repeatable process. In the context of the theater problem, the occurrence of someone in a town attending a movie at the theater where Anthony works can be considered a random phenomenon and then the result of the random process is attending a movie. The problem asks whether or not Anthony knows someone who attends the theater on days that he works, so each day that Anthony works at the theater represents a repetition of this process. What follows are exemplars garnered from the interviews that illustrate understanding held by students who were characterized as having either a stochastic conception, a situational stochastic conception, a nonstochastic conception, or no image of a repeatable process for the theater problem.

Theater problem image of repeatable process: Stochastic conception. Four students held a stochastic conception of the theater problem, and two of these students also evidenced a stochastic conception of the quiz problem addressing births at the hospitals. Students with a stochastic conception of the theater problem indicated that they were thinking about more than one incidence of people coming to the movies and that the number of people Anthony knew would vary. Thinking about repetition was indicated by envisioning repeated nights at the movie theater. For example, Student 113 talked about repetition in terms of the movie happening every night when he said, "Let's say the movie happened every night and 250 new people every single [night], you know there'd be, quite possible that on one or two nights you'd see no one." Student 214 said that he was thinking about "the 250 people he would encounter in a movie theater in a night," and wrote that the number of people Anthony would see on any given night is "not known."

Two of the students who held stochastic conceptions of the theater problem also mentioned thinking about sample size and repeated sampling as indicative of a repeatable process. Student 113 explicitly stated, "Like we learned in class, you take a sample of 250 from that 30,000 population … so I expect one percent of 250 is 2.5. So I said that is what I would expect him to see if the sample is representative of the population." He went on to explain, "You know, you were having this movie over and over and over again. And you were, you know, sampling with replacement with the population." He also talked about taking many samples and thought about expectation in relation to "if he took a sample of 250 over and over and over again." Student 313 also mentioned sampling but appeared to confuse sample size with the size of the population when he stated, "I took in the sample size just cause the amount we're given is concrete, the sample size of 30,000 people and not so much anything else." However, later in the interview Student 313 mentioned that "in the theater of 250 people I can have from 0 to 250 people Anthony knows." He also said that "the results differ depending on the situation and … we can never predict exactly how many people are going to come." These statements indicate thinking about a sampling process that is repeating.

Student 313 connected thinking about repetition in the theater problem to a model of repeatedly flipping coins he experienced in the labs. When asked how work on the labs influenced his thinking about the two problems, he replied:

> I know for problem number two (the theater problem) definitely it was, I can think of the first or second lab, where we were using coin flips … . There was one part where we had to calculate the expected, how many I guess, I think, if I can remember back, was like how many coin flips you expect to be either heads or tails. And that's how I approached this problem.

Student 313 also explained, "With randomness and probability that there's a wide range of samples and that they can vary so." Perhaps Student 113 was also thinking about a model when he connected the notion of expectation to repeating an experiment: "Because you'd expect him to see 2 ½ every time if you completely conducted this experiment over and over."

Theater problem image of repeatable process: Situational stochastic conception. Two students did not hold a stochastic conception of the theater problem but held a stochastic conception of the hospital problem and held an image of a repeatable process. These students are classified as holding a situational stochastic conception. When describing their thinking about the theater problem, they talked about repetition in terms of nights at the movies and going to the movies. These students also talked about sampling, however, none indicated thinking about a model in relation to the theater problem. When Student 325 talked about her thinking related to sampling and the process of sampling she said:

> I'm thinking in terms of size, um, that people aren't just going to the movies because they know him, so I kinda feel like it just kinda like random. … Some people are more likely to go to the movies and some aren't as likely to go to the movies, so that's why I see the sample, large samples.

Theater problem image of repeatable process: Nonstochastic conception. Two students held an image of a repeatable process but did not hold stochastic conceptions of either problem. These students talked about repetition in terms of days that Anthony worked at the theater. Student 711 talked about what might happen on different days: "Like if one day he sees 2 people he knows, like that may be more random as opposed to if, um, if like every day, if later on he sees 2 people he knows." Although this student did not specifically mention sample size, she indicated thinking that was related to taking a sample when she said, "And on any given day you're only getting 250 exposed. But like, and this is what makes it hard, is because it's 30,000 people. That's so many, and, like, 250 (pause) this would be for one day." Student 715 talked about sampling and conceived of a process that involved taking many samples:

> You wouldn't go there one day and say how many people do you know this one night and base your entire experiment on one sample. The more, I guess, credibility a sample has is based on how many samples are taken and how large the sample size becomes and the fact there is still a trend there.

Student 715 also mentioned repeated sampling and said, "You wouldn't sample one night or five nights, you would sample maybe 20 or even more than that." She did not mention sample size, but envisioned sampling a number of nights.

Theater problem no image of repeatable process. Four students did not indicate an image of a repeatable process. Two of these students claimed that the problem was challenging. When asked about the theater problem, Student 428 immediately said, "Oh this is hard, I had no idea what to do with this. … If I had more time I think I would have thought about other stuff." She wanted to take a mathematical approach to solving the problem and struggled with how to approach the problem in the time given:

> I didn't have enough time to think of it and (pause) more mathematical ways so I was thinking more about, like, other things that could happen like … there's nothing special about the movie, and 100 people are coming that know Anthony, people still definitely not a random sample ... some people never go to the movies, some people go multiple times a week (Student 428).

Student 1016 said that he "completely guessed on this," but settled on an approach that involved thinking of probability as a ratio. He said, "I took this as a population, the town, and the people he knew as a sample from that. And I took the ratio of the people he knew. He knew only 1% of the population."

In fact, all four of these students approached the problem by thinking of probability as a ratio and used the numbers to find answer. For example, Student 811 explained:

> I made the amount of people he knew a proportion, because he knew, or probably because he knew 300 people out of 30,000, which would be 1% of the population. And then the expected value of that would just be the number of theater goers times the probability. So, he would expect to know 2½ people out of every two hundred.

Student 811 summed up his thinking by saying that he focused on numerical calculations, "When I went through this problem, basically took all the numbers in context, pulled that value out and then worked on it. Completely numerical; then plugged that back into the problem." Student 811 also indicated that his primary goal was to find an answer: "So, I don't consider this problem in relevance to any of the thought processes I did until I was coming to a conclusion." Students who did not indicate an image of a repeatable process were thinking about the using the numbers to find an answer using mathematics.

Specifications of Conditions of the Process

This section describes evidence of students' conceptions involving specification of the process for each of the two problem contexts. First, specification of conditions of the process is depicted for the hospital problem context. Following this depiction are descriptions garnered from students who specified conditions of the process for the hospital problem and descriptions garnered from students who did not specify conditions of the process for the hospital problem. Next, specification of conditions of the process is depicted for the theater problem context. Following this depiction are descriptions garnered from students who specified conditions of the process for the theater problem and descriptions garnered from students who did not specify conditions of the process for the theater hospital problem.

Hospital problem specification of conditions of the process. In the context of the hospital problem, each repetition of the process results in the birth of a boy or a girl. Hence the outcome of the process is either a boy or girl. One condition of the process is stated in the stem: 50% of all babies born in the town are boys. It should be assumed that the births are independent. Thus, the outcome of a baby boy is equally likely for each repetition of the process. This problem required an interpretation of the number of boys born each day, and understanding that there will be variability in the number of boys born each day at each hospital. Each hospital can be conceived of as a sample of births; each day the number of boys born at the hospital will vary, as will the proportion of boys born at that hospital. The repetition of the process will produce a collection of births at each hospital. What follows are exemplars garnered from the interviews that illustrate understanding held by students who were characterized as having a stochastic conception and as specifying conditions of the process, as well as the understanding of students who were characterized as not specifying conditions for the hospital problem.

Hospital problem specification of conditions of the process: Stochastic conception. Five students held a stochastic conception of the hospital problem and all perceived an outcome of the repeatable process as the birth of either a boy or a girl. They also recognized that there was a 50% chance each birth would result in a either a boy or girl and connected the probability with an assumption of independence. Student 329 said that he assumed "the typical day would be about 50% for every single day for boys and girls in both hospitals." When asked to tell about any assumptions or conditions he was thinking about when solving the problem Student 817 replied, "I guess I thought it was normal. …As in true or false, like also boy or girl. I guess I thought having 50% each."

Students who held a stochastic conception of the hospital problem connected the probability of each outcome with an assumption that each birth (outcome) was independent. For example, Student 329 said, "Well that's just the probability of having boys is independent of having a girl, so either way they're both 50% all the time." He explained independence as meaning: "… two variables don't have any impact on each other." Student 214 said, "I assume that you would have 50% chance of it either being boys or girls, which isn't quite true in nature but we're given that in the problem so I stuck with that." When asked about his assumptions, Student 113 stated, "I assumed independence. I made the assumption that each trial was independent, and I also made the assumption that boys and girls were equally likely." He also explained that "having a boy doesn't affect or does not influence the chances of having a boy next time."

Students with stochastic conceptions perceived each hospital as a sample and realized the importance of thinking about sample size. For example, Student 113 explained:

> So, it seems like each hospital is a sampling of the total births in the town. … I'm thinking sample size is important because when you're conducting an experiment more and more you're going to get closer to what would be expected. … I was drawing more really on sample size. I felt like the relative sample size is between the two hospitals.

Student 329 related the smaller hospital to a smaller sample and the larger hospital to a larger sample, and although she struggled to articulate her thinking about a sampling process, she wondered about the sample sizes and "how the smaller ones vary a lot more than the larger ones."

These students recognized that there would be variability in the number of boys born at each hospital and that sample size was related to variability. Student 817 connected variability with sample size and explained, "I was thinking about (pause) the smaller hospital (pause) has a more lesser sample size, so it would be more extreme. …It only has 15. Gonna be more of a bias." Student 329 said, "Since there is only 15 births each day in the smaller hospital. I realized that one would have more variance just because it's so much of a smaller sample, and 45 would be closer to the same thing day to day. Where, 15 would vary a lot more." Student 214 explained thinking about variability in relation to sample size in this manner:

> Here I was actually visualizing in my head this situation kind of happening. Where once I had adhered to the math, I thought about it (pause). For 60% of the children to be boys in a hospital with 15 beds, that would make a lot of sense because the sample size in small and that means that for a couple of children they ended up going one way rather than the other. But a hospital of 45 (pause), 27 to 18, I think. … At smaller numbers, 1 or 2 off results will give you a really high percentage one way or the other.

Four of the students with stochastic conceptions indicated thinking about how the sample might represent the population, and three of these students used the term "random" to describe the sampling process that they conceived. For example, when asked to describe the sampling process, Student 329 stated:

> I assume a random sample because (pause), I guess you can typically expect that would be the case in a hospital because they are not changing things. There's always going to be a certain number of boys and girls, so I assumed that's random. … On average they'll tend to represent the population.

When asked to explain his thinking about what it meant to take a random sample, Student 329 related this to the notion of independence: "If one person has a boy, that doesn't make the next person more likely to have a girl. In both situations it's equally likely, no matter what the previous outcome." Student 329 also mentioned, "To take a random sample, I'd feel like that's representative of the population. … You'd all make sure you had different types of ethnicities and races and genders and, like, just a random sample of the whole population." Student 329 added, "You'd want to make sure to count for everyone." Student 214 said, "I think you assume random sampling. I think you have to assume it's fair and holds true to assumptions that you're given." Hence these students understood a random sample to be "fair" and "representative" of the population.

Four of the five students who held a stochastic conception of the hospital problem also mentioned thinking about an underlying model. When explaining their thinking about the hospital problem, Students 113, 214, and 329 connected thinking about conditions of the process to the model of flipping a fair coin. Student 329 recalled work he had done the in the stochastic lab assignments:

> It's like if you flip a coin two times, you could see heads both times and that wouldn't mean that a 100% of the time it would be heads. Whereas if you do it, like in the lab where we did it and we increased the sample size each time, it got closer to 50%, which is what it actually was. So I figured that the smaller one was like when we only flipped it one or two times, whereas the, [larger one], 45 was closer to 100.

Student 113 also indicated thinking about a model of flipping a coin:

> If you were flipping a fair coin, if you flip it ten times, you'd say, 'Oh well, most often I have 5 heads.' But you go ahead and flip it just 10 times, and all of a sudden get 7 heads. All of a sudden, 70% heads and you say, 'Whoa, wait a second. That's not the probability of that happening.' And, that's why with the smaller sample space it's you have more. It's more likely to happen.

In addition to thinking about repetition in the hospital problem as being modeled by a flipping a fair coin, Student 113 conceived of independence as a condition of process and said, "Just like flipping a fair coin, having a boy in the delivery previously does not affect, or does not influence, the chances of having a boy the next time." Student 214 also thought about the model of flipping a fair coin where a head represents a boy and a tail represents a girl. Student 214 explained:

> In larger sample sizes, things tend to fall closer to average than you would expect. Like if you flip a coin 100 times. You're more likely to get closer to an even 50-50 than if you flipped a coin 4 times. [In that case] you'd more likely to have 3-1 or 4-0.

Student 325 envisioned the model of fulcrum in physics with a distribution of weights. She said, "Because there were less births each day, then you're going to get, like, it's going to be a little weird. You're going to have things further away from the like typical average, and so, because there are less things to balance out in a sense." When asked to explain what she meant, she drew a picture with dots representing weights and said, "So there's only like 50. So the average is like in the middle. And, 15 there, should be like, you know here, here. So, you know, like all over the place." Her model illustrated her thinking that there would be more variability in the number of births in the smaller hospital than in the larger hospital.

Hospital problem no specification of conditions. Seven students did not specify conditions of a stochastic process. Four of the seven students indicated an image of a repeatable process but did not anticipate that repeating the process would produce a collection of outcomes. Instead of perceiving variability in outcomes in terms of births, these students focused on birthrates for each hospital and variability in birthrates.

All four students who held an image of a repeatable process assumed that the hospitals were independent. Student 1016 stated, "The births between the two hospitals are independent and the difference between boys and girls is independent." Student 313 said, "I considered two different scenarios. I didn't consider them affecting each other. …That's the extent to which I see independence in this problem, that they don't affect the results. One hospital's results is not going to affect the results of the other hospital."

Three of the students who indicated in image of a repeatable process for the hospital problem concentrated on the different birthrates for the two hospitals. They perceived that it was important to think about sample size, but at the same time connected sample size to birthrate. Student 313 said, "But, when I'm seeing rates, I'm seeing there is a greater sample size. … So, I definitely see the sampling, the sampling has, (pause) gave us two rates." These students conceived that each hospital represented a different sample size and different birthrate or birth ratio. Student 313 said that he thought about sample size "primarily because it will affect the probability." When explaining why he thought sample size was important he said, "The way I see it there's three factors in these problems. There's the rate, there's the probability itself, and also the sample size… . Those are the three primary tools I use."

These three students also perceived variability in terms of the birthrate at each hospital. Student 313 said that he approached the problem by looking at the "rates per hospital." When comparing the birthrate in the small hospital to the large hospital he said:

> If an extra boy was born as opposed to a girl that would create a bigger difference in the ratio. Where if the girls as opposed to if one more boy was born than one more girl, in 45 children born in hospital one. … So when I was saying this, it's because the smaller hospital has a smaller sample size and they created a probability of error.

Student 711 focused on how variability in the number of births would impact the birthrate. He explained:

> It's more likely the smaller hospital will record more such days because in terms of variance, an outlier would affect the variance of the, like, an outlier would affect the average more. … If you're delivering 45 babies and, like, you deliver 5 more boys than expected … .The percentage of boys isn't going to be much higher as opposed to if you deliver 5 more boys in the smaller hospital. That's like a huge number. That will make a huge difference.

Two of these students also talked about expected value. Student 313 said, "I rely on probability as expected value. So, I just always approach a problem imagining what is the conditions were ideal, what the value would be. …I'm just using expected value to calculate how many boys and girls there will be. It just gives me a basis to start off." Student 711 explained:

> So, then it was talking about the days in which the number of boys born was 60% or more. Um, so that was in comparison to the 50% that was the expected probability boys born each day. …By expected probability, just means that it would be the mean, the μ, so the expected probability. I guess I'm thinking of, like, E of X, which would be about 7.5 for the smaller hospital and (pause) 27.5 for the larger hospital.

Other students who did not specify condition of a stochastic process, perceived probability as a percent or ratio, but did not indicate they were thinking about variability. These students focused on calculating a birthrate. For example, Student 518 said:

> So I didn't try to assume anything about the chance of getting a boy or girl, just by ignoring that piece of information. So, I looked at the problem that I was stuck with: a birth rate per hospital and accounting for days that are 60% or more. And given that two pieces of information, that came down to, should simply be, just determining how many births per day would need to take place instead of trying to worry about percentages.

One of these students held an image of a repeatable process. This student mentioned sampling but did not indicate that sampling was related to variability in the hospital problem. Student 715 said:

> So, it's the percent of sampling that you're doing. So of the sampling you're doing, there's the same chance that you will always, that you will always have that 60% on that day in that hospital. … Because you're talking about percentages, the sample size shouldn't be necessary, because the percent is, because the percent scales to whatever the percent of the population is. So, it's whether you have 45 or 15 babies, you'll still have the same percentage of that sample.

Student 715 used the term "random sampling," but to her, random sampling meant "not biased." The other students, who focused on birthrate in the absence of variability, did not indicate that thinking about sampling was relevant to the hospital problem.

None of the students who focused on birthrate in the absence of variability connected notions of independence to the hospital problem. Student 518 said, "Independence, I did not think about." Student 811 said he assumed "…that everyone had an equal chance to go to each of the hospitals." Furthermore, most of these students seemed to shy away from making assumptions about the problem situation. For example, Student 518 said, "When I first read the problem, my initial assumption was that it was a fairly simple probability. It was determined, very simple calculations." When asked to explain any assumptions she made, Student 428 said, "I really try not to make assumptions, which is why I had trouble answering the question."

Theater problem specification of conditions of the process. In the context of the theater problem, each repetition of the process results in someone going to the movie theater where Anthony works. The outcome of this process is that Anthony knows the person who attends the movie on a day he works or Anthony does not know the person. The problem states that the reader is to assume that people are not coming to the theater because they know Anthony and that there is nothing special about the type of movie. This assumption indicates that it is equally likely for any individual in the town to go to the movie theater. A stochastic conception of this problem requires consideration of the assumption of independence. Because people attending the movie are drawn from a large population, the independence assumption is met. Thus, each day that Anthony works at the theater can be conceived of as a random sample from a large population. This sampling process is repeated each day that Anthony works. The sample size is 250 based on an assumption that all theater seats are filled. The proportion of people Anthony knows on any given day is a sample proportion. Conceiving of the conditions of the stochastic process requires understanding that the number of people that Anthony knows at the theater each day will vary, thus the proportion of people that Anthony knows will vary. Repeating the process of filling the theater with people on nights that Anthony works produces a collection of outcomes, which are the number of people Anthony knows on any given night.

Theater problem specification of conditions: Stochastic conception. The four students who indicated a stochastic conception of the theater problem conceived of the number of people that Anthony knows each night as an outcome. They also understood that the number of people Anthony knows each night will vary and that repeating the process of people coming to the theater would yield a collection of outcomes. For example, Student 313 said, "I can have from 0 to 250 people that Anthony knows. … So randomness allows me to say … the results can differ depending on the situation. … We can never predict exactly how many people are going to come." Student 113 explained, "Let's say the movie happened every night, and 250 new people every single, you know, night. There'd be quite possible that on one or two nights you'd see no one." Student 113 also said:

> For him to see 2 people, that's very close to what you'd expect. You really wouldn't expect him to see 10 or 15. And, that's sort of how I was thinking about it, that you know, a lot of samples … and you assume the sample is random. …. The number of people he would see would be close to 2.5 if he took the sample over and over and over again.

Two of the students who held a stochastic conception of the hospital problem indicated that they were thinking about sampling from the population of the town. Student 214 said:

> I assumed that the 250 people that the theater would hold in a night, would be a fair estimate of the population. … So for every 100 people he knows in the city, he knows 1 of them. Then, I related this to the 250 people he would encounter in the movie theater in a night.

Student 113 also indicated that he was thinking about sampling:

> So he knows 1% of the people. Like we learned in class, you take a sample of 250 from that 30,000 population and you assume that it'll be the distribution similar to, (pause) or that it'll be normal compared to the town, the total population. So I expect one percent of 250 is 2.5. So that's what I would expect to see if the sample is representative of the population. … You know, you were having this movie over and over and over again, and you were, you know, sampling with replacement from the population.

These students connected thinking about sampling and randomness to an assumption of independence. Student 113 talked about sampling with replacement and said, "The process of taking a random sample, you'd want it to completely be the chance of you choosing one person over the other completely, you know, just random." Student 113 explained that random meant, "It's equally likely to choose anyone in the population as you're choosing that one person. If you're going to go with a random sample, you're going to want ideally … completely pulling numbers out of a hat and you'd have similar representation." Student 214 explained that he was thinking about a random sample:

> The given assumptions that the fact Anthony was there and the type of movie playing don't have any effect in the number of people attending. … The people he knows going to the movies might be random because if he really liked a ton of people who loved going to the movies, they'd more likely be represented.

Student 214 defined random as meaning "that given a population or given a probability, it isn't biased by other factors that we're given in the problem." Student 329 explained that he thought about a random sample in the theater problem context: "I think that the 250 people that go to the movie theater are, it's not the same people every time. It's a random sample of the population." Student 329 also assumed that people coming to the theater had an equal chance of selection and explained: "…Just that the people who are going to the theater are equally distributed throughout the population." Student 329 also assumed that the sample was representative of the population: "The people that go to the theater are an accurate representation of the town."

The student who did not talk about sampling said that he connected what he learned in the stochastic lab assignments to his thinking about the theater problem. He thought about the stochastic lab which investigated connections between an empirical distribution and a theoretical probability distribution model in the context of 20 people selected at random to do a blind taste test of bottled water and tap water. Student 313 explained:

> There was one where we had to calculate the expected, how many. One of the questions was, 'How many people would we [expect]?' I guess, if I can remember back, it was like, 'How many coin flips would you expect to be either heads or tails?' And, that's how I approached this problem.

Theater problem specification of conditions: Situational stochastic conception. One student held a stochastic conception of the hospital problem and specified conditions of a stochastic process for the theater problem, but did not hold a stochastic conception of the theater problem. The student is classified as holding a situational stochastic conception. Student 817 viewed outcomes as "incidences" of knowing people: "He's going to know at least one person. That's what I was getting at. … Incident of knowing people [is] one person, so it would be more than one." Student 817 thought of the process of people coming to the theater as random sampling from the population. Student 817 also said:

> I guess that a person coming in is going to be independent of each other, like picking a person out of 30,000 people and it's not going affect any other people coming along with him. I guess, like, having an equal chance of people going to the theater. … People coming in, I thought should be random.

Theater problem no specification of conditions. Seven students did not specify conditions of a stochastic process for the theater problem. Three of these students held an image of a repeatable process for the theater problem and each one talked about sampling. However, these students also indicated thinking that was contradictory to what it means to take a random sample. Student 711 connected the notion of randomness to an equal chance of selection and said, “I guess random is when you have, when you choose your sample at random so that every person in the population has an equal, and equal probability of being chosen for the sample.” She also understood that there would be variability in the number of people Anthony knew at the theater each night, but she did not perceive that repetition of the process would yield independent random samples. Student 711 explained:

> And so with these 250 people, not everybody in 30,000 people have an equal chance of being chosen because there are all those factors influencing whether or not a person is, whether or not a person goes to the movies. … The frequency of customers over time and getting to making an effort to get to know them … would take away from randomness. And that would make it so there wasn’t an equal, the same probability for every single person in the population to go to the theater.

The other two students who held an image of a repeatable process and who did not accurately specify conditions of the process seemed to think that random meant anything could happen. Student 715 said:

> If you sampled it over the course of time, you’d see that you have so many people and it would go back and forth between maybe a group of 5 friends came one night and you knew all 5 plus the 2.5 people, then you’d know 7.5 people one day. And, maybe nobody came one day, so you’d know zero. So the more you sampled it you’d find variation between the 2.5 mean, which I guess would be the expected value of people you would know.

Students 325 said, “There’s nothing special about the movies, so people are going to the movies just because they wanna go to the movies. So, that’s why I said those, like, that the probability one person going to the movies is all equal, like.” Student 325 also thought that sample bias was inevitable and explained:

> Everybody would, like, I guess some people are more likely to go to the movies and some people aren't as likely to go to the movies. So, that's the way I see the sample, large samples. … In terms of random, I'm thinking … there's nothing special about the group. They're just regular people.

Four students who did not specify conditions of the process and did not hold an image of a repeatable process for the theater problem struggled with how to think about people coming to the theater and how to think about randomness in the context of this problem. Student 428 said, "Some people are coming that know Anthony. People, still definitely not a random sample I would think. You know people, some people never go to the movies, some people go, like, multiple times a week." Student 811 said, "In most towns, if you have 30,000 people, it's not likely that they will all have the same chance of going to the movie on a given night." Student 811 also indicated thinking about possible bias: "So, if the sample isn't random, it changes the actual amount of people he would, the actual population you're choosing from."

These students also struggled with making assumptions related to the theater problem. Student 1016 said, "I completely guessed on this." Student 428 said, "I was just speculating, like, well this couldn't be true. There wasn't a lot that I thought I could know." When asked what assumptions he made, Student 811 replied:

> I had to assume … he saw all 250 people that went to the theater. From what I've seen of ticket takers, they don't always notice the people going through. So, if it's not random, the same people come back to the theater. He has a much higher chance of knowing somebody.

Image of Distribution of Outcomes

This section describes evidence of students' conceptions of an image of a distribution of outcomes for each of the two problem contexts. First, an image of a distribution of outcomes is depicted for the hospital problem context. Following this depiction are descriptions garnered from students who revealed an image of a distribution of outcomes for the hospital problem and descriptions garnered from students who did not reveal an image of a distribution of outcomes for hospital problem. Next, an image of a distribution of outcomes is depicted for the theater problem context. Following this depiction are descriptions garnered from students who revealed an image of a distribution of outcomes for the theater problem and descriptions garnered from students who did not reveal and image of a distribution of outcomes for the theater hospital problem.

Hospital problem distribution of outcomes. Repetition of a stochastic process produces a collection of outcomes, and associating a numerical value with some characteristic pertaining to the outcome of a stochastic process results in a random variable. In the context of the hospital problem, the outcomes are births, where each birth results in a boy or a girl. Observing the number of boys born each day produces a numerical value that is a value of a random variable. This random variable represents the number of boys born at a hospital on a given day. For each day, there is only one value of the random variable. Because repetition of the stochastic process will produce a collection of outcomes, it will also produce a collection of values of the random variable. The random variable is discrete, and several days may yield the same value of the random variable. Thus, there will be a frequency distribution of values of the random variable representing the number of boys born at a hospital on a given day. If relative frequencies are considered, the distribution of the relative frequencies of the number of boys born at a hospital on a given day can be modeled by the binomial distribution.

Liu and Thompson (2007) define having an image of a distribution of outcomes as anticipating that the relative frequency of outcomes will have a stable distribution. In the context of the hospital problem, there are two distributions pertaining to the two hospitals. For the smaller hospital, the random variable representing the number of boys born each day can take on the values {0, 1, 2, …, 15}. For the larger hospital, the random variable representing the number of boys born each day can take on the values {0, 1, 2, …, 45}. A stochastic conception of the hospital problem entails an image of each of these distributions as well as the distinctive features of each distribution. The means and the variances of the two distributions differ, thus the central values and variability of the values of the random variable differ for each distribution. The important distinction is that the variability in the number of boys born each day is greater for the smaller hospital than for the larger hospital.

The hospital problem states that the record of births for the two hospitals was kept for one year. Thus, the stochastic process is repeated 365 times over the course of a year. A stochastic conception of the hospital problem entails understanding that over the course of a year, there will be stabilization in the frequencies of the random variable representing the number of boys born each day. Because the hospital problem frames the question in terms of which hospital is more likely to record more days where 60% or more of the total deliveries for that day were boys, the random variable representing the number of boys born each day must be transformed to sample proportion. Hence, each hospital is represented by distribution of sample proportions.

Hospital problem image of distribution: Stochastic conception. The five students who held stochastic conception of the hospital problem indicated thinking about distributions of outcomes, which were births. Student 815 stated, “The average of the outcomes is going to be different. … There tends to be more in the middle (for the larger hospital), and the less (the smaller hospital) is tend to be whatever." All five students mentioned that it was important to think about the difference between the hospitals in terms of sample size, which was the number of babies born each day at each hospital. They also compared the two cases: "First, I started trying to figure out what the difference would be" (Student 214). These students perceived that hospital size was a distinctive difference, as well as an important aspect on which to focus when thinking about distributions of outcomes. For example, Student 113 said, "The trials will be the number of deliveries each day. And, I look at this as hospital one, being experiment one, and hospital two as being experiment two. You have 45, [that is] $n = 45$, and $n = 15$ respectively for each of the trials that you have here." Student 325 explained, "So, the larger hospital, it would be a greater probability of basically anything. Like, it would be closer to the typical standard, the normal average. Like, just because there was more. It was larger, and I know it in terms of the Central Limit Theorem, … has the mean closest to the basic mean, to the, like, 50% or the 60."

The students who held a stochastic conception conceived of two distributions, where each distribution was associated with one of the two hospitals. Student 325 explained her thinking about the two distributions: "So, since there's only like 50. So the average is like in the middle. And, 15, there, should be … you know, like, all over the place." These students indicated understanding that there would be less variability in the distribution for the larger hospital than in the distribution of outcomes for the smaller hospital. For example, Student 817 said:

> I was thinking about, [the] smaller hospital has a more lesser sample size so it would be more extreme. … It only has 15; gonna be more of a bias. … Larger [hospital] more even; smaller [hospital] tends to be more random. There's a lot fewer numbers. … It's a gonna be more susceptible to changes.

Student 329 explained that he was thinking about variability and a different distribution associated with each hospital:

> [For] the larger sample, there'd be less of a sampling distribution because it's larger. Whereas the hospital that only has 15 births each day, that would have much larger distribution. ... I guess more of a distribution would mean the mean fluctuates more. It would have a bigger standard deviation.

The students who held a stochastic conception also indicated that the expectation for number of boys born at each hospital differed and was connected to variability. For example, Student 113 said:

> I wanted to know what the deviation could look like, or how often it could deviate from expected. You know, like I said earlier, about the lab with flipping the coin and percent. The deviation from what you're expecting in the experiment can become greater or lower, you know, with smaller samples.

Student 329 also related his thinking about expectation, variability, and distributions to what he had learned about distributions in the lab assignments: "Just that the 45 babies would more closely followed the normal distribution, just because of what we learned in the labs." Student 329 explained how the labs influenced his thinking:

> We had the one … where you were flipping the coin and heads and tails are going to even out more closely as we increase the samples. So, just the normal distribution would be that the mean is actually 50%. So there'd be 22½ babies in the one hospital and 7½ in the other, boys each day.

Student 329 also talked about variability in the sampling distribution. Student 329 explained his thinking about the relationship between variability and sample size:

> The larger sample, there'd be less of a sampling distribution because it's larger. Whereas the hospital that only has 15 births each day, that would have much larger distribution. ... I guess more of a distribution would mean the mean fluctuates more. It would have a bigger standard deviation.

Four of the five students who held a stochastic conception of the hospital problem indicated thinking about stabilization of the distribution of relative frequencies over a large number of trials. Student 329 said:

> I think relative frequencies is, … if you keep doing the same sample over and over and over again and the same, like, experiment, that you would get some type of consistency in your numbers. So just if you keep doing this every single day and going back to how much the mean, the sample mean, would differ from the expected mean.

Similarly, Student 214 explained:

> At early stages it's more likely to be different than the 50%. But, as time goes on, even if it's not exactly 1 to 1, each side will have streaks where it will be more of one than the other, and then it's end up evening out, … probably at sample sizes in the thousands. ... Well, the assumption I was just talking about is that it gets closer as you get to larger samples.

Student 113 thought about what may be unusual in the short run as opposed to the long-run relative frequency. Student 113 explained his thinking and recalled what he experienced when running simulations for one of the stochastic lab assignments:

> If you were flipping a coin, if you flip it 10 times you'd say, 'Oh, well, most often you'll have 5 heads.' But you could go ahead and flip it just 10 times and can all of sudden get 7 heads. All of a sudden 70% heads, and you say, 'Whoa, wait a second. That's not the probability of that happening.' And that's why with a smaller sample space, it's you have more, (pause) its' more likely that'll happen. But if you have a lot and lot of trials, it's more likely that it'll be evened out ratio wise. I remember in one of the simulations we did, … I saw the ratio itself was closing, but because of the vast number of trials, the real difference [between the number of heads and tails] was becoming really large. So, I thought that was fascinating.

Of the four students who indicated that they were thinking about stabilization of frequencies in the long run, two talked about the law of large numbers. Student 113 explained:

> The law of large numbers, I mean, this is why insurance really works. … Once your sample gets bigger, larger and larger and larger, it very, very, very, seldom that it'll deviate from you know, the expected, or what's expected in the

> population. You know, like we saw in the *Fathom* experiment. You know, you run the trial in *Fathom* thousands and thousands of times, and the line is really, really, really close to what you expect to be, 0.5. So I sorta felt like 45 is getting closer to [the] law of large numbers than 15 is. So that was another thing I was thinking about.

Student 325 wrote on her quiz: "It is stated in the law of large numbers that with a sample size greater than 30, the sample average will more likely be the town average than a sample with less than 30." When asked to explain how the law of large numbers was related to her thinking about the hospital problem, she said, "Like, intuitively, it makes sense if you have more and it's larger, more trials, or whatever. It's, um, you're going to be closer to, like, what you want." When prompted to think about the law of large numbers, the other students replied they had not learned anything about the law of large numbers in the course.

Students with stochastic conceptions also mentioned thinking about an underlying model, when explaining their thinking about the hospital problem. For example, Student 214 connected the outcomes to flipping a fair coin: "Here it's binomial again, so like coin flipping you're either going to have a success, which you can call a head or a boy here, or a failure, which would be a tail or a girl. At early stages it's going to be more different than 50%." Student 329 related her understanding of sampling in the hospital problem to thinking about flipping a fair coin and running simulations using *Fathom*:

> … Like in the lab where we did it and we increased the sample size each time it got closer to 50%, which is what it actually was. So I figured that the smaller one was like when we only flipped it one or two times, whereas the 45 was closer to 100. … The 45 babies would more closely follow the normal distribution, just because of what we learned in the labs.

Student 325 related the hospital problem to a physical model involving a distribution of weights. When asked to explain what she meant when she stated, "The larger hospital would have a more even distribution due to its large sample while the smaller hospital probably recorded more days with the male birth of 0.6 because it's small," Student 325 drew a picture with dots representing weights and explained:

> I was thinking of it more in terms of physics. And like if you have to balance something out and you put balancing in the middle. Like, you can do if they are evenly distributed. But if it's something like weight distributed and all in weird positions, you kinda have to move that average. So the average is going to change for the smaller one, (pause) there is less like points and there, (pause) it's not balanced.

Hospital problem no image of distribution. Seven of the interviewed students did not hold an image of a distribution of outcomes. Of these seven, four students held an image of a repeatable process and the other three students held no images related to a stochastic conception of the hospital problem.

Two of the students who held an image of a repeatable process indicated that they thought about the normal distribution and focused on expectation as a fixed value. Student 711 explained:

> I guess I'm thinking of, like, E of X, which would be about 7.5 for the smaller hospital and 27.5 for the larger hospital. So then, the probability is 50%. I mean the probability is just what's expected. So, I guess expected probability is really redundant. … The probability is the same. Well, it tells us the probability is the same in both hospitals, but I guess I assumed the variance would be proportional along the normal curve.

Student 313 struggled with how to think about expectation and said, "So, I always approach a problem imagining what if the conditions were ideal, what the value would be." Student 313 also said that she was "just using the expected value to calculate how many boys and girls there will be, and it gives me a basis for where to start off.: Student 313 also explained that she was most comfortable focusing on calculations:

> That's how, not how we were raised, but how we were taught. Just that the expected value is that, [and] everything in math and statistics, not statistics,

> mostly math, is concrete. And that the expected value, I feel like, I'm always looking for a specific value. Not uh, the fact that there is uh, the fact that there's a probability of error is kinda hard for me to contemplate.

One of the other two students who held an image of a repeatable process also thought about the normal distribution but did not mention expected value. Student 1016 explained: "I just thought about in terms of how, like, when you have a greater sample size, you have more of a normal distribution."

One of the students who held an image of a repeatable process thought about distribution as meaning allocation. Student 725 explained:

> I mean distribution as deliberately distributing the babies, so where the moms are distributed to the hospitals. That becomes the distribution and then you take your sample from that distribution. … So the distribution has to be random, as well as the samples that you are getting from the two separate hospitals that are part of the distribution.

The students who did not hold any images related to a stochastic conception focused on calculations. One student talked about distribution and said, "If you're going to go ahead and actually compute it, then there'd just be the calculation in itself. Another distribution, and you know, for the binomial the expected value is dependent upon the number, the sample. So everything depends on the sample in binomial" (Student 428). Also, the other two students who did not indicate any stochastic conceptions did not indicate thinking about distribution when solving the hospital problem.

Image of distribution: the theater problem. A stochastic conception of the theater problem requires understanding that the process of repeatedly filling the theater is a random process that produces a distribution of outcomes. Each night the theater is filled represents a random sample from the population of the town. The number of people Anthony knows each night varies and can be considered a discrete random variable, which can be modeled by a binomial probability experiment with n = 250 and p = 0.01. Plotting the distribution of values of the random variable yields a frequency distribution. A plot of the relative frequencies produces a density distribution with $\boldsymbol{\mu = E(X) = np = 2.5}$.

A stochastic conception requires understanding that there will be stabilization in the frequencies of the random variable representing the number of theater attending people whom Anthony knows each night. The proportion of people at each night's movie whom Anthony knows is a sample proportion. There will also be a stabilization of relative frequencies for the random variable representing the proportion of people at the theater whom Anthony knows each night. Students were also asked to determine if it would be unusual for Anthony to see at least two people he knew. A stochastic conception requires an interpretation of unusualness that includes a conception of an underlying repeatable process. A stochastic conception also involves understanding that the notion of unusualness is quantifiable and is a statistical value (Liu & Thompson, 2007).

Theater problem image of distribution: Stochastic conception. Four students indicated a stochastic conception of the theater problem and held an image of a distribution of outcomes. In addition to conceiving of the number of people Anthony would see each night as an outcome of a random process, these students talked about the expected number of people Anthony would see and mentioned thinking that number would vary. Student 113 said, "Let's say the movie happened every night… . There'd be quite possible that on one or two nights you'd see no one, but it'd be really close to 2.5. So, the fact that he'd seen 2, I thought, you know, that would be close to what would be expected." Students with a stochastic conception recognized variability in the outcomes and connected thinking about variability with an expected value. For example, Student 313 said, "I would expect Anthony would know 2½ people, so 2 or 3. … I got an expected value, and randomness says that I can have from 0 to 250 people that Anthony knows." Student 313 also explained, "Even though we can calculate expected value, we'll never know exactly what that, we can never predict exactly how many people are going to come."

When asked to explain their thinking about distribution in relation to the theater problem, students who indicated a stochastic conception mentioned thinking about the normal distribution model. Student 313 said, "Again, I'm looking at this from the same curve I'm used to." Pointing to a bell-shaped curve that he had drawn, Student 313 continued his explanation: "And that would probably be 2.5 here, and the distribution lies [here]. As the amount of people increase that he knew, probability would decrease. As the amount of people decrease to not knowing anyone, the probability would also decrease." Student 214 said, "I think what I got here was the expected average or mean [number] he would encounter per night. And then for the standard deviation, I would use that with the average to figure out where his estimate of 2 lies on a normalized curve."

All four students with a stochastic conception of the theater problem quantified unusualness as deviation from the expected value or mean of a distribution model. Student 313 said, "Unusual would be (pause) not near the expected value. … Not being at the expected value [is] having a certain amount of variability." Student 113 explained, "Unusual would mean deviation from what is really expected, a result that is seemingly far away from what you would expect to see. … What would be unusual would be him knowing no one since the sample size is fairly [large]. 250 is a fairly large sample even though there is 30,000 people. I say that he knows nobody would be, that would be unusual, or if he knows 10 people." Student 113 explained that he was thinking about the normal distribution model because the sample size was large: "You know, a lot of samples were, they were normally distributed." Student 214 explained his thinking about unusual:

> So then I thought that he said that he noticed 2 people every night or so, and that is below average. So, it would happen more than half the time. So then, it wasn't too strange. … I believe that the number of people he sees on average compared to what you would expect dictates what is usual or unusual.

Theater problem no image of distribution. Nine of the interviewed students did not hold an image of a distribution of outcomes for the theater problem. Two of these students held an image of a repeatable process and specified conditions of the process for the theater problem. Three students only held an image of a repeatable process, and four students did not indicate any images related to a stochastic conception of the theater problem.

The student who held an image of a repeatable process and specified conditions of the process indicated thinking about the expected number of people Anthony would see. He also indicated thinking the number of people Anthony knew on any given night would vary, however he did not indicate coherent understandings related to a distribution of outcomes. Student 817 thought that "there'd be the greatest incidence where he's going to know 1, and there'd be a lesser at 2. … He's going to know at least one person. That's what I was getting at." Student 817 also explained that "unusual" means, "I guess, having a really low probability," and said an example of unusual "is having no person that's he is going to recognize."

The three students who held an image of a repeatable process for the theater problem struggled with how to think about distribution in the context of this problem. Student 325 said, “For me, this problem, like, I assume that it’s like, it is, like, normal distribution. Like, that people, like, there would be a large group of people who, like, would want to go to the movies.” When asked to explain how one would use normal distribution in this problem, Student 325 replied, “I’m trying to think of that now, and I feel like I completely left that on the final.” Student 711 conceived of distribution as having something to do with probability:

> Um, distribution is like, how, is how your sample (pause). If you’re looking at the probability of a certain number, then your distribution shows, like, the probability. Kinda like the probability of each different possible, possible number for each different … so, like these diagrams would be distributions.

Student 711 also mentioned thinking that “unusual” would mean “unexpected.” Student 711 explained: “So, especially because you’d be doing the expected values. Like, unusual goes to unexpected, which would mean not average.” Student 715 thought that distribution referred to allocation in the context of the theater problem: “I think about distribution along with randomness is the idea that there are people distributing to the movie theater.”

The four students who did not indicate any images related to a stochastic conception of the theater problem also struggled with how to think about distribution. Three of these students said that distribution had something to do with probabilities. Student 528 said, “I use the word distribution in a way of, (pause) of just the likelihood of just each person getting to the theater. Not in terms of, (pause) not in terms of a graph, like a bell curve.” Student 811 said, “The chance, I guess, the distribution for me is that [in a] graphical sense. I only see distributions as graphs. It’s the graphical sense of probabilities." The fourth student could not articulate the meaning of distribution in the context of the theater problem. Student 428 said, "We learned a lot about distribution. Again, I have a concept of what a distribution is, but not a clear definition."

Analysis of Stochastic Conceptions

The analytic framework, which is described in Chapter 3, was used to investigate interview participants' written work on the conceptual quiz and thinking exhibited during their end-of-course interview. Responses to each problem on the conceptual quiz were analyzed, as were students' responses to interview questions aimed at probing notions related to stochastic reasoning. This analysis aims to characterize interview participants' thinking across the three categories described in the analytic framework: image of a repeatable process, image of specification of conditions, and image of distribution. Results of the analysis frame four ways of reasoning about probabilistic situations: stochastic, situational, nonstochastic, and no image. The following are summaries and analyses of characterizations of these four ways of reasoning across the three categories of the analytic framework.

No Image

Interview participants who did not exhibit any images related to stochastic reasoning are described as holding no image related to a stochastic conception. Three students did not exhibit evidence of stochastic reasoning in any of the three categories of the conceptual framework, either on their written answers or when questioned about notions related to a stochastic conception during the interview. Two of the three students were in the calculus review lab group and one student was in the stochastic reasoning group. The characterizations that follow portray their non-stochastic perceptions in each of the three framework categories.

Image of repeatable process. Students who are characterized as holding no image related to a stochastic conception of the two quiz problems did not recognize the problem situation involved repetition of a repeatable process. They did not describe outcomes of a process and did not indicate thinking about repeated sampling. Rather, these students indicated thinking about probability as a fixed ratio or rate. In addition, they focused on numerical calculations and a strictly mathematical approach to solving the problems

No consideration of outcomes. Students who are characterized as holding no image related to a stochastic conception did not describe outcomes of a process, nor did they think about repeated sampling in these contexts. Instead of thinking about outcomes in the context of the hospital problem, they thought about birth rates. For example, Student 518 explained, "So I just looked at the problem that I was stuck with a birth rate per hospital and accounting for days that are 60% or more. And, given that, those two pieces of information, that it came down to, should simply be just determining how many births per day would need to take place instead of trying to worry about percentages." Student 811 also approached the hospital problem by thinking about birth rate determined from data collected over a large number of days:

> Well, if it's an average per day that implies that this data was taken over a large quantity of days. And, in my mind if there's an average, even average, integer average, my mind automatically assumes that it's either a very small quantity or a very large amount of days were averaged together in order to achieve that. Of course, I don't know if they would word it half a baby is born a day.

Students characterized as having no image related to a stochastic conception did not consider "outcomes" when thinking about solving the problems. Student 428 struggled with the notion of "outcome." When asked what the outcomes would be, Student 428 replied, "Outcomes? What do you mean?" And when prompted further she said, "I still don't really understand." When Student 811 was asked to explain how he thought about outcomes in the context of the theater problem, he answered, "Doing the problem I wasn't really thinking about outcomes." Then, Student 811 said that he perceived three possible outcomes in the theater problem context: "Outcomes in this case were: Anthony noticing he saw at least two people, or not noticing two people, or noticing less than two people." Student 518 correctly explained what a theoretical outcome would be the context of the theater problem; however he did not think about outcomes as something generated by a repeatable process. Student 518 said he was thinking about combinations of successes and failures:

> An outcome to me in the simplest form, if I had to write it out or discuss it would be a series of successes and fails. If I actually had to put it on paper, I would picture it as, I don't really want to do this but like 250 spots, and one specific outcome would be just a number of S's and F's. And, an S representing someone that Anthony does know and an F representing someone he doesn't know. And one specific unique outcome would be for example, 1 S and the rest of them are F's. ...I'm thinking of combinations, because combinations and order doesn't matter if one S is in the first seat or the third row, or fifth row, sixth seats to the left or something. So that, that's one specific outcome. It's just as likely to happen, like I said, if the S was in the first seat or the S got there later and was in like the second row. That to me is an outcome, just a series of passes and fails, at least in that situation.

Student 518 also talked about using proportions and "simple math" to solve the theater problem and did not consider a repeatable process. Although he articulated a theoretical understanding of "outcome," he appeared to confuse finding the probability associated with a random variable with finding the probability of an outcome. First, Student 518 said, “So this one I looked at, he said he knew at least 2 people. So, I looked at it, in my head the probability of, like, *X* being greater than or equal to 2.” Student 518 went on to explain, “If I would’ve had a table or more of a calculator problem (pause), worried about trying to find, like, out of a large number *n* in this case would’ve been 30, 30,000, or 250, find the values and use the complement of. But, I knew I didn’t really have that option. So, I knew that’s not the approach that I should have or could have taken.” Then, Student 518 explained that he took a mathematical approach using combinations, “So, I was using 30,000 choose 250… And then I explained that since he knows 1 out of every 100 students, which means to have at least two people he knows required at least a 1 in a 1,000 shot at having two people show up.”

Probability as a fixed ratio. Students who held no image related to a stochastic conception indicated thinking about probability as a fixed ratio or percent. For example, Student 811 said, “I still visualize probability as fractions.” Student 811 that he thought about probability as a “set ratio” in the context of the hospital problem:

> I guess I look at probabilities on a basis of, I look at it as a fraction of something, that it’s a set ratio. In this one, it said 50% are born in the town are all boys. So, it’s obvious that any of the data, it should be that ratio is going to hold for both cases, small and large because that was a population.

Student 518 also explained thinking about a fixed value for probability in the context of the theater problem:

> There’s 30,000 people in the town and he says he knows about 300. So, I looked at, okay, so out of the 30,000 people was the actual probability of him knowing a person, which worked out to be about 1 in 100. So, 1 in a hundred people he should have known by name.

Focus on mathematical calculations. All three students who are characterized as holding no image related to a stochastic conception indicated taking a mathematical approach to solving the problems on the conceptual quiz. For example, Student 518 explained that he approached the hospital problem “like any math problem” and used calculations:

> The way I think about math and probability is the same as way I think about anything… And that’s why I write down different things like 45 a day. Just write down all the information that I believe is relevant. And then like any math problem that I look at, it has 45 babies were delivered each day in the larger hospital, smaller hospital is 15 births each day, and then calculating the 50% born in the town are boys and just working from there. …I know 50% of 45 is 22 ½, so they would need somewhere above that number to get to 60%, while half the number of the smaller hospital would be 7.5. And, I did the math and it’s 9.27.

When talking about the hospital problem, Student 428 said:

> You know how to compute the probability of the something being greater than something for binomial. So it’d be like a pretty straight forward thing to do. …If you’re going to go ahead and actually compute it, then there’d just be the calculation in itself.

Student 811 explained that he focused on numerical calculations in order to reach a conclusion in the context of the theater problem:

> In this case, probability is that ratio of proportion of people, like the proportion of people in the theater, or proportional of the population. I guess, I still look at the probability to that fraction of a value. …When I went through this problem, I basically took all the numbers in context, pulled that value out and then worked on it. Completely numerical, then plugged that back into the problem.

Student 518 also focused mathematical calculations when he explained his thinking about the theater problem:

> So, it’s just the chance of getting in. And, just that’s just my simple math, had to determine what my notes were. So I used the (pause), since order wouldn’t have been any issue, I was looking at permutations versus computations. Because, since there was only 250, you had to figure out how many different ways are there to fill the 250.

Image of conditions of the process. Students who are characterized as holding no image related to a stochastic conception of the two quiz problems did not specify conditions relevant to a repeatable process. These students did not indicate thinking about a collection of variable outcomes. Although some of the students mentioned independence, they were reluctant to make assumptions related to the problem situations. Students who held no stochastic image did not think sampling was relevant to these problem contexts other than thinking samples should be a random samples. They did not indicate thinking about sample size in relation to variability, but thought about sample size as a number used in calculations. These students indicated the importance of noting if the sample size was large because that meant one could use the normal distribution.

Assumptions. Students who held no image of a stochastic conception did not indicate thinking about a collection of variable outcomes. Because they did not conceive of a repeatable process, they did not make any assumptions related to a process. However, when prompted during the interview, these students indicated thinking about some assumptions they used to solve the problems.

Two of the three students who held no image related to a stochastic conception expressed a reluctance to make assumptions. For example, Student 428 said, “I really try not to make assumptions, which is why I had trouble answering the question.” She explained that she felt more comfortable working the problems she encountered in the class because those problems were more clearly defined with regards to assumptions. “They tell you straight up.”(Student 428). Student 428 also said, “For me it was, you don’t know this, and you don’t this, and you don’t know this. So, how can you assume this? But again, it completely changes if you make some probably inheritable assumptions.” Student 518 said that he tried to avoid making an assumption in the hospital problem context; thus ignored some information that was given: “So that there’s equal chance of getting a boy or a girl per births. So, I didn't even try to assume anything about the chance of getting a boy or girl, just by ignoring that piece of information."

Although these students were reluctant to make assumptions, Student 428 hesitantly said she assumed the births were independent:

> Well, they don't tell you anything though. I mean, like, you could assume, like, it's just like, I know it sounds probably ridiculous, but, like, they don't, I don't know, like, say that the births are independent, like. Which I mean, obviously they mostly probably would be. But what about, like, triplets?

Student 518 did not assume the births were independent, but said he assumed each hospital was independent:

> But, I considered them independent because each hospital is giving birth to a set number of boys per day, and they're not expecting any kind of a roulette. So, I just considered them automatically independent, and without worrying about having to take out some kind of overlapping number of births.

Student 811 said that he only made two assumptions in the hospital problem context: "I'm assuming in this problem, I assumed that there were, that everybody had an equal chance going to each of the hospitals, and that there's an actually even chance of boys to girls which is kinda off for a town."

In the context of the theater problem, Student 428 was asked to explain what underlying assumptions or conditions she was thinking about when she solved the problem and said, "I was just speculating, like, well this couldn't be true. There wasn't a lot that I thought that I could know. …I really didn't know what to do with it." When prompted to explain her thinking about the problem if she could make assumptions, she said, "I guess if you make the assumptions again, then you could probably do another binomial thing. You're just choosing the number of people of the theater, the probability that Anthony knows then and the probability that Anthony doesn't know them." Both

Student 518 and 811 talked about making assumptions related to people coming to the theater and using the given proportion to determine the probability. Student 518 stated:

> My assumptions as I read through, I took them as assumptions but I probably should have taken them as what they told me. But, my assumptions were everyone had an equal chance of showing up at the theater. But, they did say later assume that people were not coming to the theater because they know Anthony. I assumed that the probability of him knowing one person would also translate into the smaller sample population of the 250 that would be in the theater. So I assumed that if he had a 1 out of a 100 chance of knowing someone out of 30,000 people that he, that, could somehow translate into the smaller population of just the 250.

Student 811 explained the assumptions he made in the context of the theater problem:

> In this case, the probability is that ratio of the proportion of people like the proportion of the people in the theater or proportional of the population. …Obviously, I also had to assume in order to see the people he knew, he saw all 250 people that went in the theater. From what I've seen of ticket takers, they don't always notice the people going through. I believe those were the only two assumptions in this case.

Sampling. Students who held no stochastic images in the problem contexts did not indicate thinking about a sampling process and did not talk about repeated sampling. However, these students talked about a random sample. For example, Student 811 understood the importance of taking a random sample: "Without a completely random sample where everything has an equal likelihood of occurring; probabilities are not an accurate portrayal of the samples. So, if there is no randomness, probabilities not accurate." In the context of the theater problem Student 811 thought about taking a random sample:

> So if the sample isn't random, it changes the actual amount of people he would, the actual population you're choosing from. So, if it's not random, and the same people come back to the theater, he has a much higher change of knowing somebody.

Student 518 also thought about taking a random sample in the context of the theater problem:

> And the theater could hold 250, so we are taking at any given time a random sample of 250 people. Making, and also that each person that arrived at the theater was independent. There was no one who was more influenced to go to the theater over anyone else, so the chance of one person going was just good as the chance of another person going.

When talking about a random sample in the theater problem context, Students 811 and 518 contrasted taking a random sample with taking a sample that was biased. Student 428 did not think people going to the movies in the theater problem context could be considered a random sample: "No, I wouldn't think it would be [a random sample]. I mean. I don't really know if it is. I said that because I thought I had to have a definite answer and it bothers me that I couldn't think of anything." Student 428 appeared to struggle with the notion of randomness and could not explain what "random sample" meant to her.

None of these students mentioned thinking about a random sample in the hospital problem context. However, in addition to mentioning the need to take a random sample, two of these students indicated thinking that the results of sampling were only useful if an extremely large sample was taken. For example, Student 518 explained that she had doubts sampling produced meaningful information when samples were small:

> To me sampling is, I guess for the lack of better words, I just don't like it. And, it goes back to the whole randomness, what random means to me. To me, I, and I understand it's, like, for example we go to a polling question. And, this is one of the big things about polling, if you take a poll of a specific population you have to use a random sample because talking to 20,000 to 30,000 people is very expensive, very time consuming. I understand why sampling is important, but to me it's not very [important]. It's efficient, but it's not very accurate or meaningful. Taking a population of 50 people out of 40,000 to me is nothing, because of the randomness.

Student 428 talked about a sample that would be infinitely large: "I don't know, I guess it's as *n* goes to infinity then you're, like, approaching the natural pattern." Student 428 also explained, "When you have a big sample then you can start to predict entire populations based on that sample. …There are a lot of things in math that approximates when you have really big things that are natural and we really don't know. So that was the really the only thing that stuck out." Both of these students thought that only large samples were useful and neither student indicated an understanding of sampling.

Sample size. Students who did not hold stochastic images did not indicate thinking about variability in relation to sample size. Furthermore, these students did not indicate that thinking about sample size was particularly important. For example, when asked if sample size was important to thinking about the hospital problem, Student 528 said, "Sample size would be important, but unfortunately it's not very. It's not very clear how, what the sample size would be representing." Student 811 explained, "The 50% was the population proportion because that's all the children born in town, not just on that given year. The way I looked at it, so, at that point the data population sample sizes is much less than important for a proportional." Student 428 mentioned thinking about sample size in the hospital problem context but not in relation to variability: "It's the only thing, you know, is different between the two hospitals. I mean, logically it's what you look for, the difference in."

These students did not indicate that thinking about sample size was important other than it was a number used in calculations, or that a large sample size meant one could use the normal distribution in calculations. All three students talked about using sample size for calculations. For example, in the context of the hospital problem, Student

528 said, “Expected value could be, depending on what we’re talking about, it could be as simple as just the number that we’re referring to, of the sample size times the probability of it happening. And, that would be the expected value.” In the theater problem context, Student 811 explained:

> Because the sample size is 250, the distribution was not particularly important to me, because the law of large numbers says anything over 35 can be approximated by a normal. So, I just assumed it was normal because the sample size was that large.

Student 428 also connected thinking about a large sample to the normal distribution: “Your big sample, you can use a normal to approximate it, which is nice.”

Image of distribution. Students who are characterized as holding no image related to a stochastic conception of the two quiz problems did not indicate thinking about a distribution of outcomes. These students struggled with how to think about distribution in relation to the problem contexts. They indicated thinking about expectation as a fixed value to calculate and did not indicate thinking about variability in relation to expectation or the notion of distribution. These students conceived of distribution as having something to do with probabilities, but were unable to articulate key notions relevant to distributions. They conceived of distributions as merely graphical representations or tables useful for calculating probabilities. Furthermore, these students recalled very little about the law of large numbers.

Notion of distribution. Students who held no image of a stochastic conception did not indicate thinking about a distribution of outcomes. However, all three students mentioned the normal distribution. In addition to talking about the normal distribution, one of the students talked about the binomial distribution, and the others mentioned the Poisson distribution, the students’ *t* distribution, and the Chi-squared distribution. These students thought about distributions as curves, graphs, and tables. Two students also commented that graphs of distributions represent probabilities. For example, Student 811 explained, “When I’m thinking about a normal distribution… I’m always visualizing the distribution curve and looking at the values on it. So the fact that it’s normal to me means that the degree of error in it is represented by *z*-values.” Student 811 also explained, “The distribution for me is that graphical sense, because I only seen distributions in graphs. It’s a graphical sense of the probabilities.” Student 518 also talked about representations of distributions:

> But, then coming to this class, I learned a different way of distribution by using the tables we were given. And, we learned like the Chi-squared, the [students’] *t*, the [normal]. And now distribution in that sense is, is probability in the sense of, like, if you have a population of maybe 300 people, the distribution spans the probability of just the one person being success all the way up to the 300 people being the success. And the graph itself, the distribution, just represents the probabilities of each instance.

These students remarked that distributions were visual representations useful for calculating probabilities. Student 518 perceived distribution as a shortcut for determining a probability:

> And distribution to me is more of, uh, it’s just the way that I learned them at class, it’s more of a shortcut. They’re specifically designed in terms of statistics class being, of being used for a set number of objects in a population and the probabilities of a success in the situation. Now you can use the probability distribution, the probability distribution defines what the value of some smaller number of that population is. I also know they can be used to represent the overall probability of it happening.

These students seemed to understand how to use a distribution to determine probabilities, but struggled to explain their thinking about the concept of distribution. For example, Student 811 said:

> I look at a distribution, I guess, the defined set of that curve based on the probabilities of something being around its mean or standard deviation.

> …Distribution, what a large enough sample has to be continuous. Has to be continuous otherwise it doesn't work out in mathematical equations.

Student 428 was forthright about her lack of understanding of the notion of distribution:

> I mostly think about the ones that we learned about. We learned about a lot of distributions. Again, I have a concept of what a distribution is but not like a clear definition. …It's just some sort of, uh, some sort of pattern I guess. Like, I don't know, you have some sort of experiment that follows some set of rules and then it has, like, basically you can reduce something to a common experiment.

Expectation. Students who held no stochastic image did not connect thinking about distribution to variability. They focused on numerical calculations and spoke about expected value, the variance or standard deviation, and wanting to determine an interval. In the context of the hospital problem, Student 518 explained:

> I wasn't really thinking about expected value, but as I remember my statistics class, I mean, expected value could be depending on what we're talking about if it could be as simple as just the number that we're referring to of the sample size times the probability of it happening. And, that would be the expected value, which I think I was unconsciously going for by taking the 60 and multiplying it by like 50%.

Student 428 said that she was thinking about intervals and that "you were more likely to be closer to what you expected when you had a bigger sample. And, here you had one, like 15, which is a lot smaller than 45, so it made sense kinda you would deviate a little. You'd be more likely for that to happen in the smaller one." Student 428 explained her perception of deviation in terms of a limiting value, not random variability: "It's just basically, like, you have a sample that approaches infinity. Your sample expected value is going to be, really it's going to approach, the population expected value." Student 811 said that he wanted a formula to use in the context of the theater problem:

> So that's why I pulled the expected value out. And, then after that I generalized… Well, in this case I didn't remember the formula for the error, so I couldn't actually run the test. ...The expected value is that mean value of that

> sample which it lies somewhere in the distribution, in the actual value of the population.

The three students who did not hold any images related to a stochastic conception thought about expectation as a calculated number and did not think of expectation in relation to random variability.

Law of large numbers. Students who did not hold stochastic images said they learned about the law of large numbers, but did not indicate an understanding of the law of large numbers. None of these students indicated thinking about stabilization of frequencies. Student 811 expressed thinking about law of large numbers in relation to large samples, but appeared to confuse the law of large numbers with the Central Limit Theorem: "Because the sample size is 250, the distribution was not particularly important to me, because the law of large numbers says anything over 35 can be approximated by a normal." Student 428 said her professor mentioned the law of large numbers in class, "but I don't remember it." Student 518 remembered encountering the law of large numbers in lecture class but said:

> My experience working with the law of large numbers was minimal. …I recall learning the rules and how to work with it, but it wasn't tested enough for it. We didn't work with it enough that I'd just, once we moved on to something new that I knew was going to be more important, I just completely forgot about it.

These students could recall the terminology, but none of them could give an indication of the meaning of the law of large numbers.

Unusual. All students were asked to think about "unusual" in the context of the theater problem. The three students who held no stochastic images thought about the notion of "unusualness" in different ways. Student 518 stated that unusual meant "not likely," and described her thinking about unusual in terms of a flipping a coin:

> The chances of getting a head then a tails would be usual. But then to me, if you were going to flip a coin a hundred times, [and] you got heads a hundred times that would be very unusual. I looked at it in terms of a real life situation, what would be considered unusual in real life. And, that's just how I thought about it.

When asked about unusual in the context of the theater problem, Student 428 declared, "Oh, this is hard. I had no idea what to do with this. I felt like this is not unusual, but I don't know how to think of why." She also that her answer to the theater problem would change depending on whether or not she made assumptions: "It'd probably be unusual. It looks unusual. …If you don't make assumptions, then I feel like it's not that unusual. I don't know. I would have had a different answer if I assumed lots of stuff." (Student 428). Student 811 thought about "unusual" in relation to an expected value. Student 811 said it would not be unusual for Anthony to see 2 people he knows because that value is close to the expected value:

> In this case there would be a greater than 50% chance that he would know two people there. ...A probability test on this it would definitely be greater than 50%. Of course I can't run a standard deviation on this, but it's obviously that is what it's asking.

Nonstochastic Conception

Three interview participants are characterized as holding a nonstochastic conception because they did not exhibit stochastic reasoning across all three categories for either of the two problem contexts. These students exhibited evidence of stochastic reasoning in only one category of the conceptual framework (an image of a repeatable process) for either one or both problem contexts on the conceptual assessment and for some interview probes. They did not indicate an overall stochastic conception for either quiz problem context. Furthermore, they did not exhibit evidence of stochastic reasoning when questioned about notions related to a stochastic conception. All three of these students were in the calculus review lab group. The characterizations that follow portray their stochastic perceptions in the image of a repeatable process category, as well as their nonstochastic perceptions in each of the three framework categories.

Image of repeatable process. All three students who are characterized as having a nonstochastic conception held an image of a repeatable process in at least one of the problem contexts on the conceptual quiz. Two of these students held an image of a repeatable process in both problem contexts, and the remaining student only held an image of a repeatable process in the hospital problem context. These students talked about a process that involved repetition under the same conditions and described outcomes of the repetitive process. None of these students indicated thinking about a repeated experiment, however two of these students connected thinking about sampling to the repetitive process. As students who held a nonstochastic conception, they conceived of probability as a ratio in both problem contexts. All three of these students indicated using this ratio in calculations, however, the two students who held an image of a repeatable process in both problem contexts said that they did not focus solely on numerical calculations. None of these students connected thinking about the process to a model.

Repetition under the same conditions. Students who held a nonstochastic conception indicated thinking about a process that repeated under essentially the same conditions. In the context of the hospital problem, all three students talked about records of births on more than one day. Student 715 said that one would need to consider sample sizes and “you would need how many days you would be doing the study over.” Student 715 also connected the notion of sampling to the repetitive process that he was thinking about: “So they’re talking about the sampling when they talk about being more or less likely the larger hospital recorded more days of boys being born. So, sampling you’d obviously go in and ask them, ‘Did you have more boys on this day?’ And then you would put a ‘yes’ or a ‘no’ kind of thing.” Student 711 and Student 1016 also indicated thinking that each day represented a repetition of the process. For example Student 711 explained, “The overall probability of the baby being born in the town being a boy is 50%; and thus you would expect that about 50% of the babies born each day would, then in each hospital, would be 50%. So, then it was talking about the days in which the number of boys born was 60% or more, um, and so that was in comparison to the 50% that was the expected probability of boys born each day.” Student 1016 said that he approached the hospital problem in a logical way: “There’s more births here and there’s less births in the smaller hospital, so the smaller hospital will have more days from which there’s a greater number of boys born, because of a greater percentage, like the ratio, because just one birth can alter the ratio. And that’s why I said the smaller hospital.”

In the context of the theater problem both Student 715 and Student 711 indicated thinking that the repetitive process consisted of people coming to the movie theater on nights that Anthony is working. Student 715 mentioned repeated sampling: “The idea [is] that you would have this many people, and if you sampled it over this course of time, you’d see that you have so many people. And, it would go back and forth between maybe a group of 5 your friends came one night and you knew all 5 plus the 2.5 people, then you’d know 7.5 people one day. And maybe nobody came one day, so you’d know 0.” Student 711 also connected the notion of sampling to repetition of the process, but made the assumption each day at the theater was not a random sample: “And if it’s not random, and people are coming frequently, and one day he sees two people he knows, um, another day he realizes it is quite frequently more than that." Student 1016 did not indicate thinking about a repetitive process in the context of the theater problem.

Outcomes of the process. All three students with a nonstochastic conception described outcomes of the process in the context of the hospital problem as babies being born. Student 715 mentioned thinking about "a random chance that you have all boys born that day or you have no boys born that day." Student 711 indicated thinking about "days in which the number of boys born was 60% or more," and said, "The 50% for each individual is the probability of the individual's birth." Student 1016 explained, "I was just thinking of an example. So, let's see, if one day you have the average per day for the small hospital, and there's 8 boys and 7 girls, so that's about 50% percent. …When you just look the boys, and you change it to 9 boys and 6 girls, just one birth difference, then you get that 60%."

In the context of the theater problem Student 715 and Student 711 described outcomes as being the number of people Anthony knew at the theater that night. For example, Student 711 explained, "But, so in terms of it being skewed, certain people come frequently and others just occasionally. So if two people he knew came, like one night, they may be more likely than other people to come more frequently because they, I mean, they like going to movies. They go there more often." Student 715 also described outcomes of the process as the number of people Anthony knew and said: "He either knows 0 people or anything between 0 and 250. I mean he could theoretically know all 250 people in the theatre because he knows 300 people in the town. So, anywhere between 0 and 250 would be relevant, I mean, besides the half numbers. The discreet numbers in the sample situation would be the sample solution." Student 715 also indicated that the outcomes would vary from night to night: "Well, you wouldn't go there one day and say, 'How many people do you know this one night?' and base your entire experiment on one sample."

Conception of probability as a fixed ratio. When asked how they conceived of probability in the context of the hospital problem, all three students who held a nonstochastic conception indicated thinking about probability as a ratio or percent. Two of these students also indicated thinking that this ratio could vary. For example Student 711 explained, "Well, it says that generally 50% is the average probability, but on several days it was much more than the probability. …Like, the overall probability of the baby being born in the town being a boy is 50%, and thus you would expect that about 50% of the babies born each day, would then in each hospital, would be 50%. So, then it was talking about the days in which the number of boys born was 60% or more, um, and so that was in comparison to the 50% that was the expected probability of boys born each day." Student 1016 said that he thought of the percentage or ratio of average births per day and that ratio could change: "So the smaller hospital will have more boys born because if a greater percentage, like the ratio, because just one birth can alter the ratio. And that's why I said the smaller hospital. …Because you're looking at the ratio of guys versus girls for both hospitals and that's the question, if the ratios are, like, 60% or so." Student 715 explained that he conceived of probability as a fixed ratio in the context of the hospital problem: "Um, I think of probability in the form of percent. There is a percent. A percent is representative of a portion of a sample and that portion of a sample implies the probability that that portion will be picked over another portion, that's not included in the percent. … I read the problem and the first thing I said was, if it's a percentage, then it would scale and logically I wouldn't need a statistical evaluation to determine that that 60% would be 60%, whether its 45 or 15."

All three students who held a nonstochastic conception also indicated thinking about probability as a fixed ratio in the context of the theater problem. For example, when asked how he conceived of probability, Student 1016 said, "As a ratio, pretty much. I guess when I did the ratio first between the people he saw and the people in the town, when I got that I saw that's the probability of the people he saw. Like a dice, when you throw a dice, you get ½. So, I got the probability of success." Student 1016 also explained: "I thought that's what probability was. You're guessing what the probability is of a random event happening. You're trying to estimate randomness." In the context of the theater problem, Student 711 also explained that he conceived of probability as fixed value to use in numerical calculations: "Yeah, so in terms of probability, um, I guess that's kinda [how] I see the unusual thing. Like, 'What's the probability that he knows at least 2 people?' Just like, 'What's the expected value, so the probability?' I mean it. You were to do it purely mathematically then. …so that's how you would use the probability to look at this." Student 715 also explained thinking about probability as a fixed value: "Well, there's the probability that, you know, how many people he knows in town. Probability-wise he knows every 1 out every 100 people in town. So, if a 100 people walk by he should theoretically know 1 person. And so speaking out of probability, if he has 250 in the theatre, then he should know 1% or 2.5 people."

Numerical calculations. Students who held a nonstochastic conception indicated that they grappled with how to answer the conceptual quiz problems and whether to logically solve the problems or use mathematical calculations. These students seemed to contend with reasoning through the problem contexts while simultaneously applying their understanding of mathematics, probability, and statistics. For example Student 711 said:

> So, I was like, after I did them, I looked around I hadn't used any math for either of these problems. And people were like taking out their calculators and like doing the problems out. 'Oh my God did I do them wrong?' But, both the questions didn't explicitly ask for numbers. I guess, whenever I can, I try to use logic instead of the numbers because a lot of the times I'll find myself and like, like, just studying negative binomials, and I realize, the formulas are, like, really easy because it's really obvious if you just think about it. But, if you don't and you're just memorizing it, then it's so hard to actually, like, like, memorize it because it doesn't mean anything to you.

Student 1016 said that he used logic to solve the hospital problem: "I just thought, what I thought was logical. I didn't see this as probability." Student 1016 also mentioned that he conceived of random as "something completely out of your control." When he explained how he figured out his answer to the theater problem, Student 1016 said, "I completely guessed on this." Then he elaborated as he explained that he "took the ratio of people he knew" and "did it by binomial." Student 715 indicated thinking about "probability that's calculated" versus probability "in practice." Student 715 said, "I think about them together. I mean probability is the chance that what you expect to happen will happen. So I can expect to win the lottery, but the probability of that happening is virtually nill, that's not going to happen. So, I can go get a lottery ticket and I can crunch the numbers and get a probability of oh, 0.05% of winning the lottery. Then, I win the lottery and that means I won it but it still means I had a 0.05% chance of winning. It wouldn't change that number." Student 715 did not think her experiences in the class influenced her thinking about the conceptual quiz problems:

> The first question was influenced very little by [this class]. That became more of a logic problem to me. The second problem, probably the only thing I ever would

> have thought about comparing to statistics, not specifically [this class], is the idea that the percentage related to the population of the people in the theater. That would be about it.

Image of conditions of the process. Students who are characterized as holding a nonstochastic conception did not consistently specify conditions relevant to a stochastic conception of a repeatable process. They did not indicate thinking about a collection of variable outcomes. However, all three students who held a nonstochastic conception mentioned sampling. Two of these students indicated thinking about repeated sampling, and both students also talked about random sampling. All three students indicated making assumptions when solving the problems. Two students mentioned thinking about independence. They assumed the samples would be independent, but did not assume outcomes would be independent. All three students who held a nonstochastic conception mentioned thinking about sample size. However, they did not indicate an understanding of the implications of sample size in relation to a stochastic conception. These students talked about the need to take a large sample. They also indicated thinking about the impact of sample size and ratios on the expected value. Students who held a nonstochastic conception indicated thinking about variability in the expected value that resulted from different sample sizes, but did not indicate thinking about variability in outcomes.

Sampling. All three students who held a nonstochastic conception indicated that they thought about sampling. Two of these students indicated thinking about repeated sampling and talked about a sampling process. For example, Student 715 explained:

> Well, you wouldn't go there one day and say, 'How many people do you know this one night?' and base your entire experiment on one sample. The more, um, I guess, credibility a sample has is based on how many samples are taken and how large the sample size becomes, and the fact that there's still a trend there.

In the context of the theater problem, Student 715 said, “So, you wouldn’t, like, you wouldn’t sample one night or five nights. You would sample maybe 20 or even more than that.”

Student 715 talked about repeated sampling in both problem contexts, while Student 711 only talked about repeated sampling in the context of the theater problem. Student 1016 did not indicate thinking about repeated sampling, but thought about sampling only as it related to sample size: “The only way I applied it is making the sample sizes. You sample a couple of births per hospital.”

The two students who indicated thinking about repeated sampling also mentioned thinking about random sampling. Student 715 said that random meant “the sample isn’t biased.” She assumed random sampling occurred in both problem contexts. For example, in the hospital problem context, Student 715 explained:

> Completely without bias and chosen arbitrarily with no form one way or the other. …um, that’s the way I assumed they would be. They were talking about the likelihood that the larger hospital recorded more days, which says to me it’s all on a basis of chance. So, out of chance, there’s a random chance that you have all boys born that day or you have no boys born that day, but out of the average, you’ll have about 60%. So for randomness, if it’s biased in any way then they could pick which days they go, and pick to go on the days they have 100% boys born. And then it wouldn’t be random.

In the context of the theater problem, Student 715 explained that randomness was relevant: “…you are assuming that the people that come into the theater are random and chosen by random. Because if they’re not chosen by random, and say a group of friends come in, then it biases the sample towards him knowing more than he should.”

Student 711 talked about random sampling and indicated understanding what it meant to take a random sample. Student 711 explained, “Random is when you have, when you choose your sample at random so that every person in the population has equal, an equal probability of being chosen for the sample." Although Student 711 indicated that it was important to take a random sample, she did not perceive each sample of 250 people attending a movie was equally likely. Student 711 explained:

> And so with these 250 people, not everybody in 30,000 people have an equal chance of being chosen because there are all those factors influencing whether or not a person, um, whether or not a person goes to the movies. … And, that would make it so there wasn't an equal, the same probability for every single person in the population to go to the theater.

Hence, Student 711 did not perceive the context of the theater problem involved random sampling. Student 711 said, "…I made the assumption that, like, with this not random sample. There is, um, like, the likeliness that his friends are to attend, like, a movie, or people he knows that attend a movie."

Assumptions. Students who held a nonstochastic conception indicated making assumptions when thinking about the problem contexts. Student 715 and Student 1016 mentioned thinking about independence in the context of the hospital problem. Student 715 assumed the samples were independent: "Independence I would view as being a part of the random sampling. If you have a random sample, then the samples are independent of each other, and one of them isn't biasing the other." Student 1016 said, "The births between the two hospitals are independent, and the difference between guys and girls is independent." Thus, both Student 715 and Student 1016 assumed samples would be independent, but did not indicate assuming that outcomes of the repeatable process were independent. Student 711 did not indicate thinking about an assumption of independence in either problem context.

When asked what assumptions or conditions they were thinking about when solving the conceptual problems, all three students who held a nonstochastic conception mentioned an assumption related to “distribution.” Student 1016 assumed the binomial distribution could be used to determine the probability in the context of the theater problem:

> I assumed that you could use a binomial distribution with this because that’s the only way I saw it. When I was working the problem I was thinking how I would approach it, and that’s the only thing that popped into my head. I guess when I did the ratio, first between the people he saw and the people in the town, and when I got that, I saw that’s the probability of the people he saw. Like a dice, when you throw a dice, you get ½. So I got the probability of success.

Student 711 assumed the normal distribution was appropriate in both problem contexts. In the context of the hospital problem Student 711 said, “The underlying assumption I had was that it was normal. The probability of, like, boys being born was normal of the babies delivered each day. …Then you could follow the variance would be along the normal curve.” Student 711 also talked about a normal distribution in the context of the theater problem: “So the ‘N’ is like if it was normal. Right, the first time I read it, it would be normal.” Student 715 also stated that she made an assumption about distribution, but her understanding of distribution was not related to a probability model. Student 715 stated, “Obviously, the sample and the distribution are random,” however, she conceived of “distribution” as meaning to dole out and assumed a random allocation of mothers to the two hospitals. Student 715 explained, “I mean distribution as deliberately distributing the babies, so where the moms are distributed in the hospitals. That becomes the distribution and then you take your sample from that distribution.”

Sample size. Students who held a nonstochastic conception indicated thinking about sample size, but did not indicate understanding implications of sample size in relation to a stochastic conception. Rather, they perceived the importance of considering sample size in the problem contexts. Students who held a nonstochastic conception exhibited a variety of perceptions as to why sample size was important to consider in the problem contexts.

Two of the students who held a nonstochastic conception focused on the birth ratio for each hospital in the hospital problem context and talked about the effect of sample size on the birth ratio. For example Student 1016 indicated thinking about "average births per day" and said:

> There's more births here [the larger hospital], and there's less births in the smaller hospital. So, the smaller hospital will have more days from which there's a greater number of boys born, because of a greater percentage, like the ratio. Because just one birth can, like alter the ratio more easily that one birth can alter here. And that's why I said the smaller hospital.

Student 1016 did not indicate thinking about repeated sampling but thought the birth ratio would vary: "The ratio is also random as the births are random." Student 711 indicated thinking about repeated sampling and also explained that she was thinking about the effect of sample size in the hospital context:

> An outlier would affect the average more. So that, if you had, I mean, if you're delivering 45 babies, and like you deliver 5 more boys than expected, then… . The percentage of boys isn't going to be much higher as opposed to if you deliver 5 more boys than expected in the smaller hospital. That's, like, a huge number. That will make a huge difference. …Yes, the sample size is important because one has 15, as I was explaining before.

Furthermore, Student 711 indicated thinking about how sample size would affect the variance:

> That's what I was trying to explain before, that's the variance proportional to each. To each, like, extra male births or extra female births is, like, it creates much more variance or a higher variance than if you have a higher sample size. So, if there are 45, then like, if you have an outlier, that's not going to alter the variance as much as if you have a smaller sample size.

When asked to specify what "variance" she was thinking about, Student 711 replied, "The variance of the population or the sample size, like, the variance of the data sets, basically." Although Student 711 appeared to confuse the notions of "sample" and "population," she perceived that the sample size affected the "average" birth ratio.

The other student who held a nonstochastic conception indicated thinking about repeated sampling but did not think it was necessary to consider sample size in the context of the hospital problem. Student 715 perceived that the percent of births at each hospital was a fixed number: "So, it's a percent of the sampling that you're doing. So, of the sampling you're doing, there's the same chance that you will always, that you will always have that 60% on that day in that hospital." Student 715 explained:

> Because you're talking about percentages, the sample size shouldn't be necessary, because the percent is, because it scales to whatever the percent of the population is. So, whether you have 45 or 15 babies, you'll still have the same percentage of that sample. So, the sample size shouldn't matter unless there is a large variant in one hospital versus the other of having more boys, which is unlikely.

Student 715 and Student 1016 also indicated thinking that it was important to have a large sample. When explaining her thinking about sampling in the context of the theater problem, Student 715 said that a large sample size contributed to the credibility of the data collected: "The more, um, I guess, credibility a sample has is based on how many samples are taken and how large the sample size becomes, and the fact that there's still a trend there." Student 1016 indicated thinking a consideration of sample size was important because that determined whether or not it would be appropriate to use the normal distribution. Student 1016 explained his thinking in the hospital problem context: "I just thought about it in terms of how, like, when you have a greater sample size, you have more of a normal distribution. But, when you have a smaller sample size it's more skewed positive or negative. That was my thinking in saying there's more variability in the smaller sample size."

Image of distribution. Students who are characterized as having a nonstochastic conception did not indicate thinking about a distribution of outcomes. Two of these students thought about the normal distribution and described thinking about variability in conjunction with distribution. All three students who held a nonstochastic conception indicated thinking about variability in relation to expected value. However, they did not indicate thinking about variability in relation to outcomes of a process. Two of these students did not recall learning about the law of large numbers, and the other student did not think the law of large numbers applied to the problem contexts. Students who held a nonstochastic conception did not conceive of "unusual" in relation to a distribution model. They indicated thinking that "unusual" meant not the expected value.

Notion of distribution. Although not thinking about a distribution of outcomes. Student 1016 and Student 711, talked about the normal distribution, and Student 1016 also indicated thinking about the binomial distribution and the Poisson distribution. Student 1016 said he thought about sample size when deciding which distribution to apply: "You would need to think when you have a large sample size like this, is it large enough for it to be normal, so you can apply normal to it." When explaining his thinking about the theater problem, Student 1016 was not sure which distribution model to use. At first Student 1016 said, "I assumed that you could use a binomial distribution with this, because that's the only way I saw it." Student 1016 also remarked that the sample size was large, so he thought about the normal distribution:

> I guess, normal distributions, because you have 30,000 people and you have 300. Well, 300 is a pretty large sample size, so you would (pause). Wow, I already

> forgot everything. …I remember you use binomial, or Poisson, or anything else if the sample or population is very skewed. And, you would use normal once you have a larger population. And, no matter what that would be normal.

Student 711 also talked about applying the normal distribution. In the hospital context, Student 711 said, “The probability of, like, boys being born was normal of the babies delivered each day. So if they, the probability was 50%, then you could follow the variance would be along the normal curve. …The probability is the same, well it tells us the probability is the same in both hospitals, but I guess I assumed the variance would be proportional along the normal curve.”

Both Student 1016 and Student 711 talked about shape and mentioned graphical representations when describing their thinking about distributions. These students also alluded to thinking about variability in conjunction with distribution. For example Student 711 explained, “If you’re looking at the probability of a certain number, then your distribution shows, like, the probability, kinda like the probability of each different possible number for each different, yeah possible (pause). And so, like, these diagrams would be distributions. And, like, ultimately all of this would add up to one. But it also shows like, like kinda the spread and the, like, gives you an idea of any, any skewing.” Student 1016 remarked that he thought about sample size, distribution, and variability when solving the hospital problem: “I just thought about it in terms of how like when you have a greater sample size you have more of a normal distribution but when you have a smaller sample size it’s more skewed positive or negative. That was my thinking in saying there’s more variability in the smaller sample size.”

The other student who held a nonstochastic conception indicated thinking that “distribution” meant to dole out items. Student 715 stated, “My first thought when it comes to distribution is the definition of the word. It's literally just the distribution of items, things, samples. So, you have people and items that are distributed generally randomly. I mean there could be a bias, but out of that distribution usually a sample becomes of that distribution." Student 715 alluded to "distribution" as meaning allocation when she described her reasoning in both problem contexts. For example, in the hospital problem context, Student 715 explained, "The only random variable would probably be the distribution into the hospitals, because the random variable would be taking out the dependence on the bias. So, if you distributed the births randomly, you'd have the random variable, I guess, between the two sample sizes. But, but then I guess you couldn't distribute them randomly between each sample, it would be like switching beds." In the context of the theater problem Student 715 stated, "I think about distribution along with randomness, the idea that there are people distributing to the movie theatre. And, then it would be a random sample, if the sample were to be not biased to him knowing more or less people."

Expectation and variability. Students who held a nonstochastic conception indicated thinking about expectation and expected value in both problem contexts and stated that the expected value was the mean or average. These students did not indicate thinking about variability in relation to outcomes of a process, but they mentioned thinking about variability or variation in relation to the expected value. Students who held a nonstochastic conception compared the expected value with varied hypothetical results. When Student 711 and Student 1016 talked about expectation and expected value in the context of the hospital problem, they mentioned thinking that the mean number of births could vary. For example Student 711 explained his thinking about expectation and expected value:

> So, then it was talking about the days in which the number of boys born was 60% or more, um, and so that was in comparison to the 50% that was the expected probability of boys born each day. By expected probability, just means that, like, it would be the mean, the μ, so the expected probability. I guess I'm thinking of, like, E of X which would be about 7.5 for the smaller hospital and 27.5 for the larger hospital. So then, the probability is 50%. I mean, probability is just what's expected.

Student 711 also talked about variance and indicated thinking about how the value would be affected by an outlier: "It's more likely the smaller hospital will record more such days. It's because in terms of variance, an outlier would affect the variance of the, like, an outlier would affect the average more." Student 1016 talked about expectation as "the average per day" in the hospital context and gave an example when explaining his thinking. Student 1016 also mentioned how variability in the number of boys born on any given day would change the average:

> If one day, you have the average per day for the small hospital, and there's 8 boys and 7 girls, so that's about 50% percent since it's an odd number, and that's the same for the larger hospital because it's an odd number too. When you just look the boys, and you change it to 9 boys and 6 girls, just one birth difference, then you get that 60%.

Student 1016 also connected thinking about variability and expected value when he said: "Like, expected values, you expect the smaller hospital to have a greater variability."

In the context of the theater problem, Student 711 indicated thinking about what might occur in practice compared to the mathematically determined expected value. She grappled with reasoning through the problem or using her knowledge of mathematics to solve the problem. Student 711 explained:

> Yep, it's just natural, like, if you see people more you're going to, like, just feel like you know them better, you're going to make an effort to meet them. That's

> more like a social thing. But, like, it is one of those things to take into account when you're looking at random sample, I guess. …And thus, like, um, if he's often seeing at least two people he knows, like, if one day he sees two people he knows, like, that maybe more random as opposed to if, um, if, like, every day, if later on he sees two people he knows. Also, I mean, I didn't write this, but also mathematically, like, you look at it, if there's 30,000 people and he knows 300 hundred of them, then that's going to be, like, he knows 1 out of every 100 people and 250 people come a day. So, like, he'd know 2.5 people, like, on average anyways. So, I mean, that's the mathematical way of looking it.

Student 711 also indicated thinking about the mean, as well as a "random distribution" and shape of a distribution, but she exhibited difficulty coordinating her thinking about these notions. Student 711 wanted to either use mathematics or consider randomness and distribution:

> So I thought that originally, and then I read the question. And, like, because it, like, it said be sure to address issues of randomness and distribution, I looked at those instead of using the math that you know. You would naturally know 2.5 as expected if it was a random distribution, a random and normal distribution. But so, in terms of it being skewed, certain people come frequently and others just occasionally. So if two people he knew came, like, one night, they may be more likely than other people to come more frequently because they, I mean, they like going to movies. They go there more often.

In addition to talking about a mathematical way to solve the problem versus a logical way, Student 711 also said that she noticed other students used calculators for the conceptual quiz and found this somewhat troubling. Student 711 explained that she did not use math to solve the problems on the quiz:

> After I did them, I looked around. I hadn't used any math for either of these problems. And people were like taking out their calculators, and, like, doing the problems out. 'Oh my God did I do them wrong?' But, both the questions didn't explicitly ask for numbers. I guess, whenever I can, I try to use logic instead of the numbers.

Student 715 connected thinking about expected value to standard deviation, and thought about using the standard deviation to determine an interval in the hospital problem context. Student 715 explained;

> Um, [the expected value] would be a more statistically relevant number than the 60%. It would give you a better mean to go off of, if you were doing the study, because it would tell you what the value you expected to have was. So, if you had that over 60% or more you could say, ‘I expect this many babies to happen.’ And, if you have a standard deviation you could say within this and this many babies. …I guess the probability is the 60%. Then, the expected value would give you the more precise definition of the probability than you’ve already given.

Student 715 explained how she coordinated probability theory with what could actually happen in the context of the theater problem. Student 715 also mentioned thinking about repeated sampling and variation of the mean, i.e. the expected value:

> That’s how you get into standard deviation. The idea that you would have this many people, and if you sampled it over this course of time you’d see that you have so many people. And, it would go back and forth between maybe a group of five of your friends came one night and you knew all five plus the 2.5 people, then you’d know 7.5 people one day. And, maybe nobody came one day, so you’d know zero. So, the more you sampled it you’d find the variation between the two-and-a-half mean, which I guess would be the expected value of people you would know.

Thus, Student 715 indicated thinking that the variation was somehow within the mean itself.

Law of large numbers. Two of the three students who held a nonstochastic conception did not recall learning about the law of large numbers. The other student, Student 1016, indicated that he about the law of large numbers, but did not think this idea applied in the hospital problem context. Student 1016 explained that his understanding of the law of large numbers was connected to thinking about sample size and when to apply the normal distribution: “That was Chapter 4, I think. It says that if you have a population and you take a sample from that, if you increase the sample size, the distribution becomes like the bell shape, normal. I don’t think it applies to this problem, because of the sample size.”

Unusual. All interviewed students were asked to explain their thinking about "unusual" in the context of the theater problem. Students who held a nonstochastic conception did not quantify "unusual" as deviation from the mean. They conceived of "unusual" as meaning unexpected or unlikely. Student 1016 said, "So, the fact that he saw someone he knew every night was very unusual, and so I said that's very unlikely." When asked to give an example of "unusualness" Student 1016 replied, "Very, very unlikely." Student 711 and Student 715 also indicated thinking that "unusual" meant not the expected value. For example, Student 711 conceived of "unusual" as "not average" and stated, "Unusual would mean, like, unexpected. Like, what would be expected and, like, what if? So I, especially because you'd be doing expected values, like, unusual goes to unexpected, which would mean, like, not average." Student 715 also conceived of "unusual" as not the expected value. She said that the expectation could be considered "normal" so "unusual" would be "not normal." : Student 715 explained:

> Unusual being there's, I guess, something you expect to occur and you're asking whether or not the expectation of that occurrence is normal or unnormal. Abnormal I guess is the word I'm looking for. So, unusual would be by definition not normal or the expectation of something that is happening that is abnormal.

Situational Conception

Three interview participants exhibited images of stochastic reasoning across all three categories for one of the two problem contexts on the conceptual assessment and for some interview probes. They are described as holding a situational conception because they exhibited evidence of stochastic reasoning for one of the two problems on the conceptual quiz, but not the other. They did not consistently exhibit evidence of stochastic reasoning when questioned about notions related to a stochastic conception. All three of these students were in the stochastic reasoning lab group. The characterizations that follow portray their stochastic perceptions, as well as their non-stochastic perceptions, in each of the three framework categories.

Image of repeatable process. All three students who are characterized as having a situational conception held an image of a repeatable process in both problem contexts on the conceptual quiz. These students talked about a process that involved repetition under the same conditions and described outcomes of the repetitive process. Only one student indicated thinking about a repeated experiment. However, all three students who held a situational conception indicated thinking about repeating sampling. Although none of these students spontaneously indicated thinking about a model of repetition, when prompted via questioning, two of the students indicated thinking about a specific model in relation to the conceptual quiz problem for which they held a stochastic conception.

Repetition under the same conditions. Students who held a situational conception talked about a process that involved repetition. For example, in the context of the hospital problem, Student 817 said, "There's gonna be a lot of number samples." Student 325 mentioned that "there were less births each day" and also indicated thinking about repetition as a number of trials: "Intuitively, like, it makes sense. If you have more and it's larger, more trials or whatever, it's um, then you're going to be closer to, like, what you want." Student 325 related her thinking about repetition of a process to repeated trials in chemistry: "With chemistry we do more trials to be at the point where we want." In the context of the theater problem Student 313 talked about outcomes as being "from 0 to 250 people that Anthony knows" and that there will be a "range of different values." Student 313 considered repetition as involved in a random process, “So the randomness allows me to say that there will be, the results can differ depending on the situation. …We’ll never know exactly what … we can never predict exactly how many people are going come.”

Outcomes of the process. Students who held a situational conception described outcomes of the process. In the context of the hospital problem, these students understood that outcomes would be the birth of a boy or a girl. Student 325 talked about “births each day,” and Student 817 said, “I guess, one is not going to affect the other, the outcome, … like, the gender.” Student 313 also talked about outcomes as births in the context of the hospital problem: “So, we’re going to see that 45 babies a day in hospital one. So, we’re going to have more children born.” In the context of the theater problem Student 313 said, “The outcomes (pause) would be, the outcome I was looking for was the fact that two people were attending the theater. Yeah, that’s pretty much it.”

Repeating an experiment. None of the students who held a situational conception used the word “experiment,” however Student 325 alluded to thinking about repeating an experiment when she talked about doing repeated trials in chemistry. She explained, “Um, like, intuitively I, like, it makes sense if you have more and it’s larger, more trials or whatever. It’s, um, then you’re going to be closer to, like, what you want, like you want. With, uh, chemistry, we do more trials to be at the point where we want. That’s what I think of the average as being, the point that you want it to be at” (Student 325).

Repeated sampling. Students who held a situational conception talked about repeated sampling. For example, Student 313 talked about a “wide range of samples, and that they can vary” and indicated thinking about sampling in the context of the hospital problem: "…'cause we're taking the sample of two hospitals. So, I definitely see the sampling." In the context of the hospital problem Student 817 described thinking about a number of samples: "There's gonna be a lot of number samples. It would start to even out but then it's doesn't have, it only has 15. Gonna be more of a bias on." In the context of the theater problem Student 325 indicated that he thought about taking more than one sample. "Everybody would, like, I guess some people are more likely to go to the movies and some people aren't as likely to go the movies. So, that's why I can see the sample, large samples."

Model of repetition. All three of the students who held a situational conception were in the stochastic lab group and ran simulations. Yet none of these students spontaneously indicated thinking about a model of repetition. However, when asked about the stochastic lab assignments two of the students indicated thinking about a model of repetition in relation to the conceptual quiz problem for which they held a stochastic conception. For example, Student 313 held a stochastic conception of the theater problem and said, "I know for problem number two, [the theater problem], I definitely, it was, I can think of the first or second lab where we were using coin flips and that whole lab. Just, there was one part where we had to calculate the expected, how many. One of the questions was how many people would we, how many. I guess, I think if I can remember back was, like, how many coin flips would you expect to be either heads or tails. And that's how I approached this problem." Student 325 held a stochastic conception of the hospital problem and said, "I think it was the pennies, and it was, I thought about that in [the hospital] problem."

Simulations using Fathom. Student 817 said that he remembered "parts of those *Fathom* labs" and that "they were helpful in terms of getting a general knowledge or general understanding of the subject." However, he did not mention thinking about a model or *Fathom* simulation when explaining how he solved the conceptual quiz problems. Students engaged in lab assignments running simulations of sampling and sampling distributions before studying theoretical sampling distributions in class, and Student 325 said the lab assignments helped her make more sense of what she was learning in class. Student 313 said, "I feel like at least a little part of each lab applied to, I could see a little part of each lab in each problem, especially on the quizzes." Student 325 mentioned thinking about the lab assignment that involved running simulations of sampling distributions of sample means:

> Yeah, and I remember it was, like, on one of the homework quizzes we had in class that I thought about the pennies. And like, and I guess we hadn't done it in class yet or something. And, like, it made more sense in class.

Image of conditions of the process. Students who are characterized as having a situational conception indicated thinking about conditions of a repeatable process in the problem context for which they exhibited evidence of stochastic reasoning. In addition, these students may have indicated thinking about some aspects of conditions of a repeatable process in the other problem context for which they did not exhibit stochastic reasoning. Students 325 and 817 only revealed evidence of stochastic reasoning in the hospital problem context, and Student 817 only revealed evidence of stochastic reasoning in the theater problem context.

All three students who are characterized as having a situational conception indicated thinking that repetition of the process would yield a collection of variable outcomes. Students who held a situation conception made assumptions related to a stochastic conception in one problem context but not the other. Their assumptions varied. They made assumptions related to what they conceived of as "normal." They also made assumptions about randomness and independent outcomes. Furthermore, students who held a situational conception indicated thinking about repeated sampling, but did not indicate thinking about a sampling process or a representative sample. These students indicated thinking about sample size in at least one of the problem contexts. The two students who only indicated stochastic reasoning in the hospital context connected notions of sample size and variability in that problem context. Two of the three students who held a situational conception connected their thinking about conditions of the repeatable process to an underlying model in the problem context for which they held a stochastic conception. Overall, students who held a situational conception exhibited conflicting notions of randomness, probability, and sampling. They also had difficulty coordinating their intuitive thinking about probability in the problem contexts with their theoretical understanding of probability.

Collection of variable outcomes. Students who held a situational conception indicated thinking about a collection of variable outcomes in both problem contexts. In the hospital problem context, these students talked about variability in genders resulting from births. For example Student 817 explained

> I was thinking about (pause), smaller hospital has a more lesser sample size, so it would be more extreme. …It's more probable that, uh, there's gonna be a lot of number samples. It would start to even out, but then it's doesn't have, it only has 15. Gonna be more of a bias on. …Somehow a person may have like more of the same gender rather than other ones.

Student 325 also talked about variability in birth outcomes:

> For me that was the only thing like it had to be because there were less births each day. Then, you're going to get, like, it's going to be a little weird. You're going to have things further away from the, like, typical average, and so because there are less things to balance it out in a sense.

Student 313 indicated thinking about variability in number of people Anthony would know at the theater each night:

> Okay, with this problem I would say that randomness is relevant cause there's the, I got an expected value and the randomness says that I can have from, in the theater of 250 people, I can have from 0 to 250 people that Anthony knows. …So the randomness allows me to say that there will be the results can differ depending on the situation and we will, even though we can calculate expected value, we'll never know exactly what that, we'll never, we can never predict exactly how many people are going to come.

Assumptions. Students who held a situational conception indicated making a variety of assumptions. The assumptions they made differed depending on the problem context and whether or not they indicated evidence of stochastic reasoning in that problem context.

Assumptions and stochastic reasoning in hospital context. Student 817 and Student 325 indicated stochastic reasoning in the hospital problem context, but not the theater problem context. When asked to reveal their thinking about any underlying assumptions or conditions in the hospital problem context, both of these students mentioned "normal." Student 817 seemed to connect an assumption about "normal" to a situation with two outcomes. For example Student 817 said, "I guess, I thought normal, as in true or false, like, also, boy or girl. I guess, I thought having 50% percent each." Student 325 did not specifically mention making any assumptions in the hospital problem context, but indicated thinking about "normal" in relation to "average" and the central limit theorem: "Answering this problem I was thinking that the hospital, so the larger hospital, it would be a greater probability of basically anything. Like, it would be closer to the typical standard, the normal average, like, just because there was more. It was larger, and I know it's in terms of the central limit theorem." Student 325 mentioned thinking about a normal distribution in the theater problem context: "For me, this problem, like, I assume that it's, like, like, it is like normal distribution. Like, that people, like, there would be like a large group of people who, like, would want to go to the movies." However, when asked to explain why she thought the normal distribution was related to this problem, Student 325 replied, "I'm trying to think of that now, and I feel like I completely left it on the final. Um…."

In the hospital problem context, Student 817 revealed that he assumed the outcomes were independent. Student 817 wrote the following statement on his conceptual quiz: "Furthermore, since each baby has independent probability, more # less chance for the births to be biased to 60%." During the interview, when asked to explain his thinking with regards to "independent probability," Student 817 replied, "I guess, one is not going to affect the other, the outcome… like, the gender." In the theater problem context, Student 817 mentioned thinking that people coming to the theater would be "independent of each other." When asked what assumptions he made when thinking about the theater problem, after a long pause, Student 817 said, "That people coming in, I thought, should be random." Student 817 connected thinking about randomness and independence and explained: "I guess that a person coming in is going to be independent of each other, like, like picking one person out of 30,000 people. And, it's not going to affect any other people coming along with him." Student 817 also explained random meant: "I guess, having equal chance of people or person going into the theater." Student 325 did not mention thinking about independence in either problem context. However, in the theater problem context Student 325 stated, “Assumptions, um… I guess I assumed that was just, like, randomly distributed, even if I can correctly, even, like, define random distribution.” When pressed to explain her assumption, Student 325 gave an example:

> Like, in terms of random distribution I thought, like, a random sample. Like, I don’t know, like, asking if there’s a survey done. You’re asking if someone, like, if people like snow. If you’re gonna go, you know, like, up north, you’re probably gonna get more people that like snow. So then, if you go to California, you know, there’s, or if you go to Maryland, you know, you’re gonna get, like, people aren’t going to say they like snow. But, if you do the whole entire, like, you go the entire East coast, like you, it can be more randomly distributed.

Although Student 325 had a difficult time articulating what “random distribution” and “random sample” meant, she assumed that people were equally likely to go to the theater and said this comprised a random sample. Student 325 explained, “I’m thinking about it in terms because of size, um, that people aren’t just going to the theater because they know him. So, I just feel, like, it’s just kinda like a random [sample]. Like, um, people are going to the movies, like, nothing special about the movies. So, people are just going to the movies just because wanna go to the movies. So, that’s why I said those, like, that the probably of one person going to the movies is all equal, like.”

Assumptions and stochastic reasoning in theater context. Student 313 only indicated stochastic reasoning in the theater context and not the hospital context. When asked about assumption and processes relevant to thinking about the theater problem, at first Student 313 said, “Just again, just thinking about proportion. I took this as a very straight forward problem. I thought the more, the mathematical approach to this.” However, Student 313 also thought about randomness when writing his answer to the theater problem conceptual quiz because he wrote: “Randomness considers the fact that the probability will not always be completely accurate.” When asked to explain what he was thinking, Student 313 replied, "Randomness is the ability to, ability not… This is not going to sound too professional, but the ability to not be around the expected value, the ability to expand and not have a concrete answer." Student 313 also explained, "Probability, the best way I can put it would be probably, I would use probability to estimate the amount of randomness and with regards to that randomness. Randomness says I can have large amount of range of, different values from the expected value. Probability says the, gives me the probability that an event is, of occurring. Like I said, probability will tell me, probability tells me the likelihood that it's going to happen, and randomness tells me that it can happen. Then, they work hand-in-hand, I guess."

Student 313 also indicated that he struggled with how to approach the theater problem and said, "I'm not used to taking statistics classes and I'm used to getting concrete answers, so the approach is different." Student 313 also explained, "I guess, I don't know how exactly to put it into words how I take into these problems. I use mostly through, I don't know how to say, life experience, but more [than] just mathematics." Student 313 went on to explain how he was thinking about underlying processes in the context of the theater problem: "And the processes were, I guess, I say, were my weaker points in this course. I'm used to mathematics classes following equations and stuff. So, the processes were a little different. But underlying, I can't really tell you a specific, but I can say I saw what we were learning in the course in this problem."

When asked about underlying assumptions and condition in the hospital problem context, Student 313 said, "I can see that just the way that I think, I'm more about the numbers as opposed to the process, I guess you could say." Student 313 said that he also considered: "…rates, probability, frequency, independence are in there. Probably not one of them [a] major player in there, but it all was considered, [also] distribution." Furthermore, Student 313 indicated thinking that the hospitals did not affect each other and talked about a sampling rate for each hospital: "I see that it in respect to that, hospital one, hospital two, that [they] had different rates, different. They are both constrained by, I guess they are both constrained by the conditions and that 50% are male and 50% are female. So as far as independence is concerned, that's (pause), I see each hospital as independent, I guess." Although he indicated thinking that the hospitals were independent, Student 313 did not indicate thinking that the birth outcomes were independent. For example, Student 313 said that the hospitals did not affect each other and he perceived that the hospitals represented different scenarios:

> With this problem I considered two different scenarios. I didn't consider them as affecting each other. I guess that's what I'm saying. So, that's how, that's the extent to which I see independence in this problem, that they don't affect the results. One hospital's result is not going to affect the results of the other hospital.

Sampling. Students who held a situational conception indicated thinking about repeated sampling in one or both problem contexts. However, these students did not indicate thinking about a sampling process. Furthermore, they did not indicate thinking about a representative sample.

Sampling and stochastic reasoning in hospital context. Student 817 indicated thinking about repeated sampling in the hospital problem context: "There's gonna be a lot of number samples. It would start to even out but then it's doesn't have, it only has 15. Gonna be more of a bias …" However, when asked about a sampling process in this problem context, Student 817 stated, "I don't think it relates or have any effect on this." Student 817 indicated he generally thought that "sampling has to be random." Student

817 also thought about sampling in relation to probability distribution: "It depends on how you sample. The probability distribution might differ, I guess, I think." Student 817 did not indicate thinking about repeating sampling in the theater problem context.

Student 325 did not clearly indicate thinking about repeated sampling in either problem context. However, in the hospital context Student 325 indicated thinking about a repeatable process in addition to thinking about the central limit theorem. She also mentioned thinking about "sample sizes." Student 325 stated, "Yeah, and I was trying to think about large sample sizes. So, like, a larger *n* of a sample has a mean closest to just the basic mean, so like, the 50%, or the 60%." Student 325 expressed thinking about samples in the theater problem context and explained:

> People aren't just going to the theater because they know him, so I just feel, like, it's just kinda like a random [sample]. Like, people are going to the movies, like, nothing special about the movie. …So that's why I kinda see it as a sample, like just a large sample.

Sampling and stochastic reasoning in theater context. Student 313 indicated thinking about sampling in the theater problem context. He expressed thinking that each group of 250 people will differ. Student 313 stated: 'So, the randomness allows me to say that there will be, the results can differ depending on the situation. And we will, even though we can calculate expected value, we'll never know exactly what that, we'll never, we can never predict exactly how many people [he knows] are going to come. Student 313 also connected thinking about a range of samples to thinking about distribution:

> Distribution … with randomness actually works exactly with randomness and probability. That there's a wide range of samples, and that they can vary, so I'm using probability distribution. I use the two graphically, I use the graphic distribution to display how the probability will occur in that sample.

Furthermore, Student 313 did not think about repeated sampling in the hospital context, but thought about sampling rates in that context:

> I'm just factoring in all the information in, I guess, just taking it as it is, trying to approach this process: How I would set it up, if I should approach it by the rates per hospital, which should be my first step? Which, is what I was considering when I was looking at this problem.

Sample size. Students who held a situational conception indicated thinking about sample size in one or both problem contexts. The two students who only indicated stochastic reasoning in the hospital problem context connected the notions of sample size and variability in outcomes in that problem context. The student who only indicated stochastic reasoning in the theater context did not connect notions of sample size and variability in either problem context.

Sample size and stochastic reasoning in hospital context. Student 817 and Student 325 indicated a situational conception with stochastic reasoning in the hospital problem context. Both of these students indicated thinking about sample size in relation to variability in outcomes in the hospital context. For example Student 325 explained:

> Okay, so for this one specifically, I was thinking that since they are looking for a 60% number, there were more boys born, I was thinking since it was further away from the like 50% typical, like, typical births. I was thinking it would have to be the smaller one and because it was further away from the, like, the average. I was thinking that because it, just to me, it just makes sense. If you have more of something, you can have it can be more precise. So, if we're looking for the abnormality, it's not normal. Then it would have to be the one with the smaller hospital.

Student 325 said that she struggled with thinking about sample size in the theater problem context: "I guess, it's, like, 30 is a smaller sample size, or 300 is smaller sample size of 30,000 where if it's, if they're distributed, like, the same, then I don't know. I struggled with this problem, like, really." Student 817 also thought about sample size and variability in the hospital context. Student 817 explained, “I was thinking about smaller hospital has a more lesser sample size, so it would be more extreme.” Student 817 did not indicate thinking about sample size in the theater problem context.

Sample size and stochastic reasoning in theater context. Student 313 indicated a situational conception and stochastic reasoning in the theater problem context. However, Student 313 did not indicate thinking about sample size in relation to variable outcomes in the theater context.

Student 313 perceived sampling rates were important to think about in the hospital problem context. He also indicated thinking about sample size in relation to differing sampling rates for the two hospitals. Student 313 explained how he conceived of sample size and probability: “Sample size, primarily because that’ll affect the probability. The way I see it there’s three factors in these problems. There’s the rate, there’s the probability itself, and also the sample size, which all affect the, those are the three primary tools I would need to use.” Student 313 also talked about “probability of error” and related that to sample size. Student 313 explained:

> If an extra boy was born as opposed to a girl that would create a bigger difference in the ratio, boys to girls, as opposed to if one more boy was born than one more girl in the 45 children born in hospital one. So, when I was saying this, it’s because the smaller hospital has a smaller sample size and they created a probability of error. If one [more], there would be a greater change from the expected value at hospital two as opposed to hospital one. A greater range if one child, if just one boy or girl was born more than what is expected, the 50%.

Underlying model and conditions of the process. Two of the three students who held a situational conception connected their thinking about conditions of the repeatable process to an underlying model in the problem context for which they held a stochastic conception. In the hospital context, Student 325 indicated thinking about equilibrium and a model that involved an equal distribution of weight. Student 325 connected thinking about variability in the number of births each day to sample size and said:

> Because there were less births each day, then you're going to get, like, it's going to be a little weird. You're going to have things further away from the, like, typical average, and so, because there are less things to balance it out in a sense.

Student 325 explained that she was thinking about a physical model involving a distribution of weights:

> Just, like, they balance each other. Like, I was thinking of more in terms of, like, in Physics, just because I'm taking a Physics class right now, and like, if you have to balance something out. And, you put, like you, it's balancing in the middle. Like, you can do that if they are evenly distributed. Like, you can just balance in the middle. But if it's something like weight distributed, and all in weird positions, you kinda have to move the average. So the average is going to change for the smaller one because there's less, like, points, and it's not balanced.

The other student who held a situational conception and conceived if an underlying model did not specifically mention conditions of a process in connection to thinking about a model. However, Student 313 said work on the lab assignments influenced his thinking about the theater problem. In particular, Student 313 indicated thinking about simulated coin flips:

> I know for problem number two, [the theater problem], I definitely, it was, I can think of the first or second lab where we were using coin flips and that whole lab. Just, there was one part where we had to calculate the expected, how many. One of the questions was how many people would we, how many. I guess, I think if I can remember back was, like, how many coin flips would you expect to be either heads or tails. And that's how I approached this problem.

The third student who held a situational conception, Student 817, did not connect thinking about a model to thinking about conditions of a repeatable process.

Image of distribution. The three students who are characterized as having a situational conception revealed evidence of stochastic reasoning related to an image of distribution, but only in the problem context in which they indicated stochastic conceptions across all three framework categories. Students 325 and 817 only revealed evidence of stochastic reasoning in the hospital problem context, and Student 817 only revealed evidence of stochastic reasoning in the theater problem context.

All three students who held a situational conception indicated thinking about notions related to distribution in one or both problem contexts. However, these students only indicated thinking about a distribution of outcomes in the problem context for which they exhibited evidence of stochastic reasoning. Students who held a situational conception also connected notions of expectation and variability in the problem context for which they indicated evidence of stochastic reasoning. The two students who exhibited evidence of stochastic reasoning in the hospital context indicated thinking about connections between sample size and distributional aspects in that context. One of the three students who held a situational conception indicated thinking about stabilization of frequencies of outcomes and the law of large numbers. Two of the students who are characterized as holding a situational conception described thinking about an underlying distribution model in the problem context for which they exhibited evidence of stochastic reasoning. Only one student who held a situational conception indicated thinking that about “unusual” as deviation from the mean. The other two students indicated thinking “unusual” meant having a low probability. None of the students who are characterized as having a situational conception quantified “unusual” as deviation from expectation in terms of a distribution model.

Distribution of outcomes. Students who held situational conception indicated thinking about a distribution of outcomes in the context for which they indicated stochastic reasoning. Student 325 and Student 817 indicated stochastic reasoning in the hospital problem context, and both of these students described thinking about a distribution of outcomes in that context. Neither Student 325 nor Students 817 indicated thinking about a distribution of outcomes in the theater context. Student 313 indicated stochastic reasoning in the theater problem context and only indicated thinking about a distribution of outcomes in that problem context.

Distribution of outcomes and stochastic reasoning in the hospital context. Student 325 and Student 817 only indicated stochastic reasoning in the hospital problem context. Both of these students described thinking about a distribution of outcomes in the hospital context. On conceptual quiz, Student 325 wrote the following statement: "The larger hospital will have a more even distribution due to its large sample." Student 325 explained how she was thinking about distribution:

> I was thinking that because, it just to me, it just makes sense. If you have more of something, you can have it can be more precise. So, if we're looking for the abnormality, it's not normal. Then, it would have to be the one with the smaller hospital. …For me that was the only thing like it had to be, because there were less births each day. Then, you're going to get, like, it's going to be a little weird. You're going to have things further away from the, like typical average, and so because there are less things to balance it out in a sense.

When asked what came to mind when he thought about distribution, Student 817 explained that he was thinking about a distribution of outcomes in the hospital context. Student 817 also drew a picture of a frequency distribution and said:

> And the numbers, this is, like, 3, 4, and 2, like that. These [are] outcomes. … Because, you're expecting that, um, one incidence is not going to affect the other one, should be 50%. …The final outcome, or, like, the average of the outcomes is going to be different. If you have more, it tends to be in the middle and if less it tends to be in the, um, whatever.

Distribution of outcomes and stochastic reasoning in the theater context. Student 313 indicated thinking about randomness as he described a distribution of outcomes in the theater problem context. Student 313 explained:

> With this problem, I would say that randomness is relevant ‘cause there’s the, I got an expected value, and the randomness says that I can have from in the theater of 250 people, I can have from 0 to 250 people that Anthony knows. …So, the randomness allows me to say that there will be, the results can differ depending on the situation, and we will, even though we can calculate expected value, we’ll never know exactly what that, we’ll never, we can never predict exactly how many people are going come.

Student 313 also indicated thinking about a distribution model in the theater context. Student 313 explained:

> Again, I’m looking at it from the same curve that I’m used to. …I guess I could say I could use this for, this is the first thing I turn to when I think of normal everyday situations. I know that there are some problems where this would not apply, but as with regards to a problem like this, where we’re considering random, random probability, and variability, this is what I would first think about in my mind.

Expectation and variability. Students who held a situational conception connected notions of expectation with variability in outcomes in the problem context for which they indicated evidence of stochastic reasoning. Both Student 325 and Student 817 connected thinking about expectation to thinking about variability in outcomes in the hospital problem context. However, neither of these students connected the notions of expectation and variability in the theater context. Student 313 connected thinking about expectation to thinking about variability in outcomes in the theater context, but not the hospital context.

Expectation, variability, and stochastic reasoning in the hospital context. Student 325 and Student 817 connected notions of expectation and variability in outcomes in the hospital context. Student 325 indicated thinking about how outcomes might vary from the “typical standard, the normal average.” Student 325 explained thinking that there would be more variability in birth outcomes for the smaller hospital:

> I was thinking that since they are looking for a 60% number, there were more boys born. I was thinking since it was further away from the like 50% typical, like, typical births. I was thinking it would have to be the smaller one and because it was further away from the, like, the average. …For me that was the only thing, like, it had to be because there were less births each day. Then, you’re going to get, like, it’s going to be a little weird. You’re going to have things further away from the, like, typical average, and so because there are less things to balance it out in a sense.

Student 817 said, “If you’re guessing the probability of being each gender is 50%, I would guess that expectation is like the half of the births. And if you have a large sample size it would be more of a group and that expectation.”

Student 325 and Student 817 did not connect thinking about expectation to variability in outcomes in the theater problem context. Student 325 said she struggled with how to think about the theater problem. Although she mentioned thinking about the normal distribution, she did not indicate thinking about expectation in the theater context. Student 817 connected his thinking about expectation to probability and the “bell curve,” but did not indicate thinking about variability in outcomes. As Student 817 explained his thinking about expectation, he moved his hands, motioning in the shape of a bell curve: “If you’re guessing the probability of being each gender is 50%, I would guess that expectation is, like, the half of the births. And if you have a large sample size it would be more of a group and that expectation.”

Expectation, variability, and stochastic reasoning in the theater context. Student 313 connected notions of expectation and variability in outcomes in the theater context. First, Student 313 explained that in the theater context he set up a proportion to determine the expected number of people Anthony would know:

> In a theater of 250 people I set that as a proportion, one over a 100 people and *x* over 250, and I would expect Anthony would know 2 ½ people, so 2 or 3 people in the theater. So, 'Is it in fact unusual that at least two people know Anthony?' So, I said, 'no,' because I expect either 2 or 3 people. That would be expected so it's not usual to know 2 or 3 people in the theater.

Then, Student 313 also explained that he considered variability in the outcomes and connected thinking about random variability to expected value:

> …I got an expected value, and the randomness says that I can have from in the theater of 250 people, I can have from 0 to 250 people that Anthony knows. Randomness allows me to do that just because we don't know. Randomness, I guess, uh, like this problem. Let's say Anthony had a special movie or Anthony told all his friends and family about and 250 people came. So, the randomness allows me to say that there will be the results can differ depending on the situation. And we will, even though we can calculate expected value we'll never know exactly what that, we'll never, we can never predict exactly how many people are going come.

Student 313 explained that he calculated the probability and then, "using that probability, I applied it to the theater and how many people are in the theater, and how many people I would expect that he would know in the theater." Thus, Student 313 appeared to coordinate his thinking about theoretical probability with random variability of outcomes in the theater context.

In the hospital problem context, Student 313 indicated thinking about probability and expected value in a different way. Student 313 explained that in this context he considered the expected value to be a number to calculate:

> I guess, I feel that I'm used to, I'm use to through other classes, expected values. Um, and I rely on probability as the expected value, so just I always approach a problem imaging what if conditions were ideal: 'What the value would be?' Now, into this concern, with this problem: I'm just using the expected value to calculate how many boys and girls there will be, and it just gives me a basis for where to start off. And, that's even though I didn't explicitly write that down for this problem, it's what I was considering for this problem.

Student 313 also mentioned thinking about error in relation to expected value, but also said this notion was kind of confusing to him. Student 313 explained:

> Error to me would be not, I guess, it's kinda confusing. Now that I kinda see where we're going, but error is, when I said error, I meant it in respect to expected values. The error was anything that differs from the expected value would be considered the error. …I guess the way I was using it in this [hospital] context was probably the same. We usually think about variability as where there is no concrete answer. In this problem where we expect, we have an expected value. We expect half of the children born will be boys and half will be girls. That would be anything that would differ from that I would, I said to myself, was an error, but it's not an exact error. But, anything without a concrete answer probably would be variability. That's just how I established it in my mind.

Student 313 did not appear to coordinate thinking about error with random variability in the hospital context. He focused on calculations and the birth rate at each hospital, and also grappled with reconciling notions of variability and error:

> Because if they're seeing 45 babies a day, 15 babies a day, with 45 babies a day I recognize that there's going to. The first things I think about, there's going to be a greater sample size that is more. Now, I'm playing with variability and error in my head. There's the probability, the statistics concepts. …Like as I said before, I know I'm repeating myself but, if the male is born that the ratio for male to female is greater than 15 babies a day. The rates, I guess, like I said. I don't know how much of this helps your interview 'cause I'm more of a math, I just like numbers and equations and plug them in, type of person. Not as much of a critical thinking type of person. I'm just using the rates to start off my problem, this is, that, I mean, I can tell that's the place I started off with this problem.

Sample size and aspects of distribution. The hospital problem context elicited thinking about sample size and various aspects of distribution. The two students who only indicated stochastic reasoning in the hospital context talked about the relationship between sample size and aspects of distributions in the hospital context. These two students did not indicate thinking about sample size in relation to distribution in the theater context. The other student was characterized as having a situational conception and only indicated stochastic reasoning in the theater context. This student considered aspects of distribution in the theater context. However, he did not specifically mention thinking about the relationship between sample size and distribution in the theater context. In the hospital context, this student thought about the effect of sample size on the rate of boys born at each hospital.

Distributional aspects, sample size, and stochastic reasoning in the hospital context. Student 325 and Student 817 indicated stochastic reasoning in the hospital context and considered distributional aspects of sample size in that context. Student 325 talked about sample size and variability in sample means:

> I was trying to think about large sample sizes. So like a larger *n* of a sample has a mean closest to just the basic mean, so like the 50%, or the 60%. …Okay, so for this one specifically, I was thinking that since they are looking for a 60 % number, there were more boys born. I was thinking since it was further away from the, like, 50% typical, like, typical births. I was thinking it would have to be the smaller one. And, because it was further away from the, like, the average.

Student 325 also thought about a distribution of outcomes and how "it's gonna balance out" with a larger sample size. Student 325 explained that the distribution of outcomes for the smaller hospital would be more variable:

> For me that was the only thing, like, it had to be because there were less births each day. Then, you're going to get, like, it's going to be a little weird, you're going to have things further away from the like typical average. And so, because there are less things to balance it out in a sense.

She conceived of a physical model that involved balancing a distribution of weights. Student 325 explained:

> Like, I was thinking of more in terms of like in Physics, just because I'm taking a Physics class right now, and like, if you have to balance something out. And, you put, like you, it's balancing in the middle. Like, you can do that if they are evenly distributed, like you can just balance in the middle. But if it's something like weight distributed and all in weird positions, you kinda have to move the average. So the average is going to change for the smaller one because there's less, like, points, and it's not balanced.

Student 325 also explained, “But, if they’re evenly distributed, like, there’s a pattern.”

Student 817 mentioned thinking about sample size and aspects of distribution in the hospital context. He thought that more births meant there was less chance for births to be biased to 60%. Student 817 explained:

> So then more sample size, I guess, it would be more to the middle, better than 60%. …I guess, in the course there are a lot of problems that dealt with large sample size. That tended to be more of an even on the percentage of the probability, and the smaller ones are, like, more like of, like, random.

Student 817 also said, “Variability is, like, varying of the mean.” Furthermore, Student 817 explained thinking about sample size in conjunction with a bell-shaped curve and expectation: “If you’re guessing the probability of being each gender is 50%, I would guess that expectation is, like, the half of the births. And if you have a large sample size it would be more of a group and that expectation.”

Distributional aspects, sample size, and stochastic reasoning in the theater context. Student 313 only indicated stochastic reasoning in the theater context and did not mention thinking about aspects of distribution in relation to sample size in the theater context. In the hospital context, Student 313 thought about sample size with regards to changes in the ratio of boys to girls:

> If an extra boy was born as opposed to a girl, that would create a bigger difference in the ratio boys to girls, as opposed to if one more boy was born than one more girl in the 45 children born in hospital one. So, when I was saying this is because the smaller hospital has a smaller sample size and they created a probability of error if one [more boy], there would be a greater change from the expected value at hospital two as opposed to hospital one, a greater range if just one child, if just one boy or girl was born more than what is expected, the 50%. …I’m just using the expected value to calculate how many boys and girls there will be and it just gives me a basis for where to start off.

In the hospital context, Student 313 explained the importance of thinking about sample size with regards to a “rate” for each hospital, but not in relation to a distribution of outcomes. Student 313 explained how he thought about "rate" in this problem context: "…We're taking the sample of two hospitals. So, I definitely see the sampling, the sampling has, gave us a two rates. And it all comes into play. It's hard for me to explain it." In addition to sampling rates, Student 313 also indicated thinking about expected value and error: "I was considering error as the fact that it wasn't exactly the expected value." Student 313 did not talk about error in relation to a distribution of outcomes, but indicated that he was more used to "set values" and felt uncomfortable dealing with error:

> Everything in math and, not statistics, but mostly math, is concrete. And the expected value, I feel like I'm always looking for a specific value, not the fact that there is, the fact that there's a probability for error is kinda hard for me to contemplate. I'm more interested in a set value as opposed to saying there's a probability that this is the approximate the error, or this is the range, or the confidence interval, or different things like that. I'm more used to set values. So, I guess, through other calculus classes I've taken before and stuff, we've always had, like, a set value, not so much an interval where this is where the values can be.

Student 313 acknowledged that considering error and thinking about intervals was new to him. He perceived that learning statistics was very different from mathematics. Student 313 said in other classes, "The vast majority of the problems we're solving for specific values, as opposed to finding a range of values where the answer could be." Student 313 also said that he was used to "getting concrete answers," but in this class "the approach was different." In the hospital context, he struggled with coordinated notions of error, variability, and expected value: "Error to me would be, not, I guess, it's kinda confusing. …When I said error I meant it in respect to expected values. The error was anything that differs from the expected value, would be considered the error." Student 313 said he was thinking about variability and probability of error as being the same thing in the hospital context:

> I guess, the way I was using it in this context, it was probably the same. We usually think about variability as where there is no concrete answer. In this problem, where we expect, we have an expected value. We expect half of the children born will be boys and half will be girls. That would be anything that would differ from that I would, I said to myself was an error. But, it's not an exact error. But, anything without a concrete answer probably would be variability. That's just how I established it in my mind.

Although he struggled with coordinating notions related to distribution in the hospital context, Student 313 indicated thinking about randomness and aspects related to a distribution of outcomes in the theater problem context: "So, the randomness allows me to say that there will be, the results can differ depending on the situation. And we will, even though we can calculate expected value we'll never know exactly what that, we'll never, we can never predict exactly how many people are going come." Student 313 explained that he was thinking of distribution, and "…looking at it from the same [bell-shaped] curve that I'm used to. …I know that there are some problems where this would not apply. But as with regards to a problem like this, where we're considering random, random probability and variability, this is what I would first think about in my mind." Student 313 also indicated that he applied what he learned in the stochastic labs to his approach to the theater problem.

Stabilization of frequencies. One student who held a situational understanding also indicated thinking about stabilization of frequencies of outcomes and the law of large numbers. This student only indicated stochastic reasoning in the hospital context. The other student, who held a situational understanding and only indicated stochastic reasoning in the hospital context, merely indicated a vague recollection of the law of large numbers. The remaining student who held a situational understanding only indicated stochastic reasoning in the theater context. This student stated that he did not remember the law of large numbers, but indicated thinking about experiences with stochastic labs that involved simulations with stabilization of frequencies of outcomes. None of these students indicated thinking about stabilization of frequencies for the problem context in which they did not indicate stochastic reasoning.

Stabilization of frequencies and stochastic reasoning in the hospital context. Student 325 held a situational understanding and only indicated stochastic reasoning in the hospital context. She indicated thinking about a mathematical definition of the law of large numbers and coordinated that with thinking about long run stabilization of frequencies. First, Student 325 explained her understanding of the law of large numbers based on what she learned in class:

> The law of large numbers is that, I know like we said in class that, like, *n* has to be greater than 30. Like, and that it's considered a large sample size. So, the more, the greater *n*, the more likely that, like, the larger sample, or like a smaller sample of the larger sample, like, their averages are gonna be equal. And when I think about that, like, I think about if you take the ratio of, like, the sample average, and like the smaller sample that really, if you take that ratio it equals one, or it gets closer and closer. Like, I think in class we said as the, like, limit approaches infinity it goes to zero.

Student 325 indicated thinking about large samples and ratios in relation to a mathematical definition of the law of large numbers. Student 325 also wrote about the law of large numbers in her answer to the hospital problem on the conceptual quiz: "It is stated in the law of large numbers that with a sample size greater than 30 the sample average will more likely be the town average than a sample with less than 30." When Student 325 explained her thinking about what she had written, she gave an example and expressed thinking about stabilization of frequencies of outcomes in the context of repeated trials in chemistry experiments:

> Um, like, intuitively I, like, it makes sense if you have more and it's larger, more trials or whatever. It's, um, then you're going to be closer to, like, what you want, like you want to. With chemistry, we do more trials to be at the point where we want. That's what I think of the average as being, the point that you want it to be at.

Student 325 also indicated thinking that more trials result in a pattern: "So, with more *n*, and, like with a greater *n*, then you have it's more evenly distributed. …It's important to think about distribution just because it's where, the points lie. Like where, if they're, you know, if they're unevenly distributed, they're just to me, like, all over the place. But, if they're evenly distributed, like, there's a pattern."

Student 817 held a situational understanding and only indicated stochastic reasoning in the hospital context. When first asked about the law of large numbers during the interview, Student 817 said, "That wasn't from the course, was that? …I don't remember that." Later in the interview, Student 817 was asked if he remembered the meaning of the law of large numbers. At this point, Student 817 indicated a vague recollection of the law of large numbers, but did not appear to be confident about his thinking. Student 817 said, "Maybe, I think it was, if you have large numbers, the probability is expected by. If you draw a graph of, for instance, it's going to peak at an expected value." On his answer to the hospital problem on the conceptual quiz, Student 817 wrote: "Since the smaller hospital has much fewer deliveries, it is more susceptible to extremes or changes in the mean." Student 817 explained that he was thinking about changes in outcomes of numbers of boys: "I guess what I was going with, going for was that it was more of, that there's a lot of numbers. Less, a lot fewer numbers than, uh, required. It's going to be more of, uh, more susceptible to changes."

Stabilization of frequencies and stochastic reasoning in the theater context. Student 313 held a situational understanding and only indicated stochastic reasoning in the theater context. When asked about the law of large numbers, Student 313 said, "I can't say I remember it too well. I remember addressing the topic." Then Student 313 said, "I don't think I'll be able to answer your question." However, Student 313 also expressed thinking his work on the stochastic lab assignments influenced his thinking about the theater problem. The labs that he mentioned involved running simulations and observing stabilization of relative frequencies of outcomes. Student 313 explained:

> I know for problem number two [the theater problem]. I definitely, it was, I can think of the first or second lab where we were using coin flips and that whole lab. Just there was one part where we had to calculate the expected. How many, one of the questions was, 'How many people would we, how many?' I guess, I think, if I can remember back, was, like, how many coin flips would you expect to be either heads or tails. And that's how I approached this problem.

Although Student 313 said he did not remember what he learned about the law of large numbers, it appeared that working through stochastic lab assignments which were aimed at developing a conceptual understanding of the law of large numbers influenced his thinking about the theater problem.

Underlying distribution model. Student who held a situational conception thought about an underlying distribution model. In the context for which they held a stochastic conception, these students related thinking about sampling and variability to thinking about an underlying distribution. All three students mentioned thinking about a normal distribution model.

Underlying distribution model and stochastic reasoning in the hospital context. Student 325 indicated thinking about variability when she thinks about distribution: "Distribution, I think of just how the data is spread. And so, like I said before, normal distribution makes sense to me." Student 325 also connected thinking about sampling and variability to distribution: "I just keep thinking normal distribution. They should be, like, pretty, like, evenly distributed. And that, I think, like, that if you have a larger sample, again then, it's gonna, its all, there is gonna be way closer to the, like, some sample means."

Student 325 said work on the lab assignments impacted her thinking about the hospital problem. In particular, she mentioned thinking about the lab that involved samples of pennies and sampling distributions. Student 325 explained:

> I think it was the pennies, and it was. I thought about that in this [hospital] problem. …But, like, it made sense to me that, like, with the larger, um, like, you're gonna have, like, the mean is going to be closer to average. …It was the one, um, with the years and the class averages and people. So, 'cause I remember that, I think my average was, I think the difference in years was 18.5, or something. And like the class average was, like, 18.3, or something. And I remember on the questions, like, I couldn't really fully explain why, like, that was true. But, like, it made sense.

Furthermore, Student 325 also indicated thinking about a distribution model for ages of pennies:

> With the ages of the pennies, I feel like it kinda had a trend. …[The] mean is, like, there's not that actually, had that, um, μ. But with the pennies, like, I think it was the ages of, like, decreasing or increasing. Like, it just kinda looked like it was, so that is what I get.

When asked how he thought about distribution in the hospital context, Student 817 said, "Um, just fuzzy on that I guess. It's a distribution, like, the number of, uh, number of, uh, thing. Like the, oh yeah, I don't remember that one." Although he had a difficult articulating his thinking about distribution, he drew a picture of a discrete distribution model. As he drew the picture, Student 817 explained, "Something like that.

And the numbers, this is like 3, 4, and 2, like that." Student 817 said that the numbers represented outcomes and the bars represented frequencies.

Underlying distribution model and stochastic reasoning in the theater context. As he explained his thinking in the theater context, Student 313 drew a bell-shaped curve and said, "Again I'm looking at it from a same curve that I'm used to." Student 313 explained what he drew, "And that would probably put 2.5 here, and distribution lies as the amount of people increase that he knew. The probability would decrease as the amount of people decrease to not knowing anyone. The probability will also decrease." Student 313 also explained that he was thinking about variability and sampling in relation to the distribution model he had drawn:

> Distribution, I consider the fact that there, that again with randomness actually works. Exactly, with randomness and probability that there's a wide range of samples, and that they can vary. So, I'm using probability distribution. I use the two graphically. I use the graphic distribution to display how the probability will occur in that sample.

Unusual. In the theater problem context, students were asked to consider whether or not it would be unusual that at least two people Anthony knows attend the movie. Students who held a situational conception revealed thinking about "unusual" in two ways. The two students who only indicated stochastic reasoning in the hospital context indicated thinking about "unusual" as meaning having a low probability. The student who only indicated stochastic reasoning in the theater context indicated thinking about "unusual" as deviation from the expected value. None of the students who are characterized as having a situational conception quantified "unusual" as deviation from expectation in terms of a distribution model.

Unusual and stochastic reasoning in the hospital context. Student 325 and Student 817 both indicated thinking that "unusual" meant having a low probability in the theater context. Student 325 said, "I guess it means the probability of that scenario happening is small." Student 817 said, "Unusual, I guess, having a really low probability." Both of these students approached the theater problem using percentages. When asked how he conceived of probability in the theater problem, Student 817 replied, "I guess, having, uh, that he's going to know at least 1% then out of 300 people. There's going to be a person he knows." Student 817 said in this situation an example of "unusual" would be "having no person that's he is going to recognize." Although Student 325 said, "I struggled with this problem," she also indicated considering aspects of randomness and random samples in the theater problem context. Student 325 explained:

> I remember when I was looking at this problem, like, I kinda didn't know where to start. So, I just started trying to figure out, like, the percent of the town that he knows by name. So that's why I said he knows about 10% of the town. And, since the theater can only hold less of the town, percentage of the town that he knows, I was just thinking that, I don't even think I actually answered the question, but right now I would say that I find it unusual that he would know two people.

Unusual and stochastic reasoning in the theater context. Student 313 said, "Unusual would be not near the expected value. Being, having, uh, I'm using error, but I guess that's because I'm tired. Not being at the expected value, having a certain amount of variability." In the theater context, Student 313 explained that he thought of a bell-shaped distribution: "I know that there are some problems where this would not apply, but as with regards to a problem like this, where we're considering random, random probability and variability, this is what I would first think about in my mind." He indicated thinking about random variability and deviation from the expected value, but did not quantify "unusual" as deviation from expectation in terms of a distribution model.

Stochastic Conception

There were three interview participants who exhibited images of stochastic reasoning across all three categories in both problem contexts and interview probes. These students are described as holding a stochastic conception because they exhibited evidence of stochastic reasoning for both the hospital problem and the theater problem. In addition, during the interview they consistently exhibited evidence of stochastic reasoning when questioned about notions related to a stochastic conception. Two of the three students were in the stochastic reasoning lab group and one student was in the calculus review lab group. The characterizations that follow portray their stochastic perceptions in each of the three framework categories.

Image of repeatable process. Students who held a stochastic conception indicated conceiving of an underlying process that was repeated under essentially the same conditions. They indicated the process produced outcomes and described outcomes of the process. The students also talked about the repetitive process as being analogous to repeating an experiment. They envisioned that this process resulted in repeated sampling. Each of these students also connected their thinking about a repeated process to a model. Two of the students who held a stochastic conception were in the stochastic lab group, and both of these students spontaneously talked about how their experiences running simulations in *Fathom* influenced their thinking about the conceptual quiz problems. While the third student who held a stochastic conception did not have the experience with *Fathom* simulations, he referred to appropriate models and examples when explaining his reasoning.

Repetition under the same conditions. In both problem contexts, students who held a stochastic conception talked about a process that involved doing the same thing over and over again. For example, Student 113 conceived of the hospital problem as: “The question was asking me, ‘How many times will you have? How many days will there be? Which one will have more days?’ I was quickly trying to see, ‘Okay, what do you expect each day the boys will be?’” Student 214 explained that the process in the hospital problem context was analogous to repeatedly flipping coin:

> Here it’s binomial again. So, like the coin flipping you’re either going to have a success, which you can call a head or a boy here, or a failure, which would be a tail or a girl. At early stages it’s more likely to be different than the 50%..., but as time goes on, even if it’s not exactly 1 – 1, each side will have streaks where it will be more of one than the other. And then it’ll end up evening out.

Outcomes of the process. Students who held a stochastic conception described outcomes of the process. They thought about outcomes in the context of the hospital problem as births, and outcomes in the theater context as the number of people Anthony knew or did not know. For example, in the hospital problem context Student 329 explained, “In this one, if you, if one person has a boy, that doesn’t make the next person more likely to have a girl. In both situations it’s equally likely no matter what the previous outcome was.” Student 214 described the process for the hospital problem context as yielding a certain number of baby boys and girls: “Here I was actually visualizing in my head this situation kind of happening. …For 60% of the children to be boys in a hospital with 15 beds that would have to be 9 boys and 6 girls.” Student 113 said, “So I felt the smaller hospital, for example, randomly the first five births that they happen to be boys. Then, it's more likely that they'll have over 60% that day. Other hospital, that's delivering more babies. So the hospital that's delivering more will tend to have more of the one to one ratio." When asked to explain how he figured out his answer to the theater problem, Student 214 talked about his assumption that the theater would hold 250 people in a night and described thinking about "the number of people [Anthony] knew in the theater that night." He thought about outcomes as people Anthony knew or did not know. In the context of the theater problem Student 329 explained:

> Outcomes. That'll be there be zero to 250 people at the movie theater that Anthony knows by name. …Let's say the movie happened every night and 250 new people every single [night]. You know, there'd be quite possible that on one or two nights you'd see no one, but it'd be really close to 2.5. So, the fact that he'd seen 2, I thought, you know, that would be close to what would be expected.

Student 817 said, "I guess outcomes are that he's going to know 1 person, or 2, or 3, and so on."

Repeating an experiment. All three of these students mentioned they thought about repeating an experiment. When explaining his thinking during the interview, Student 113 continually gave examples of problem situations and experiments that involved repetition. He also talked about "conducting an experiment more and more." In the context of the hospital problem Student 113 mentioned, "The way I thought about it was if I flipped a coin 45 times. And, another experiment I flip a coin 15 times. Which one will more likely to get over 60% heads?" Student 113 also indicated that he was thinking about repetition of a binomial probability experiment: "If it is a binomial experiment with a probability of 0.5, if you ran a million trials you'd be more likely to have probability of 0.5." When talking about the theater problem, Student 113 envisioned a process where "you were having this movie over and over and over again and you were you know sampling with replacement with the population. …You completely conducted this experiment over and over and over." Student 329 conceived of the hospital problem as an experiment that "you keep doing every single day." He said, "If you keep doing the same sample over and over and over again and the same, like, experiment, then you would get some type of consistency in your numbers." Student 329 also talked about a repeated experiment when explaining his thinking about relative frequencies: "I think relative frequencies is, if you keep doing the same sample over and over and over again and the same like experiment, that you would get some type of consistency in your numbers, so just if you keep doing this every single day and going back." Student 214 also described his thinking about the hospital problem in terms of an experiment: "At early stages it's more likely to be different than the 50%.… But as time goes on, even if it's not exactly one-to-one, each side will have streaks where it will be more of one than the other. And, then it'll end up evening out." Student 214 went on to say, "Here I was actually visualizing in my head this situation happening."

Repeated sampling. Students who held a stochastic conception envisioned that repetition of the process resulted in repeated sampling. For example, Student 329 envisioned a process where the people who go to the movie theater were not "one certain group." He explained, "It's not the same people every time. It's a random sample of the population, just random. It's, like, not always the same people. It's always different people." Student 113 also thought about a sampling process in both problem contexts. When talking about the theater problem, Student 113 said:

> Let's say the movie happened every night and 250 new people every single, you know, there'd be quite possible that on one or two nights you'd see no one. … You know, you were having this movie over and over and over again and you were you know sampling with replacement with the population.

On his answer to the conceptual quiz Student 113 wrote, “Since there is randomness in the theater sampling, it is possible for him to know no people or more than three on a given night.” When describing his thinking about the hospital problem Student 113 mentioned sampling and said, “So in this situation it seems like each hospital is a sampling of the total births in the town.”

Model of repetition. All three students who are characterized as holding a stochastic conception connected their thinking about a repeated process to a model. In the context of the hospital problem Student 113 and 214 both mentioned thinking that the underlying process was similar to flipping a coin and could be modeled by a binomial probability experiment. Student 214 said:

> Here it’s binomial again. So, like the coin flipping, you’re either going to have a success, which you can call a head or a boy, here, or a failure, which would be a tail or a girl. At early stages it’s more likely to be different than the 50% because in order for it to be closer to 50% in the early stages you need 1, 1, 1, 1 failure-success. Otherwise you’re going to have very skewed answers. But, as time goes on even, if it’s not exactly 1-1, each side will have streaks where it will be more of one than the other. And then it’ll end up evening out.

Student 113 talked a great deal about repeated trials in a binomial experiment. He explained, “This is what I thought about it. [The] probability of having a boy and a girl is approximately modeled by a binomial. …. So, if it is a binomial experiment with a probability of 0.5, you know, if you ran a million trials, you’d be more likely to get closer to that probability of 0.5.” Student 329 did not specifically mention thinking about a binomial experiment, but talked about the hospital problem in relation to flipping a coin:

> It’s like if you flip a coin two times you could see heads both times, and that wouldn’t mean that a 100% of the time it would be heads. Whereas, if you do it like in the lab where we did it and we increased the sample size, each time it got closer to 50%. Which is what is actually is. So, this one I figured that the smaller

> one was like when we only flipped it one or two times whereas the 45 was closer to 100.

Simulations using Fathom. Student 329 and Student 113 were in the stochastic lab group and during the interview, both students spontaneously mentioned thinking about the *Fathom* simulations they had run during the semester when they explained their thinking. Both Student 329 and Student 113 coordinated their thinking about a repeating a binomial experiment with what they had observed when conducting simulations in *Fathom*. For example, Student 113 explained:

> Well I know that it is something where you have a certain number of trials, a fixed probability. So, I was thinking similar to flipping a coin. So, the way I thought about it was, you know, if I flipped a coin 45 times, and in another experiment I flip a coin 15 times. Which one will more likely to get over 60% heads? So I was thinking very quickly you'll see that it would be easier through the things we were doing all semester running these experiments through *Fathom*. That's what I was doing. You'd see that you could have, if you only ran 15 trials, you might end up with a result that definitely far away from 50-50.

Images of conditions of the process. Students who held a stochastic conception in both problem contexts indicated thinking about sample size, repetition of a sampling process, random sampling, variability, and the notion of independence. In both problem contexts, these students recognized that repetition of the process would yield a collection of outcomes. Students who held a stochastic conception also indicated thinking that the outcomes must be independent. In addition, they recognized that there would be variability in the outcomes resulting from the process. These students described a sampling process where each selection was equally likely and said the process would produce a random sample that was unbiased and representative of the population. Furthermore, students who held a stochastic conception indicated thinking that variability in outcomes was related to the size of the sample.

Collection of variable outcomes. The three students who indicated a stochastic conception in both problem contexts recognized that repetition of the process would yield a collection of outcomes. In addition to describing outcomes of a repeated process, students who held a stochastic conception recognized that there will be variability in the outcomes resulting from the process. For example, Student 329 mentioned that there would be variability in the number of people Anthony would know at the theater, "I guess you would see more than 2 people more times than he didn't see 2 people. So he would see 2 people most of the time." Student 113 also indicated thinking about variability in outcomes in the context of the theater problem when he said, "Let's say the movie happened every night and 250 new people every single [night]. You know, there'd be quite possible that on one or two nights you'd see no one." In the context of the hospital problem, Student 329 wrote about variability in outcomes of numbers of boys born in the smaller hospital and the larger hospital, and related variability to sample size. During the interview Student 329 explained his thinking: "And, since there is only 15 births each day in the smaller hospital, I realized that that one would have more variance just because it's so much of a smaller sample. And the 45, meanwhile, would be closer to same thing day to day, where the 15 would vary a lot more."

Independent outcomes. Students who held a stochastic conception assumed that the outcomes of the process were independent. For example Student 113 said, "Okay, so I definitely assume each birth was independent from the one before. So just like flipping a coin, having a boy in the delivery previously, does affect or does not influence the chances of having a boy the next time." Student 329 explained, "Well that's just that the probability of having boys is independent of having a girl, so either way they're both

50% all of the time." Student 214 inferred that he was thinking the outcomes were independent because he talked about the problem as being a binomial experiment analogous to flipping a coin: "Here it's binomial again. So, like the coin flipping you're either going to have a success, which you can call a head or a boy, here, or a failure, which would be a tail or a girl." None of the students who held a stochastic conception used the word "independent" when describing their thinking in the context of the theater problem, however, they described notions of randomness that involved an assumption of independence. For example, when explaining why the notion of randomness was relevant to the theater problem Student 214 said, "I think it's important in that it helps us generate the assumptions. Like we said, what the type of movie is random and it doesn't affect who's coming."

Sampling process. All three students who held a stochastic conception talked a great deal about notions of randomness and sampling. They talked about a sampling process where each selection would be equally likely. For example, when talking about the theater problem, Student 113 said:

> If you have all the, the entire population of the town, he knows about 300. And so, if you pick a random person from the town, it's likely that he'll know 1 out of 100 people. ...Sorta means that for every, the way it's distributed is that, every hundred people you'll expect to get one.

Student 113 also explained:

> If you're thinking of a random sample and the population, …you'd want it to be completely, the chances of you choosing one person over the other completely, you know, just random. Can't use the same word. It's equally likely to choose anyone else in the population as you choosing that one person. If you're going to go with a random sample, you're going to want ideally your random sample, you know, you're just completely pulling numbers out of a hat. And, you'd have similar representation. You'd have the people in your sample come from similar backgrounds as the population.

Student 329 said that he thought about a sampling process that would result in a random sample: “I feel like it’s random, but, like, you’d all make sure you had different types of ethnicities and races and genders. And, like just like a random sample of the whole population. And like different age groups and stuff like that. You’d want to make sure to count for everyone.”

Representative sample. Students who held a stochastic conception assumed that the results of the sampling process would be a random sample that would be unbiased and representative of the population. Student 329 connected the notion of taking a random sample with the idea that the result would be representative of the population and said, “To take a random sample I’d feel like that’s like an accurate representation of the whole population.” Student 113 explained, “But I know that you know that you assume the sample is random. …Since the sampling is completely random, it’s not biased towards any group. Or, a completely random sample will be well suited to the population you’re working with.” Student 214 explained his assumption about sampling in this way:

> I think you assume a random sampling. I think you have to assume that its fair and it holds true to any assumptions that you’re given, like we said. What the type of movie is random and it doesn’t affect who’s coming. The people that he knows going to the movies might be random because if he really liked a ton of people who loved to go to movies they’d more likely be represented.

These students connected the notion of randomness to taking an unbiased representative sample: “For example, I think [randomness] means that a given population or a given probability isn’t biased by any other factors than what we’re given in the problem” (Student 214). When explain his thinking about the hospital problem Student 329 said:

> This one I assumed was a random sample because… I guess you can typically expect that would be the case in a hospital because they are not changing things.

> …In this one, if you, if one person has a boy that doesn't make the next person more likely to have a girl. In both situations it's equally likely no matter what the previous outcome was.

Student 329 also mentioned random sampling when explaining his thinking about how probability is related to randomness: "Well I think they go kinda hand in hand. Like probability doesn't occur if it's not a random sample. Or, it's not an accurate representation of the probability if it's not a random sample. So randomness is really important."

Variability and sample size. Students who held a stochastic conception indicated that sample size is related to variability. "So in this situation it seems like each hospital is a sampling of the total birth in the town. I think that the smaller sample will tend to (pause). The results of the smaller sample will not fit the population as well as the larger sample would" (Student 113). These students also recognized the implications of sample size in producing variability in outcomes of the process. For example Student 214 said, "To me that seems like that would make a lot of sense because the sample size is small and that just means that for a couple of children they ended up going one way rather than the other. But a hospital of 45 (pause), 27 to 18 I think it is. …So even though the sample size is proportionality larger I think it would be closer to even." Student 113 talked at length about the implications of sample size in the context of the hospital problem:

> I was drawing more really on the sample size. I feel like the relative sample size is between the two hospitals. … One birth, or one that resulted in more births than expected of boy would be, would influence the ratio of the smaller hospital no matter what. So I, I didn't get hung up on the assumption. I didn't say, you know, what if it's not really 50 percent because I could tell the problem wanted you to sorta assume it's 50-50 and think it through. You know, like, it seems really the influence here is on sample size rather than the actual. It's not focusing

> on process of the birth, or the probability of one or the other happening. It's more about the results through the sample size.

Underlying model and conditions of the process. Students who held a stochastic conception connected their thinking about conditions of the repeatable process to an underlying model. Thinking about a model was particularly apparent in the hospital problem context where all three students who held a stochastic conception mentioned thinking about flipping a coin. For example Student 113 indicated variability in the birth outcomes could be modeled by coin flips:

> If you were flipping a fair coin [and] if you flip it ten times, you'd say, 'Oh well most often you'll have five heads.' But you could go ahead and flip it just ten times and all of a sudden get 7 heads. All of a sudden 70% heads and you say, 'Whoa wait a second. That's not the probability of that happening.' And, that's why with the smaller sample space it's you have more. It's more likely that that'll happen. But, if you have a lot and lot of trials, it's more likely that that'll be evened out ratio wise.

Student 329 also indicated thinking about the model of flipping a coin:

> It's like if you flip a coin two times. You could see heads both times and that wouldn't mean that 100% of the time it would be heads. Whereas if you do it, like in the lab where we did it and we increased the sample size, each time it got closer to 50% which is what is actually is. So this one I figured that the smaller one was like when we only flipped it one or two times, whereas the 45 was closer to 100.

Student 214 was not in the stochastic lab group, but he also mentioned thinking about coin flips:

> And then I ended up coming up with my answer not seemingly through any math here, but through pre-conceived knowledge I guess that I had that in larger sample sizes, things tend to fall closer to average than you would expect. Like, if you flip a coin 100 times you're more likely to get closer to an even 50-50 than if you flipped a coin 4 times, you'd be more likely to have like a 3-1 or a 4-0. …Here it's binomial again. So, like the coin flipping, you're either going to have a success, which you can call a head or a boy, here, or a failure, which would be a tail or a girl.

Image of distribution. Students who held a stochastic conception indicated thinking about a distribution of outcomes. These students indicated their thinking about distribution by talking about expectation and how outcomes would deviate from what was expected. They indicated thinking about variability in the context of distribution and connected their thinking about expectation to variability in outcomes. These students recognized there would be a stabilization of frequencies over a large number of repetitions of the process and indicated thinking about the law of large numbers. Students who held a stochastic conception also indicated thinking about an underlying distribution model. In addition, they quantified unusual as deviation from expectation in relation to a distribution model.

Distribution of outcomes. Student who held a stochastic conception described notions of distribution when talking about outcomes of the repeated process. For example, Student 214 indicated thinking about a distribution of outcomes that could shift, depending on the assumptions. Student 214 explained:

> I think randomness is important because if there isn't randomness then you can't be sure of your distribution. Like, it could be heavily shifted one way or the other. Like in the example, if his friends really liked going to the movies, the entire curve would be shifted to the right, and you would expect him to see more people that he knew.

Student 329 mentioned thinking how the mean of the outcomes would vary with repetition of the process:

I think relative frequencies is, if you keep doing the same sample over and over and over again, and the same like experiment, that you would get some type of consistency in your numbers. So, just if you keep doing this every single day and going back how much the mean, the sample mean, would differ from the expected mean.

Student 113 described his thinking with regards to about how outcomes would vary by talking about outcomes of flipping a fair coin and relating that to what he observed in one of the *Fathom* simulations. Student 113 explained, “Ah, you know, like I, you know, I said earlier about the lab with flipping the coin and the percent. The deviation from what you’re expecting in the experiment, can become greater and lower, you know, with smaller samples.”

Expectation and variability. Students who held a stochastic conception talked about expectation and how outcomes would deviate from what was expected. When thinking about how to solve the problem, they connected notions of expectation and variability. For example Student 113 said, “In a 50:50 situation it’s giving you about the number of boys you would have at each hospital. And then I said, ‘That’s useless. I want to know what the deviation could look like, or what how often it could deviate from the expected’.” Thus, Student 113 seemed to understand the importance of considering how outcomes could vary from what was expected. When Student 214 explained his thinking about the hospital problem, he also connected the notion of expectation to variability in outcomes:

> …In larger sample sizes things tend to fall closer to average than you would expect. Like, if you flip a coin 100 times, you’re more likely to get closer to an even 50:50 than if you flipped a coin 4 times. You’d be more likely to have, , a 3:1 or a 4:0.

Student 329 talked about variability when explaining his thinking about expectation:

> The expected mean is what you expect, as it says, in the sample mean. Like, if you did one particular day, and you went to the hospital and there’s 45 babies born, the sample mean would be how many of them were actually boys. So, they’d probably differ from the expected value.

When he talked about expectation in the context of the theater problem, Student 113 mentioned variability in outcomes as he drew a diagram representing the distribution:

> I would say, like, if you’re going to the movie, he’s expected to see 2.5 people. ‘Cause this is the distribution of this sample. So I sorta thought about, you know,

> sorta being a normal distribution where it'd be for him to see 2 people. That's very close to what you'd expect. You really wouldn't expect him to see 10 or 15.

In the context of the theater problem, Student 214 also thought about how outcomes might vary in relation to what would be expected. First he talked about determining the expected ratio, then he mentioned that the expected value is only an estimate:

> So I used the ratio we were given to calculate expected ratio. Well I guess it'd be expected because he only estimated. …I think what I actually got here was the expected average or mean he would encounter per night. And then for the standard deviation I would use that with the average to figure out where his estimate of 2 lies on a normalized curve, if it was normalized.

Variability: An important aspect of distribution. These students indicated thinking about variability in the context of distribution. For example, when explaining his thinking about the theater problem, Student 329 said:

> The expected value that I got was 2.5. So you would expect, like you know, if it followed the normal distribution, you'd expect it to be 2.5 most of the time. And slightly vary from that a little bit. …So, like, sometimes it would be more than 2.5 and sometimes it would be less. But most of the time it would be about 2.5. It would average out to that.

In the context of the hospital problem Student 329 also thought about variability and a distribution of outcomes:

> So, for the smaller one, ur, the larger sample, there'd be less of a sampling distribution because and it's larger. Whereas the hospital that only has 15 births each day that it'd have much larger distribution. I guess more of a distribution would mean the mean fluctuates more. It would have a bigger standard deviation.

When explaining his thinking about distribution, Student 214 said, "So I would expect a lower variability would mean the numbers are more closely grouped together around the average."

Students in the stochastic lab group ran simulations and observed distributions of outcomes resulting from the simulations. Student 113 said that at first he thought about approaching the hospital problem by strictly looking at numbers and doing calculations, but he explained that his experiences doing the stochastic lab assignments prompted him to think about variability and consider aspects of distribution that were involved with solving the hospital problem:

> I guess the one thing that made me look at this, that sorta took me away from, you know, at first starting to do this, because I looked at the question and said, "I know, you should know this." And, I tried to go. "Okay, so let's try to calculate it out." And, I realized, you know, wait a second, you just have to stop and think. And, one thing I learned this semester was that in probability a lot of times things you that would expect, you know, things aren't usually like as they seem. …But I thought, you know, wait a second. This problem is trying to tell me something. It can't be that easy. So that's how I sorta thought through, uh, you know, like I, you know, I said earlier about the lab with flipping the coin and the percent. The deviation from what you're expecting in the experiment can become greater and lower, you know, with smaller samples.

Stabilization of frequencies. Students who held a stochastic conception recognized that frequencies of outcomes would stabilize over a large number of repetitions of the process. For example Student 214 said, "Like, if you flip a coin 100 times you're more likely to get closer to an even 50:50 than if you flipped a coin 4 times, you'd be more likely to have like a 3:1 or a 4:0." Student 214 contrasted what would happen in the long run with what would happen in a short run of repetitions:

> At early stages it's more likely to be different than the 50% because in order for it to be closer to 50% in the early stages you need 1, 1, 1, 1 failure-success. Otherwise you're going to have very skewed answers. But, as time goes on even, if it's not exactly 1-1, each side will have streaks where it will be more of one than the other. And, then it'll end up evening out.

Student 329 explained his thinking about stabilization of frequencies this way: "I think relative frequencies is if you keep doing the same sample over and over and over again and the same, like, experiment, that you would get some type of consistency in your numbers. So just if you keep doing this every single day." Student 113 talked about stabilization of frequencies when explaining his thinking about the hospital problem: "I thought about it was, well, you have a larger hospital, so the experiment's being conducted more and more. I felt the larger one would have less deviation from that 50% mark, because you have an anomaly, like, same thing like flipping a coin, you flip 4 heads in a row or something. It won't be that much affected if you're flipping a coin 100 times." Student 113 also connected his thinking about stabilization of frequencies to what he observed simulating experiments in *Fathom*:

> If you, once your sample gets bigger, larger and larger and larger, it very, very, very seldom that it'll deviate from, you know, from the expected, or what's expected in the population. You know, like we saw in the *Fathom* experiment. You know, you run the trial in *Fathom* thousands and thousands of times and the line is really, really, really close to what you expect to be, 0.5. So, I sorta felt like 45 is like getting closer to law of large numbers than 15 is. So that was another thing I was thinking about.

All three of these students also indicated thinking about the law of large numbers. Student 113 appeared to have the most robust recollection of the law of large numbers and said:

> The law of large numbers, I mean, this is why insurance really works. …That means that the law of large numbers, if your sample is very, very large, then your result is going to be very, very, very tight with the mean, what your expected [mean] is. For example, the law of large numbers, we've looked in regards in economics with people, with insurance companies pricing, pricing how much premiums going to be.

Student 329 indicated a bit more hesitation when asked about the law of large numbers. He said, "I think the only thing I really kinda remember, I don't even know if it's right or not, is that the law of large numbers means that certain probabilities apply once you have, like, a certain number of the population in the sample." (Student 329). Student 214 made an assumption that indicated an understanding of the law of large numbers: "Well, the assumption that I was just talking about, that it gets closer as you get to larger samples sizes." However, when asked about the law of large numbers in the early part of the interview, Student 214 said, "I didn't see anything about that." But then later in the interview, when presented a problem designed to probe his understanding of sampling distribution, Student 214 indicated thinking about the law of large numbers: "But the sample size being 4, uh, with the law of large numbers, …the more likely you would stray from the average here, or rather the mean of the population."

Distribution variance, shape, and sample size. Students who held a stochastic conception described how sample size affected variance and shape of the distribution of outcomes. For example Student 214 said, "…In larger sample sizes things tend to fall closer to average than you would expect. Like, if you flip a coin 100 times, you're more likely to get closer to an even 50:50 than if you flipped a coin 4 times." Student 214 indicated that he when he talked about sample sizes he was thinking about "number of births" and said:

> I think a given sampling should follow a general distribution. Or, I feel the larger sampling you have the more likely it would hold true to that. But you could also have samples that are not representative of the population and could fall into outliers. … In the smaller sample, I realized that that one would have more variance, just because it's so much of a smaller sample. And the 45, meanwhile, would be closer to same thing day to day, where the 15 would vary a lot more.

Student 113 also explained the importance of sample size when thinking about a distribution of outcomes in the context of the hospital:

> I'm thinking sample size is important because you know, just like I said, when you're conducting an experiment more and more you're going to get closer to what would be expected from, you know. So, if it is a binomial experiment with a probability of 0.5, you know if you ran a million trials, you'd be more likely to get closer to that probability of 0.5. … So I felt that with the greater sampling in the larger hospital you'd be more representative. …You wouldn't be skewed towards boys or towards girls. So, I figured it would be more likely to be less days where more than 60% of boys were born.

Student 329 also talked about the importance of sample size and said, “I just remembered like wondering about the sample sizes and how smaller ones vary a lot more than the larger ones.” In the context of the hospital problem, Student 329 thought about the differences between the two hospitals in relation to sampling and said, “I realized that [the smaller hospital] would have more variance just because it’s so much of a smaller sample. And the 45, meanwhile, would be closer to same thing day to day, where the 15 would vary a lot more.” Student 329 also explained, “…the larger sample, there’d be less of a sampling distribution because it’s larger. Whereas the hospital that only has 15 births each day, that, it’d have much larger distribution. …I guess more of a distribution would mean the mean fluctuates more. It would have a bigger standard deviation.”

Underlying distribution model. Students who held a stochastic conception also indicated thinking about an underlying distribution model. For example, Student 329 thought about the problem contexts as experiments and outcomes in relation to a distribution model: “I guess just when you do an experiment, [you think], ‘How closely does it follow the normal distribution?’ or ‘Is it skewed left or right?’” Student 329 explained that he thought about a normal distribution model in context of the hospital problem:

> Just that the 45 babies would more closely followed the normal distribution, just because of what we learned in the labs. Like where we had the one where, the one I talked about earlier, where you were flipping the coin and heads and tails are going to even out more closely as we increase the samples. So, just the normal distribution would be that the mean is actually 50%, so there’d be 22½ [boy] babies in the one hospital and 7½ in the other.

Student 113 explained that he thought about a binomial distribution model in the context of the hospital problem: “Just ‘cause you have a fixed number of trials. They’re telling you how many. …You assume that the probability is 50:50.”

Student 329 also thought about the normal distribution model in the context of the theater problem and explained:

> If you say that the mean of the whole population is 2.5, then that's normally distributed. Assuming that that's the correct answer for the probability, or the average number of people he sees each time when he's there. …if it's a normal population, normal distribution where you kinda expect the bell curve type thing or one side is favored higher than the other side.

As Student 113 explained his thinking about the theater problem, he drew a diagram of a distribution (Figure 4.x) and said, "So I sorta thought about, you know, sorta being a normal distribution. Where it'd be for him to see 2 people, that's very close to what you'd expect. You really wouldn't expect him to see 10 or 15. And that's sorta how I was thinking about that, you know, a lot of samples were, they were normally distributed." Then, pointing to the center of the distribution he had drawn, Student 113 explained:

> So, you assume it's normal in that you know that's the greatest probability right in here. …Then you have, you know, you're expected result is, you know, the peak of the curve. And since it's inherently random, it shouldn't be skewed towards any of the sides. So the bell curve centers around what you expect and you have the probabilities on either side of something else.

Quantification of unusualness. Students who held a stochastic conception quantified "unusual" as deviation from expectation in relation to a distribution model. In the context of the theater problem, they thought about "unusual" in relation to the expected value or mean of a distribution model and conceived "unusual" as meaning not near the expected value. For example Student 113 said, "Unusual would mean a deviation from what is really expected, a result that seemingly, I guess, far away from what you would expect to see." He also explained, "It's random…, you can never tell with certainly what's going to occur. You can sorta understand with confidence or over a lot of trials what you would be expected or what would be unusual." Student 214 explained his thinking about outcomes in relation to a distribution model:

> We have an estimated value of 2, which just given that it's approximately here. This part of the curve would be the times, or the part of the distribution where it actually occurred. … Then I used the normalized curve to show that he thought that it was unusual that he would see at least 2 people.

Student 214 also explained that he thought about unusual in comparison to what would be expected, "But each of those correspond to different values that we find in this distribution, I believe, because the number of people he sees on average compared to what you would expect dictates what is usual or unusual." Similarly, Student 329 thought "unusual" would mean an outcome not near the expected value:

The question was asking if it was unusual that he knew at least 2 people, but I found that on the average he would know 2.5 people. So, a little bit more than 2. Therefore, I found that it wasn't unusual for him to find, to see at least two people that he knew. In this situation I'm sorta saying that it's less than 50%, but I feel like unusual is probably not common. So it would probably be a much smaller percent. …I guess you would see more than 2 people more times than he didn't see 2 people.

Summary of Stochastic Conceptions

The analysis presented in Chapter 5 addressed four research sub-questions related to students' conceptions of probability distribution. This section summarizes findings of the qualitative analysis by research question. The first research sub-question addresses evidence of students' overall stochastic conception as evidenced by reasoning across three hierarchical framework categories: image of a repeatable process, image of specification of conditions, and image of distribution. The three components of research sub-question 1 address students' conceptions evidenced in each of these categories. A summary of the analysis of students' conceptions and characterization of their stochastic conceptions is followed by summaries of analyses of students' conceptions evidenced for each of the three framework categories.

Research Sub-question 1:

What is the nature of students' reasoning when confronted with a probability situation? Analysis of the students' interview responses along with their written work on the conceptual quiz revealed evidence of stochastic reasoning across all three framework categories: image of a repeatable process, image of specification of conditions, and image of distribution. Evidence garnered from qualitative interview data confirmed that these categories are hierarchical. Individuals who presented an image of distribution of outcomes for a given problem, the highest level of conceptualization, also presented an image of specification of conditions and an image of a repeatable process for that problem. Similarly, individuals who presented an image of specification of conditions for a given problem also presented an image of a repeatable process for that problem. Some individuals only presented an image of a repeatable process for a given problem context, while others presented no images indicative of a stochastic conception.

Table 5.1 provides a summary of stochastic images indicated by students during the end-of-course interviews. The table is partitioned to represent evidence of stochastic images presented by each student for each of the two problem context on the conceptual assessment. Each row of the table denotes an interview student and columns specify framework category images. If a student presented evidence of a stochastic image for particular framework category, the table displays an "X" in that column. Table 5.1 shows that 100% of the interview participants indicated thinking along the proposed hypothetical learning trajectory. This finding provides evidence which confirms the hierarchical nature of the framework categories and substantiates this hypothetical learning trajectory for development of stochastic understanding of probability distribution.

Table 5.1

Summary of Interview Participants' Stochastic Images

Hospital Problem	Image of Repeatable Process	Specification of Conditions	Image of Distribution of Outcomes
Student Number			
113	X	X	X
214	X	X	X
313	X	-	-
325	X	X	X
329	X	X	X
428	-	-	-
518	-	-	-
711	X	-	-
715	X	-	-
811	-	-	-
817	X	X	X
1016	X	-	-

Theater Problem	Image of Repeatable Process	Specification of Conditions	Image of Distribution of Outcomes
Student Number			
113	X	X	X
214	X	X	X
313	X	X	X
325	X	-	-
329	X	X	X
428	-	-	-
518	-	-	-
711	X	-	-
715	X	-	-
811	-	-	-
817	X	X	-
1016	-	-	-

The analysis of the students' interview responses along with their written work on the conceptual quiz revealed evidence of stochastic reasoning across all three framework categories and the hierarchical nature of the categories: image of a repeatable process, image of specification of conditions, and image of distribution. Students who were characterized a holding a stochastic conception for a given problem context, indicated stochastic reasoning across all three framework category images for the problem context. Table 5.2 summarizes evidence of stochastic reasoning presented by interview participants across the three image categories of the conceptual framework for each problem context.

Table 5.2

Evidence of Stochastic Reasoning across Categories of the Conceptual Framework

Framework Categories	Student Numbers
Hospital Problem	
Image of distribution, specification of conditions, image of a repeatable process	113, 214, 325, 329, 817
Specification of conditions, image of a repeatable process	113, 214, 325, 329, 817
Image of a repeatable process	113, 214, 325, 329, 817, 313, 711, 715, 1016
Theater Problem	
Image of distribution, specification of conditions, image of a repeatable process	113, 214, 313, 329
Specification of conditions, image of a repeatable process	113, 214, 313, 329, 817
Image of a repeatable process	113, 214, 313, 329, 817, 325, 711, 715

Drawing on this analysis, a hierarchical numerical measure was defined to yield scores indicative of levels of stochastic reasoning. This measure was applied to the qualitative data. A numerical measure was defined to yield scores indicative of the levels of stochastic reasoning demonstrated by interview participants. A score of 0, 1, 2, or 3 was assigned for each stochastic image within a problem context. Individuals who presented evidence of an image of distribution, specification of conditions, and image of a repeatable process were assigned a score of 3 for the problem context in which these images were reported. Individuals who specified conditions and reported an image of a repeatable process were assigned a score of 2 for the problem context in which these images were reported. Individuals who reported an image of a repeatable process were assigned a score of 1 for the problem context in which these images were reported. Individuals who did not presented any images related to a stochastic conception were assigned a score of 0 for that problem context. There were two problem contexts, so individuals could receive a score from 0 to 6, inclusive.

Three students consistently exhibited evidence of stochastic reasoning across all three categories of the conceptual framework for both problem contexts and interview probes. These three students are characterized as holding a stochastic conception. The students who held a stochastic conception received a score of 3 for each of the problem contexts, which results in a combined score of 6. A combined stochastic score of 6 indicated that a student held an image of a repeatable process, an image of specification of conditions, and an image of distribution for both problems on the conceptual assessment.

Three students exhibited stochastic reasoning across three categories of the framework for only one of the two problem contexts and exhibited stochastic reasoning in only one or two framework categories for the other problem context. Students assigned a score of 4 or 5 and are characterized as holding a situational conception. The students who held a situational conception received a score of 3 for exactly one of the problem contexts and a score of 1 or 2 for the other problem context, which resulted in a combined score of 4 or 5. A combined stochastic score of 4 or 5 indicated that a student held an image of a repeatable process, an image of specification of conditions, and an image of distribution for exactly one of the two problems on the conceptual assessment and reported a nonstochastic conception for the other problem context.

None of the students who were interviewed exhibited reasoning across only two categories of the framework for both problem contexts. Furthermore, none of the students exhibited stochastic reasoning across all three categories of the framework for exactly one of the two problem contexts and simultaneously exhibited no images related to a stochastic conception in the other problem context. Therefore, no student received a score of 3.

Three students reported stochastic reasoning in one category of the framework for either one or both problem contexts. Students assigned a score of 1 or 2 and are characterized as holding a nonstochastic conception. The students who held a nonstochastic conception received a score of 1 for at least one of the problem contexts and a score of 0 or 1 for the other problem context, which resulted in a combined stochastic score of 1 or 2. A combined stochastic score of 1 or 2 indicated that a student only held an image of a repeatable process for either one or both problem contexts on the conceptual assessment.

Three students reported no images related to a stochastic conception for either problem context. These students received a score of 0. A combined stochastic score of 0 indicated that a student did not report an image related to a stochastic conception for either problem context. Table 5.3 summarizes interview participants' combined stochastic scores for the interview over four mutually exclusive categories describing students' stochastic conceptions.

Table 5.3

Distribution of Student Identification Numbers over Categories of Stochastic Conceptions

	Stochastic Score (6)	Situational Score (4 or 5)	Nonstochastic Score (1 or 2)	No image Score (0)
	113	313	711	428
Student numbers	214	325	715	518
	329	817	1016	811

Stochastic image. Interview students who indicated a stochastic conception for both problem contexts consistently described their thinking in relation to a model of repetition. They coordinated thinking about experimental probability with what might happen and with their thinking about theoretical probability. These students used examples to illustrate their thinking and described their thinking in terms of models that represented the problem situation. When explaining their theoretical understanding they often used an example of an empirical situation. Furthermore, students who were characterized as holding a stochastic conception appeared confident when answering interview probes. Their answers to the interview probes were relatively concise and to the point. This manner of response contrasted with students who were characterized as holding a situational conception.

Interview students who held a situational conception displayed conflated thinking at times during the interview, as evidenced by lengthy explanations and some contradictions. In one of the two problem contexts, these students confidently coordinated thinking about experimental probability with theoretical probability in one of the problem contexts, but struggled to coordinate these notions in the other problem context.

These qualitative findings elaborate quantitative findings in this study. Students who were characterized has having a stochastic image are students who exhibited a stochastic conception for one or both problem contexts on the conceptual assessment. Students who held a stochastic image indicted coordination of a data-centric perspective with a modeling perspective of distribution and coordinated thinking about experimental and theoretical probability in one or both problem contexts on the conceptual assessment. Completion of the stochastic reasoning lab assignments had a statistically significant impact on students' stochastic conceptions. The mean stochastic image for students in the stochastic reasoning lab group who completed all of the lab assignments was significantly higher than the mean stochastic image for all students in the calculus review lab group and was significantly higher than the mean stochastic image for students in the stochastic reasoning lab group who did not complete all of the lab assignments. These findings indicate that completion of the stochastic reasoning lab assignments facilitated movement along a hypothetical learning trajectory towards a stochastic understanding of probability distribution. Students who were characterized as holding a stochastic image indicated a principled understanding of probability distribution which included integration of stochastic understandings with instrumental understandings of probability.

Nonstochastic image. Interview students who were characterized as having either a nonstochastic conception or no image related to a stochastic conception generally approached the problems by looking for calculated answers. They focused on using formulas and doing mathematical calculations and seemed to resist interview questions aimed at probing stochastic reasoning. In general, the understanding of these students did not appear conflated by attempts to resolve notions of experimental probability with theoretical probability. They focused on determining the theoretical probability and did not perceive a conflict because they were confident they had used the correct numbers. Their approach to solving the problems was identifying the pertinent information, deciding what formula to use, and then completing a mathematical calculation to obtain the answer.

These qualitative findings elaborate quantitative findings in this study. Students who held a nonstochastic conception or indicated evidence of holding no stochastic images were characterized as having a nonstochastic image. The majority of the participants in this study held a nonstochastic image of probability distribution. Although evidence supports the claim that the stochastic reasoning lab assignments facilitated movement along a hypothetical learning trajectory towards a stochastic understanding of probability distribution, most students in this study did not indicate stochastic conceptions. Qualitative evidence indicates that these students did not integrate thinking about experimental probability and theoretical probability, but their thinking about probability focused on an algorithmic approach and the use of formulas and calculations to determine a correct numerical answer. Thus, students who were characterized as holding a nonstochastic image indicated solely instrumental understanding of probability distribution.

Research Sub-question Component 1(a) Image of a Repeatable Process

How do students characterize a probability situation in terms of an image of a repeatable process? Students who held a stochastic conception consistently indicated conceiving of an underlying repeatable process. They indicated thinking about aspects of a repeatable process in both problem contexts and interview probes. Furthermore, they envisioned that the process produced outcomes, and they talked about repetition when describing outcomes of the process. Students who held a stochastic conception related thinking about the process to repeating an experiment. These students also talked about repeated sampling when describing their thinking about the process. In contrast to other students who were interviewed, all three students who are characterized as holding a stochastic conception spontaneously mentioned thinking about a model when describing their reasoning. Two students who are characterized as holding a stochastic conception mentioned thinking about running simulations in *Fathom*, and the third reported thinking about coin flips.

The interview participants who held a situational conception also indicated thinking about a repeatable process for both problem contexts and indicated this process would be repeated under essentially the same conditions. Furthermore, these students recognized that the process would yield outcomes, and they described the outcomes of the process. Students who held a stochastic conception indicated thinking about a repeated experiment and repeated sampling in both problem contexts, whereas students who held a situational conception only indicated thinking about a repeated experiment and repeated sampling in the problem context for which they exhibited stochastic reasoning. Students who held a nonstochastic conception gave some indication that they thought about a repeatable process and outcomes of the process, but they did not indicate thinking about a repeated experiment. Students who held no image of a stochastic conception did not indicate thinking about a repeatable process and did not indicate thinking about outcomes of a process. Students who held a nonstochastic conception, along with those who held no stochastic image, strictly conceived of probability as a fixed ratio or percent. They also tended to focus on numerical calculations. Table 5.4 summarizes images presented by interview participants that related to a repeatable process. The summary of stochastic images reported for each stochastic conception category includes evidence indicated by all students who are characterized as holding that particular conception. *May* is used to indicate evidence reported by some explanations, but this evidence was not required criteria.

Table 5.4

Summary of Images for Repeatable Process

Conception	Image for Repeatable Process
Stochastic	*Indicates understanding that the process is repeated under essentially the same conditions*
	Indicates the repeatable process yields outcomes and *describes outcomes of the process*
	Indicates thinking about a repeated experiment.
	Indicates that repetition of process results in repeated sampling
	Connects thinking about the process to a model
	May connect thinking about the process to running a simulation
Situational	*Indicates understanding that the process is repeated under essentially the same conditions*
	Indicates the repeatable process yields outcomes and *describes outcomes of the process*
	May indicate thinking about a repeated experiment
	May indicate that repetition of process results in repeated sampling
	May connect thinking about the process to a model
	May connect thinking about the process to running a simulation
Nonstochastic	*May indicate understanding that the process is repeated under essentially the same conditions*
	May recognize that repetition yields outcomes and *may describe outcomes of the process*
	No indication of thinking about a repeated experiment
	May recognize that the process involves repeated sampling
	Indicates thinking about probability as a ratio
	Does not connect thinking about the process to a model
No Image	*Does not indicate thinking about a repeatable process*
	Does not consider outcomes of a process
	Indicates thinking about rates and/or proportions, not a repeated experiment
	Indicates thinking about probability as a fixed ratio or percent
	Focuses on mathematical calculations
	Does not indicate thinking about a model

Note. Italicized images were only evidenced as a result of the interview. Images not italicized were evidenced as a result of the interview and via student's individual written answers for the conceptual quiz.

Research Sub-question Component 1(b) Specification of Conditions

How do students characterize a probability situation in terms of specification of conditions of a repeatable process? Students who are characterized as holding a stochastic conception specified conditions related to the repeatable process they envisioned. These students described thinking about a sampling process where each selection was equally likely. They also indicated thinking that the results of the sampling process would yield a collection of variable, independent outcomes. Students who held a stochastic conception were attuned to the notion of variability and indicated thinking about variability in outcomes in relation to sample size. Students who are characterized as holding a situational conception indicated similar thinking, but only in the problem context for which they exhibited stochastic reasoning. Only the students who held a stochastic conception indicated thinking the sampling process would produce samples representative of the population.

Students who were characterized as holding a nonstochastic conception as well as those who held no images related to a stochastic conception did not indicate thinking about a collection of outcomes. Furthermore, these students did not report thinking about a sampling process, nor did they mention thinking about variability. At times, some of the students who held a nonstochastic conception indicated thinking about repeated sampling and sample size. However, their explanations focused on sampling rates and the impact of sample size on a sampling rate. Students who held a nonstochastic conception and those who held no stochastic image merely indicated thinking about how differing values for sample size affected calculated results. These two groups of students did not specify conditions related to a stochastic image of a repeatable process. Table 5.5 summarizes images presented by interview participants that related to specifications of conditions of a repeatable process. The summary of stochastic images reported for each stochastic conception category includes evidence indicated by all students who are characterized as holding that particular conception. *May* is used to indicate evidence reported by some explanations, but this evidence was not required criteria.

Table 5.5

Summary of Images for Specification of Conditions

Conception	Image for Specification of Conditions
Stochastic	Indicates that repetition of the process yields a collection of variable outcomes
	Assumes outcomes are independent
	Describes a sampling process where each selection is equally likely
	Indicates the sampling process produces samples that are representative of population
	Indicates that variability in outcomes is related to sample size
	Indicates conceiving of conditions of the process in relation to an underlying model
Situational	Indicates that repetition of the process yields a collection of variable outcomes
	May assume outcomes are independent
	May describes a sampling process where each selection is equally likely
	May indicate thinking about repeated sampling; does not indicate sampling process produces samples that are representative of population
	May indicate that variability in outcomes is related to sample size
	May conceive of conditions of the process in relation to an underlying model
Nonstochastic	Does not indicate that repetition of the process yields a collection of variable outcomes
	May assume samples are independent, but does not assume outcomes are independent
	May indicate thinking about repeated sampling, but does not describe a sampling process
	Indicates thinking about sample size, but does not relate sample size to variability in outcomes
	May indicate thinking about the impact of sample size on sampling rates or ratios
	Does not indicate thinking about an underlying model
No Image	Does not indicate thinking about a collection of outcomes
	Expresses reluctance to make assumptions; may mention independence
	Does not indicate thinking about a sampling process or repeated sampling
	Does not indicate thinking about sample size in relation to variability in outcomes
	Indicates thinking that consideration of sample size is only important for use in calculations
	Does not indicate thinking about a model

Note. Italicized images were only evidenced as a result of the interview. Images not italicized were evidenced as a result of the interview and via student's individual written answers for the conceptual quiz.

Research Sub-question Component 1(c) Image of Distribution of Outcomes

How do students characterize a probability situation in terms of an image of a distribution of outcomes? Students characterized as holding a stochastic conception consistently indicated thinking about distribution and a distribution of outcomes. These students attended to variability when describing their thinking about a distribution of outcomes. They also connected thinking about expectation to thinking about variability. Furthermore, they connected notions of distribution, variance, shape, and sample size. Students characterized as holding a situational conception reported thinking about a distribution of outcomes and connected thinking about expectation to variability, but only in the problem context for which they exhibited stochastic reasoning. Students who held a situational conception did not consistently attend to thinking about variability. Students characterized as holding a nonstochastic conception, as well as those characterized as holding no image related to a stochastic conception, did not mention or refer to a distribution of outcomes. Furthermore, students who held no image of a stochastic conception reported thinking about expectation as a fixed value and did not connect thinking about expectation to variability. They interpreted distributions as a curve, graph, or table of values used when doing calculations.

The most apparent distinction between students who are characterized as holding a stochastic conception in both problem contexts and students who held other conceptions is that students who held a stochastic conception spontaneously talked about how they connected their thinking about a repeatable process to a model. They conceived of conditions of the process in relation to an underlying model. They also indicated thinking about an underlying distribution as a model. Furthermore, students who held a stochastic conception quantified the meaning of *unusual* as deviation from expectation in terms of a distribution model. Two of the students who held a situational conception indicated thinking about a model of repetition for the problem context in which they exhibited stochastic reasoning, but only after being prompted during questioning. These two students also connected their thinking about some conditions of the process to an underlying model for the problem context in which they exhibited stochastic reasoning. They also described thinking about an underlying distribution model. However, students who held a situational conception did not conceive of *unusual* as deviation from expectation in relation to a distribution. One of these students indicated thinking about *unusual* as not being near the expected value. The other two students who held a situational conception indicated thinking *unusual* means having a low probability. Students who held a nonstochastic conception and students who held no stochastic conception did not report thinking about an underlying distribution as a model. These students described distributions as something used to calculate probabilities. Table 5.6 summarizes images presented by interview participants that related to a distribution of outcomes. The summary of stochastic images reported for each stochastic conception category includes evidence indicated by all students who are characterized as holding that particular conception. *May* is used to indicate evidence reported by some explanations, but this evidence was not required criteria.

Table 5.6

Summary of Images for Distribution of Outcomes

Conception	Images for Distribution of Outcomes

Stochastic	*Indicates thinking about a distribution of outcomes*
	Connects thinking about expectation to variability
	Attends to variability when thinking about distribution of outcomes; connects notions of distribution variance, shape, and sample size
	Indicates thinking about the law of large numbers in relation to stabilization of frequencies over a large number of repetitions of the process
	Indicates thinking about an underlying distribution model
	Quantifies "unusual" *as deviation from expectation in terms of a distribution modal*
Situational	*May indicate thinking about a distribution of outcomes*
	May connect thinking about expectation to variability
	May indicate thinking about sample size in relation to various aspects of distribution
	May indicate thinking about the law of large numbers in relation to stabilization of frequencies over a large number of repetitions of the process
	May indicate thinking about an underlying distribution model
	May quantifies "unusual" *as deviation from expectation or may indicate thinking "unusual" means having a low probability*
Nonstochastic	*Does not indicate thinking about a distribution of outcomes*
	May connect thinking about expectation to variability or variance
	May indicate thinking about distribution variance or variability
	Does not indicate thinking about law of large numbers in relation to stabilization of frequencies; when prompted recalls the law of large numbers as a rule
	Conceives of distribution as something used to calculate probability; mentions normal distribution
	Does not quantify "unusual"*; indicates thinking that "unusual" means not the expected value*
No Image	*Does not indicate thinking about a distribution of outcomes*
	Indicates thinking about expectation as a fixed value; does not connect thinking about expectation to variability
	Indicates thinking about distribution as a curve, graph, or table used for calculations
	Does not indicate thinking about law of large numbers in relation to stabilization of frequencies; when prompted recalls the law of large numbers as a rule
	Names distributions; indicates thinking distributions have something to do with probability
	Does not quantify "unusual"*; may indicate thinking about "unusual" in relation to expected value, or may indicate "unusual" means not likely*

Note. Italicized images were only evidenced as a result of the interview. Images not italicized were evidenced as a result of the interview and via student's individual written answers for the conceptual quiz.

CHAPTER 6: DISCUSSION

This study investigated the impact of an instructional intervention designed to support development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory, calculus-based probability and statistics course. An analysis of research and literature related to students' understanding of probability suggested the need for a large-scale study investigating understandings developed in the context of learning in a college-level probability and statistics course. A major learning goal for an introductory probability and statistics course is an understanding of statistical inference. Statistical inference is multifaceted and involves a coordination of probabilistic and statistical thinking grounded in conceptions of randomness. Stochastic reasoning supports normative conceptions of randomness, a stochastic understanding of probability distribution, and understanding of statistical inference. A stochastic understanding of probability distribution undergirds coordination of probabilistic and statistical thinking, facilitates conceptual connections between empirical distributions of data with theoretical probability distribution models, and undergirds understanding of statistical inference. This study implemented an instructional intervention which consisted of supplemental lab assignments with anticipatory tasks designed to elicit stochastic conceptions and to support development of a stochastic understanding of probability distribution.

This study investigated the impact of the instructional intervention through a control-treatment design. The study employed a mix of both quantitative and qualitative research methods to examine students' understandings of probability and statistics that resulted from instructional course interventions in the form of supplemental lab materials designed to support a stochastic understanding of probability distribution. In addition to the usual statistics instruction delivered in the course through a lecture-discussion format, participants in the treatment group received supplemental lab assignments aimed at development of stochastic reasoning and participants in the control group received supplemental lab assignments consisting of calculus review and practice problems. Quantitative data sources consisted of a student background survey, a conceptual assessment of students' stochastic understanding of probability, ARTIST assessment items from the Confidence Interval and Sampling sub-scales, and final course examinations. At the end of the course, 12 students participated in individual interviews. Six of the interview participants were selected from the treatment group and six were selected from the control group. The purpose of the interviews was to probe students' generalized stochastic reasoning as well as their reasoning about two contexts, which students encountered on the conceptual quiz. Qualitative data from the interviews provided insight into the thinking and reasoning underlying students' written answers on the conceptual quiz and informed development of a scoring rubric for the conceptual assessment. This rubric was subsequently used to score all 186 study participants' written answers on the conceptual quiz. Analysis of the collected qualitative data collected also served as a means of examining the validity of the stochastic conceptual assessment.

This chapter provides a discussion of the study and is presented in four sections. The first section gives a summary of results and overall conclusions of the study. The second section presents contributions and implications of the study, while the third section explains the limitations of the study. The final section offers suggestions for future research.

Summary of Results and Overall Conclusions

This research was designed to investigate the impact of an instructional intervention on undergraduate students' stochastic understanding of probability distribution in a calculus-based introductory probability and statistics course. The instructional intervention consisted of six supplemental lab assignments, which students completed outside of regular class-time. The students in the treatment group received lab assignments which were designed to support development of stochastic reasoning and development of principled knowledge of probability distribution. The lab assignments consisted of a series of anticipatory tasks designed to engage students in coordinating their thinking about complementary probabilistic and statistical notions related to a stochastic understanding of probability distribution. Components of the lab assignments were used as bridging tools (Abrahamson, 2007) to enable students to connect their intuitive understandings with formal mathematical, probabilistic, and statistical knowledge. The stochastic lab assignments were designed to develop stochastic anticipations (Simon, 2013) and support stochastic understandings of probability distribution and undergird deeper understanding of course content subsequently presented by their professors and discussion teachers.

This study investigated probabilistic and statistical conceptions held by undergraduate students in a calculus-based introductory probability and statistics course with a relatively strong background in mathematics, having completed at least two semesters of calculus prior to enrollment in the course. In addition, the study was designed to investigate characteristics of students' understanding of probability distribution resulting from an instructional intervention consisting of supplemental lab assignments. The stochastic-reasoning lab assignments were carefully structured to move the learner along a hypothetical learning trajectory aimed at development of stochastic understanding of probability distribution, while calculus-review lab assignments provided students with an opportunity to review and practice calculus content required for the course. This study investigated the impact of the instructional intervention on students' understandings at the end of a one-semester, calculus-based introductory probability and statistic course by addressing the following research question: What is the impact of an instructional intervention designed to support development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory, calculus-based probability and statistics course? Four research sub-questions were also addressed by this study.

(1) What is the nature of students' reasoning when confronted with a probability situation?

  a) How do students characterize a probability situation in terms of an image of a repeatable process?

  b) How do students characterize a probability situation in terms of specification of conditions of a repeatable process?

  c) How do students characterize a probability situation in terms of an image of a distribution of outcomes?

(2) Does instruction designed to support development of stochastic understanding of probability distribution impact students' stochastic

conceptions of a probability situation as evidenced on a conceptual assessment?

(3) Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding of confidence intervals as measured by the ARTIST assessment?

(4) Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding as evidenced on final course examinations administered in an introductory, calculus-based probability and statistics course?

In order to address the first research sub-question, qualitative data were collected and analyzed. Research sub-questions (2), (3), and (4) were addressed via collection and analysis of quantitative data. This section includes a summary of the qualitative results followed by a summary of the quantitative results. Finally, this section offers a concluding response to the main research question and provides an integrated analysis of both the qualitative and quantitative results.

Summary of Qualitative Results

Twelve students were selected to participate in an end-of-course interview and their stochastic conceptions were analyzed using a conceptual framework that was informed by three hierarchical image categories which characterize a stochastic conception of probability (Liu & Thompson, 2007). Students' overall stochastic conception was evidenced by reasoning across all three hierarchical categories: image of a repeatable process, image of specification of conditions, and image of distribution. Because the image categories are hierarchical, students who presented an image of distribution for a particular problem context, the highest level of conceptualization, also presented an image of specification of conditions, and an image of a repeatable process for that problem context. Students who presented an image of specification of conditions also presented an image of a repeatable process.

Responses to interview probes and written work for the two problem contexts on the conceptual quiz were analyzed. Students who exhibited evidence of stochastic reasoning across all three categories of the conceptual framework for both problem contexts and interview probes were characterized as holding a stochastic conception. Students who exhibited evidence of stochastic reasoning across all three categories for only one of the two problem contexts were characterized as holding a situational conception. Students who exhibited an image of specification of conditions and/or an image of a repeatable process in one or both problem contexts, but did not exhibit an image of distribution for either context, did not exhibit evidence of stochastic reasoning for either problem context and were characterized as holding a nonstochastic conception. Some students did not exhibit any image related to a stochastic conception for either problem context. These students are characterized as holding no images related to a stochastic conception.

Summaries of results addressing each of the three component categories of the framework are followed by a summary of results addressing the probabilistic nature of students' overall conceptions for the two problem contexts. The three components of research sub-question (1) address each of the three categories of the conceptual framework: image of a repeatable process, specification of conditions of the process, image of distribution of outcomes.

How do students characterize a probability situation in terms of an image of a repeatable process? Students who were characterized as holding a stochastic conception indicated an image of a repeatable process for both problem contexts. These three students indicated thinking that this process was repeated under essentially the same conditions. They also indicated thinking that the repeatable process yielded outcomes and described outcomes of the process. In both problem contexts, students who held a stochastic conception related thinking about the process to repeating an experiment. They also indicated thinking about the process as repeated sampling. In contrast to other students, those who held a stochastic conception spontaneously talked about a model of repetition when explaining their reasoning about the repeatable process. Of the three students who held a stochastic conception, two had completed the stochastic lab assignments, and they also spontaneously mentioned thinking about simulations. Furthermore, both of these students spontaneously talked about connections between thinking about the repeatable process they envisioned with regards to the two problem contexts and their experiences running simulations using *Fathom*.

The three students who were characterized as holding a situational conception also indicated an image of a repeatable process for both problem contexts and indicated thinking that this process was repeated under essentially the same conditions. However, students who held a situational conception only indicated thinking about repeating an experiment or repeated sampling for the problem content in which they indicated stochastic reasoning. None of the students who held a situational conception spontaneously talked about a model of repetition. However, when prompted, these students connected thinking about a model of repetition to reasoning about a repeatable process for the problem context in which they indicated stochastic reasoning. All three students who held a situational conception completed the stochastic reasoning lab assignments. However, none of these students specifically connected thinking about simulations to thinking about a repeatable process for the problem contexts on the conceptual quiz.

There were three students who were characterized as holding a nonstochastic conception; they indicated thinking about a repeatable process for at least one of the two problem contexts and described outcomes of the process for that context. However, these students did not consistently express thinking the process would be repeated under essentially the same conditions. Although two of the students talked about repeated sampling in the theater problem context, one of these individuals expressed thinking about a process that changed from day to day. In contrast to students who held either a stochastic conception or situation conception, students who held a nonstochastic conception gave no indication of thinking about repeated experiment. Even when prompted, none of these students expressed thinking about the process in relation to a model. All three students who held a nonstochastic conception indicated thinking about probability as a fixed ratio.

Students who were characterized as holding no stochastic image did not indicate thinking about a repeatable process. These three students indicated thinking about rates or proportions and thinking about probability as a fixed ratio. Furthermore, these students focused on mathematical calculations and did not indicate thinking about a model.

How do students characterize a probability situation in terms of specification of conditions of a repeatable process? Students who were characterized as holding a stochastic conception indicated thinking that the process yielded a collection of independent, variable outcomes. These students described a sampling process where each selection was equally likely and the resulting samples were representative of the population from which they were sampled. Students who held a stochastic conception indicated thinking that the variability in outcomes was related to sample size. Furthermore, they spontaneously indicated thinking about conditions of the process in relation to an underlying model.

Students who were characterized as holding a situational conception indicated thinking that the process yielded a collection of variable outcomes, but these students did not consistently assume the outcomes were independent. For the problem context in which they indicated stochastic reasoning, students who held a situational conception described a sampling process where each selection was equally likely. However, they did not express thinking that the resulting samples would be representative of the population from which they were sampled. Students who held a situational conception indicated thinking that variability in outcomes was related to sample size only for the problem context in which they indicated stochastic reasoning. None of these students spontaneously indicated thinking about conditions of the process in relation to an underlying model, but when prompted they were able to connect thinking about a model to condition of the process for the problem context in which they indicated stochastic reasoning.

Students who were characterized as holding a nonstochastic conception did not indicate thinking about a collection of variable outcomes. Some of the students who held a nonstochastic conception indicated thinking about repeated sampling and assuming samples would be independent. However, these students did not describe the sampling process and did not assume outcomes would be independent. Although students who were characterized as holding a nonstochastic conception indicated thinking about sample size, they did not connect sample size to variability. Two of these students indicated thinking about the impact a change in sample size had on a sampling rate.

Students who were characterized as holding no image related to a stochastic conception did not indicate thinking about a collection of outcomes. Furthermore, these students expressed reluctance to make assumptions. They did not indicate thinking about a sampling process or repeated sampling and thus did not indicate thinking about sample size in relation to either variability or outcomes. Students who held no image related to a stochastic conception indicated that thinking about sample size was only important for use in doing calculations.

How do students characterize a probability situation in terms of an image of a distribution of outcomes? Students who were characterized as holding a stochastic conception indicated thinking about a distribution of outcomes for both problem contexts. These students expressed thinking about variability when describing their thinking about a distribution of outcomes. Furthermore, they connected notions of shape, variances, and sample size when describing distributions. Students who held a stochastic conception connected thinking about expectation to variability. They also indicated thinking about an underlying distribution model and connected thinking about the model to empirical outcomes. These students quantified *unusual* as deviation from expectation in terms of a distribution model. Students who held a stochastic conception recognized that frequencies of outcomes of the repeatable process would stabilize over a large number of repetitions and indicated thinking about the law of large numbers in relation to variability in frequencies of observed outcomes.

Students who were characterized as holding a situational conception only indicated thinking about a distribution of outcomes for the problem context in which they indicated stochastic reasoning and only two students expressed thinking about sample size in relation to distribution. Students who held a stochastic conception only connected thinking about expectation to variability for the context in which they indicated stochastic reasoning. Furthermore, these students only indicated thinking about an underlying distribution model for the context in which they indicated stochastic reasoning. None of the students who held a situational conception quantified *unusual* in terms of a distribution model. One student, who held a situational conception, indicated stochastic reasoning in the theater problem context and quantified *unusual* as deviation from the expected value. The other two students indicated thinking that unusual means having a low probability. One of the students who held a situational conception recognized that frequencies of outcomes of the repeatable process would stabilize over a large number of repetitions and indicated thinking about the law of large numbers in relation to the hospital problem context. Another student who held a situational conception related thinking about stabilization of relative frequencies to his experience with running simulations using *Fathom*.

Students who were characterized as holding a nonstochastic conception, as well as the students who were characterized as holding no image related to a stochastic conception, did not indicate thinking about a distribution in relation to outcomes of a process. These students conceived of distribution as an object used to calculate probability. The students who held a nonstochastic conception indicated thinking about distribution variance and variability within a distribution, but only one of these students connected thinking about expectation to variance. None of these students indicated thinking about stabilization of frequencies of outcomes. Two of the students who held a nonstochastic conception did not recall learning about the law of large numbers. The other student who held a nonstochastic conception, along with the students who held no stochastic images, recalled the law of large numbers when prompted and indicated thinking it was used as a rule for making a decision about when to apply the normal distribution. None of these students quantified *unusual* as deviation from the mean, but thought of *unusual* as not the expected value or not likely.

What is the nature of students' reasoning when confronted with a probability situation? Students who were characterized as holding a stochastic conception consistently described their thinking in terms of a model that represented the problem situation and gave examples to illustrate their thinking. Students who held a stochastic conception consistently coordinated their thinking about experimental and theoretical probability and consistently coordinated their thinking about empirical distributions and theoretical probability distributions. When explaining their theoretical understanding, they often gave an example of an empirical situation to which the theory applied. The manner of response to interview probes exhibited by students who held a stochastic conception differed from the other interview participants. Students who held a stochastic conception appeared more confident, and their answers were relatively concise and to the point.

Students who were characterized as holding a situational conception displayed conflated thinking during the interview, evidenced by predominantly lengthy explanations and some contradictions in thinking. These students demonstrated coordination of their thinking about empirical distributions and theoretical probability distributions in one problem context, but struggled with coordination of these notions in the other problem context. It was apparent that for students who held a situational conception, the context of the problem situation acutely influenced their stochastic thinking. While students who held a situational conception exhibited stochastic reasoning, this reasoning appeared fragile and was context dependent.

Students who were characterized as holding either a nonstochastic image or no image related to a stochastic conception approached the problem situations by anticipating a formula to use and an answer to calculate. These students indicated a focus on procedures, formulas, and mathematical calculations. None of these students attempted to coordinate thinking about empirical distributions and theoretical probability distributions. These students exhibited confidence that they had used the correct numbers and therefore had the correct answer. Students who were characterized as holding either a nonstochastic image or no image related to a stochastic conception did not express conflated thinking, but focused on determining a value for the theoretical probability.

Summary of Quantitative Results

Does instruction designed to support development of stochastic understanding of probability distribution impact students' stochastic conceptions of a probability situation as evidenced on a conceptual assessment? Results of quantitative analyses indicated that the instructional intervention had a statistically significant impact on students' stochastic understanding of probability. Students who were in the stochastic reasoning lab group were significantly more likely to indicate a stochastic conception for one or both problem contexts on the conceptual quiz than students who were in the calculus review lab group. Furthermore, investigation of an interaction between lab group and completion of all lab assignments showed that students in the stochastic reasoning lab group who completed all of the lab assignments had a higher mean stochastic image than students than students in the stochastic reasoning lab group who did not complete all of the lab assignments. In addition, students in the stochastic reasoning lab group who did not complete all of the lab assignments had a mean stochastic image that was on par with students in the calculus review lab group who completed all lab assignments. Students in the stochastic reasoning lab group who did not complete all of the lab assignments also had a mean stochastic image on par with students in the calculus review group who did not complete all of the lab assignments. Thus, the evidence suggests that completion of all of the stochastic lab assignments promoted development of stochastic reasoning.

Although this study did not specifically investigate instruction that took place during the lecture and discussion classes, quantitative results show that this instruction mattered. Students enrolled in one of lecture classes were significantly more likely to exhibit evidence of a stochastic conception. The evidence suggests that lecture-class instruction impacted development of students' stochastic reasoning.

Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding of confidence intervals as measured by the ARTIST assessment? Analysis of quantitative results provided no evidence that the instructional intervention impacted students' understanding of confidence intervals as measured by ARTIST assessment items. Scores on the confidence interval assessment for students in the stochastic reasoning group did not differ from scores on this assessment for students in the calculus review lab group. However, analyses showed there was a statistically significant relationship between students' mathematical background and their score on the confidence interval assessment. Students who earned credits for their first semester of calculus via AP Calculus as opposed to earning those credits while in college scored significantly higher on the confidence interval assessment. There was also a statistically significant relationship between students' course background in statistics and their score on the confidence interval assessment. Some study participants had taken an algebra-based statistics course prior to enrollment in the course in which the study took place. Students who had earned credits for a prior college-level statistics course scored significantly higher on the confidence interval assessment. These results suggest that students with stronger backgrounds in mathematics or statistics achieved a higher score on confidence interval items in the ARTIST assessment.

Does instruction designed to support development of stochastic understanding of probability distribution impact students' understanding as

evidenced on final course examinations administered in an introductory calculus-based probability and statistics course? Analysis of quantitative results provided no evidence that the instructional intervention impacted students' scores on the final course examination. Final examination scores for students in the stochastic-reasoning lab group did not differ from final examination scores for students in the calculus-review lab group. However, analyses also showed that students who completed all of lab assignments scored significantly higher on the final course examination. Completion of all lab assignments may be indicative of student diligence; students who completed all of the lab assignments may have completed more of the work required in the course. Analyses also showed that students who earned credits for first-semester calculus via AP Calculus scored significantly higher on the final course examination than students who earned credits for first-semester calculus while in college. Earning credits for AP Calculus can be considered a proxy for mathematical background and prior academic achievement. This statistically significant relationship may mean higher achieving students with a stronger background in mathematics performed better on the final course examination.

Conclusion

What is the impact of an instructional intervention designed to support development of stochastic understanding of probability distribution of undergraduate students enrolled in an introductory calculus-based probability and statistics course? Taken together, qualitative and quantitative results show that the instructional intervention designed to support development of stochastic understanding of probability distribution impacted students' stochastic thinking and promoted a stochastic understanding of probability distribution. Students who completed all of the stochastic reasoning labs were significantly more likely to exhibit evidence of stochastic reasoning. Many of these students indicated that they consistently held a stochastic conception in spite of the problem context, which, as demonstrated in the qualitative analyses, included consistency in coordination of thinking about experimental and theoretical probability, as well as consistency in coordination of thinking about empirical distributions and theoretical probability distribution models. Students who held a stochastic conception of probability distribution were able to apply their theoretical understanding of probability distribution to differing empirical contexts. These students demonstrated evidence of thinking that included both a modeling perspective of distribution and a data-centric perspective of distribution. Furthermore, the evidence showed that these students consistently coordinated an experimental perspective of probability with a theoretical perspective of probability and were able to apply their stochastic thinking in different problem contexts.

Some students demonstrated situational stochastic conceptions which were context dependent. For these students, stochastic thinking was evident, but their stochastic reasoning was inconsistent across differing problem contexts. Thus, their stochastic reasoning appeared to be more tenuous. When prompted during the interview, students with situational stochastic conceptions coordinated a modeling perspective of distribution with a data-centric perspective of distribution. This evidence suggests that the stochastic lab assignments impacted students' stochastic reasoning and supported coordinated thinking about empirical distributions and probability distribution models.

The type of instructional lab intervention, stochastic reasoning labs or calculus review labs, did not impact students understanding of confidence intervals as measured by ARTIST items. In addition, there was no evidence that the type of instructional lab intervention impacted students' scores on the final course examination. While understanding confidence intervals includes an understanding of probability distribution, ARTIST assessment items were not designed to specifically assess stochastic conceptions of probability distribution. Questions on the final course examination required an understanding of probability distribution, but students could successfully answer these questions using knowledge of theoretical probability and calculus. In other words, correctly answering questions on the final course examination did not necessarily require coordinated thinking about empirical distributions and probability distribution models.

The stochastic reasoning lab assignments were designed to support development of stochastic reasoning along a hypothetical learning trajectory aimed at a stochastic understanding of probability distribution. While not all students in the stochastic-reasoning lab group presented evidence of stochastic reasoning, students in the stochastic-reasoning lab group had statistically significant higher odds of presenting a stochastic image than students in calculus-review lab group. These results indicate the stochastic reasoning lab assignments impacted students' movement along the hypothetical learning trajectory. Students who completed all of the stochastic reasoning labs showed a higher propensity towards indicating a stochastic understanding of probability distribution. Other research shows that stochastic conceptions of probability distribution support thinking about statistical inference (Liu & Thompson, 2007).

Contributions and Implications

Contributions

A major contribution of this study is empirical evidence documenting that supplemental instructional supports designed to fit a hypothetical learning trajectory that is aligned with the conceptual understandings addressed within the primary mode of instruction can impact students' understanding. A main finding of this study is that completion of supplemental assignments designed to undergird development of stochastic conceptions of probability and statistics had a statistically significant impact on students' understanding of probability distribution. Students who completed all of the stochastic reasoning lab assignments were significantly more likely to exhibit evidence of stochastic reasoning. Students in this study who presented stochastic conceptions demonstrated integrated reasoning of variation and coordinated a data-centric perspective with a modeling perspective of distribution, which confirms the importance of integrated reasoning within and across a data-centric perspective and modeling perspective of distribution (Peters, 2011; Pfannkuch, 2005). Students in this study who held stochastic conceptions presented images of sampling and sampling variability which were consistent with a multiplicative conception of sampling (Saldhana & Thompson, 2002). Furthermore, students who held stochastic conceptions consistently presented images of probability and probability distributions which support understanding of statistical inference (Liu & Thompson, 2005, 2007; Makar & Rubin, 2009; Rossman, 2008). The results of this study confirm and extend findings that instruction designed to promote stochastic conceptions leads to development of principled understandings of probability distribution and other key concepts related to statistical inference (Meletiou-Mavrotheris, 2003).

As documented in other studies, this study confirms that students, who had considerable knowledge of probability distribution models, do not necessarily indicate stochastic understanding of probability distribution and appear to have compartmentalized knowledge of probability and statistics (Abrahamson, 2007; Barragues, et al., 2007; Noll & Shaughnessy, 2012). However, stochastic reasoning lab assignments may impact students' movement along a hypothetical learning trajectory towards a stochastic understanding of probability distribution. Students who completed these labs were more likely to demonstrate evidence of stochastic thinking.

A few prior studies have offered evidence of post-calculus students' understanding of probability distribution, and this study contributes to the literature on post-calculus students' understanding after a one-semester, calculus-based introductory course in probability and statistics. Students in this study who did not hold stochastic conceptions indicated thinking that was similar to the majority of students in previous studies which involved post-calculus students (Abrahamson, 2007; Abrahamson & Wilensky, 2007; Barragues, et al., 2007; Hernandez, et al., 2006; Ives, 2007; Lunsford, et al., 2006; Wilensky, 1997). This study confirms previous findings that the majority of students in a calculus-based probability and statistics course demonstrate thinking that is aligned with high school students and college-level students in algebra-based introductory probability and statistics courses. This study also confirmed that most students in a calculus-based course do not present images related to a stochastic conception of probability and neither perceive the impact of randomness, nor understand the nature of random phenomena. Like the majority of post-calculus students in previous studies, when confronted with a problem situation, students in this study who held a nonstochastic conception focused on memorized procedures, formulas, and doing mathematical calculations. These students did not indicate a conceptual understanding of probability distribution, but merely viewed it as a tool to find the answer to a probability problem.

Results of this study confirmed that students who do not possess stochastic reasoning have difficulty perceiving the difference between empirical distributions and probability distribution models (Batanero, et al., 2004). This study confirms that without a stochastic conception of probability, students present an instrumental knowledge of probability and statistics rather the principled knowledge which is necessary for application of probability and statistics in practice (Barragues, et al., 2007).

Hypothetical learning trajectory for probability distribution. The hypothetical learning trajectory used in this study posits development of a stochastic understanding of probability distribution. This hypothetical learning trajectory was based on the theoretical framework Liu and Thompson (2007) developed resulting from analyses of eight high-school-statistics teachers' understanding of probability. This study used Liu and Thompson's (2007) framework as a tool for designing the stochastic reasoning instructional invention to support students' learning, which consisted of a series of six supplemental lab assignments aimed at moving learners along a hypothetical learning trajectory towards a stochastic understanding of probability distribution. Sequencing and design of the stochastic-reasoning lab assignments were intended to develop a series of ways of thinking posited by Liu and Thompson (2007), which were theorized to support stochastic understandings of probability and statistical inference. Results of this study confirm the usefulness of this framework as a tool for designing instruction to support development of students' stochastic conceptions of probability. In this large-scale study, students in the stochastic-reasoning lab group (treatment) were significantly more likely to indicate a stochastic conception for one or both problem contexts on the conceptual quiz as compared to students in the calculus-review lab group (control). This study offers empirical evidence which provides validity for the theoretical framework to support development of stochastic thinking as posited by Liu and Thompson (2007).

Liu and Thompson's (2007) framework was also used to investigate post-calculus students' understanding of probability distribution. Qualitative findings in this study confirm that students who did not hold stochastic conceptions exhibit compartmentalized understandings of probability and notions related to statistical inference. These findings also confirm Liu and Thompson's (2007) finding that students who did not hold stochastic conceptions exhibited conceptions of probability that were not grounded in a conception of distribution that supports thinking about statistical inference. On the other hand, students in this study who exhibited stochastic conceptions indicated evidence of coordinating their thinking about empirical distributions of data and probability distribution models, thus making connections between experimental and theoretical probability. The results of this study confirm that Liu and Thompsons' (2007) framework can be used to examine and assess probabilistic thinking.

Stochastic reasoning instructional supports. The results of this study advance a growing body of knowledge about instruction in probability and statistics at the college level. In this study, students in the control group received conventional instruction in probability and statistics similar to engineering students in Barragues' (2006) study, and results confirmed conventional instruction did not promote development of stochastic conceptions. However, students in the treatment (stochastic reasoning) group received instructional support in a learning environment designed to shape stochastic intuitions via anticipatory tasks and utilization of *Fathom*. This study found that the stochastic-reasoning lab assignments had statistically significant impact on students' understanding of probability, and the evidence suggests that well-designed instructional supports aligned with Liu and Thompson's (2007) framework can promote development of stochastic reasoning.

The stochastic-reasoning supplemental lab assignments were designed to promote stochastic anticipations (Simon, 2013). These lab assignments consisted of activities which were intended to elicit stochastic conceptions and to shape students' thinking in ways that would impact their learning in the lecture and discussion sessions. The rationale for lab assignment tasks was a hypothesis that students learn through their activity in all situations and that learning can be engendered through a well-designed task sequence (Simon, et al., 2010). The results of this large-scale study show that a well-designed, sequence of anticipatory tasks has the potential to impact development of students' stochastic understanding in an instructional setting.

The results of study confirm that coordination of thinking about empirical distributions and theoretical probability distribution models via simulations in a dynamic statistical environment supports development of stochastic reasoning (Meletiou-Mavrotheris, 2003). Findings from this study support other research which suggests a pedagogical approach to learning probability and statistics should employ features of a dynamic statistical environment such as *Fathom* (Lee, etal., 2010; Maxara & Biehler, 2010, Prodromou, 2012). Research suggests there is a link between the types of tools students use and their reasoning about variation and variability (Pfannkuch, 2005). The technology used in this study to support development of stochastic reasoning was *Fathom* dynamic software. This software was chosen specifically because it offers particular affordances, which can support development of stochastic conceptions. This study found that use of dynamic software, such as *Fathom*, in conjunction with well-designed tasks, can support development of integrated thinking regarding a data-centric perspective of variation with a modeling perspective of variation.

Implications

The results of this study show that particular anticipations and perspectives of distribution are related to a stochastic conception of probability distribution. Students who were characterized as holding stochastic conceptions attended to the notions of variability which included both a data-centric perspective and a modeling perspective of variability. The implication is that coordination of thinking about empirical distributions with probability distribution models is an important aspect of a stochastic conception of probability distribution and more robust stochastic conceptions include strong conceptual links between empirical and theoretical distributions. All three students who were characterized as holding a stochastic conception of probability distribution spontaneously indicted thinking about a model in relation the problem situations presented on the conceptual quiz, which implies that coordination of data-centric perspective with a modeling perspective could be an essential aspect of stochastic thinking.

The results of this study also imply that conceptions of repeated sampling are important to a stochastic conception of probability distribution. However, while notions of repetition of a process are necessary for a stochastic conception of probability distribution, these are not sufficient. The evidence presented implies that anticipation of repeated sampling, as well as thinking about repetition of an experiment, are both needed for a stochastic conception of probability distribution. All students who were characterized as holding stochastic conceptions indicated thinking about repeated sampling in relation to the repeatable process in addition to thinking about repeating an experiment. This way of thinking about sampling contrasted with that indicted by students who did not hold stochastic conceptions.

While some of the students who did not hold stochastic conceptions mentioned sampling, it was clear they were not thinking of a sampling process, and none mentioned thinking about sampling in relation to repetition of an experiment. Furthermore, students who did not hold stochastic conceptions did not indicate a perception of variability in outcomes, but indicated a thinking that focused on formulas and calculations. The implication is that anticipation of repeating sampling is a necessary component of stochastic thinking. Furthermore, thinking about repeated sampling while simultaneously thinking about repeating an experiment may be related to coordination of thinking about empirical and theoretical probability distributions, thus coordinating a data-centric perspective with a modeling perspective of distribution.

This analysis leads to yet another implication: Students need experiences with distributions of data generated via random processes along with experiences that support development of a modeling perspective of distribution. Experiences with distributions of data should include activities which support development of stochastic anticipations of repeated sampling. Furthermore, these experiences should also include supports which undergird connections between repeated sampling and probability experiments. Students in the stochastic-reasoning lab group engaged in activities which utilized *Fathom* to run virtual simulations and to represent distribution of outcomes resulting from the simulations. As these simulation runs were repeated, *Fathom s*oftware afforded visualization of connections between pseudo-random empirical distributions and theoretical distribution models. The implication is that these experiences with simulations in *Fathom* appear to support development of normative conceptions of randomness and random phenomena which undergird stochastic conceptions of probability distribution.

The findings show that completion of the stochastic reasoning lab assignments had a statistically significant impact on students' stochastic conceptions. The mean stochastic image for students in the stochastic reasoning lab group who completed all of the lab assignments was significantly higher than the mean stochastic image for all students in the calculus review lab group and was significantly higher than the mean stochastic image for students in the stochastic reasoning lab group who did not complete all of the lab assignments. Furthermore, the findings indicate that completion of the stochastic reasoning lab assignments facilitated movement along a hypothetical learning trajectory towards a stochastic understanding of probability distribution. Students who were characterized as holding a stochastic image indicated a principled understanding of probability distribution which included integration of stochastic understandings with instrumental understandings of probability. The implication of these findings is that in the future, lab assignments such as these should be a required component of a calculus-based, introductory probability and statistics course and not simply an optional assignment for extra-credit.

The results of study also imply that supplemental course instruction can have a significant impact on students' understanding. One of the implications related to supplemental course instruction is that the type of task and design of the instruction matter. A main objective of the stochastic-reasoning lab assignments was to elicit development of stochastic reasoning by means of anticipatory tasks. These tasks were designed to promote stochastic conceptions of probability distribution resulting from learned anticipations (Simon, 2013). The tasks were designed to prepare students to learn by promoting development of stochastic anticipations. Simon (2013) explains that sequencing of activities along a hypothetical learning trajectory is essential to instructional design focused on developing conceptions by developing anticipations from the activities. This implies that the design of anticipatory activities along a hypothetical learning trajectory aimed at stochastic understanding of probability distribution can impact students' conceptions. The significant difference in students' conceptual understanding of probability distribution was evidenced by particular ways of thinking about variability, distribution, and probability that necessarily undergird stochastic thinking. These results imply that tasks designed to support development of stochastic anticipations have the potential to change students thinking.

The results of this study have implications for curriculum and instruction in probability and statistics. While a curricular goal for all college-level probability and statistics courses is an understanding of statistical inference, statisticians focus on modeling variation for purposes of explanation (Gould, 2004; Pfannkuch, 2005). Curriculum focused on exploratory data analysis contrasts with curriculum focused on classical statistical inference. However, the results of this study show that coordination of a data-centric perspective with a modeling perspective of distribution undergirds development of stochastic conceptions of probability distribution. Because stochastic conceptions support thinking about statistical inference and modeling variation, the implication is that development of stochastic conceptions is important for all students learning probability and statistics and should be included in the curriculum. The results of this study offer empirical evidence pointing to the need for college-level curriculum in probability and statistics to include instruction which promotes development of stochastic conceptions of probability distribution.

Limitations

This study contributes to the field of statistics education, but there are limitations to the conclusions that can be drawn from these findings. As is the case for all studies in the field of education, the conclusions of this study are limited by the overall design. The participants in this study were students enrolled in a traditional, calculus-based, introductory probability and statistics course at a large university; caution should be exercised in generalizing the results of this study to other populations. While this study was implemented in conjunction with regular course instruction during the semester, the lab assignments were supplemental and not an integral part of the instruction which took place during the lecture and discussion sessions. Although students may have raised questions related to supplemental lab assignments during their lecture or discussion sessions, the professors and teaching assistants did not specifically address lab assignment material as part of that instruction.

Another possible limitation of this study is the level of student engagement with the lab assignments is unknown and most likely varied from student to student. At the beginning of the semester, a few students complained to the professor about the level of difficulty of the calculus-review lab assignments. These students expressed anxiety regarding their perception that the calculus-review assignments were too difficult. Students in this study had differing perceptions regarding relevance of lab material to course content, as well as differing perceptions regarding relevance of lab material to personal learning in the course.

There are other limitations related to lab assignment completion. While all students in the course had the opportunity to complete the supplemental lab assignments, students in the stochastic reasoning group were required to access *Fathom* software in order to complete each of the stochastic-reasoning lab assignments. This software was only available in the on-campus computer labs and students in the stochastic reasoning group had to go to one these locations to complete the lab assignment. Students in the calculus review group did not need to go to a specific location to complete their lab assignments. It is not known whether this distinction introduced bias into the study. Another limitation is possible cross-contamination of sample. It is unknown whether students worked individually or together on lab assignments. Furthermore, students in the calculus reasoning groups may have studied with friends who were in the stochastic reasoning group and vice versa.

In addition, there are limitations due to constraints regarding assessment of students' understanding. The assessment of stochastic conceptions was in the form of quiz which contained two items. The length of the conceptual assessment was constrained by access to students and by the time required to complete an in-class assessment. Students were given 20 minutes to complete the quiz in class during a discussion session at the end of the semester. Students in the two lecture classes took different final exams. Different content may have been emphasized during lecture and discussion class instruction. Furthermore, it was not feasible to add stochastic conception item(s) to the final course examination.

Finally, there are limitations associated with the qualitative results. While care was taken to choose interview participants who were representative of the students in each of the two groups, treatment and control, those selected to participate in interviews were a convenience sample of students who volunteered and, as such, may not be representative of the entire population. Each of the interview participants is an individual with unique characteristics and unique understandings and was chosen to maximize the potential for varied responses to interview probes. Caution should be exercised in generalizing the qualitative results of the study to other populations.

Future Research

Given the results of this study, there are several follow-up studies that build on this work in order to advance the field. The results of this study indicate that instruction designed to support development of stochastic reasoning moved students toward an understanding of probability distribution which included more integrated thinking about probability, variability, and distribution. However, the instructional supports were supplemental to regular course instruction. Given the promise of the instructional supports used in this study, a follow-up study should investigate the impact of instructional supports which are truly integrated into regular course instruction.

Analysis of data in this study revealed the importance of conceptions related to sampling and an understanding of sampling distribution. Thus, further studies should investigate the impact of stochastic reasoning instructional supports on students' understanding of sampling distribution. Research addressing students' understanding of statistical inference is still in its infancy. Stochastic reasoning is important to statistical inference, and further studies should investigate the impact of stochastic reasoning instructional supports on students' understanding of statistical inference.

This study was intentionally conducted with a select population of learners who had a stronger background in mathematics than the typical population of undergraduate students. While this study focused on instructional supports for an introductory, calculus-based probability and statistics course, similar studies should investigate the impact of instructional supports addressing development of stochastic reasoning for students enrolled in college-level, algebra-based courses. Curricular goals for calculus-based courses and algebra-based courses include understanding statistical inference, but these curricula differ with respect to coverage of probability distribution and use of mathematics. Research indicates there is a relationship between students' prior achievement in mathematics and statistics achievement, but questions remain regarding the nature of this relationship (Dupois, et al., 2012). Further studies should also investigate the relationship between students' prior mathematics achievement and stochastic understanding in algebra-based probability and statistics courses.

In this study, the hypothetical learning trajectory for understanding probability distribution was adapted from Liu and Thompson's (2007) investigation of high school mathematics teachers' understanding of probability. Participants in Liu and Thompson's study, as well as the participants in this study, had stronger backgrounds in mathematics than most students enrolled in a college-level, algebra-based probability and statistics course (Dupois, et al., 2012). Further studies should delineate a hypothetical learning trajectory for learners in algebra-based probability and statistics courses versus learners in calculus-based probability and statistics course.

While the results of this study indicate that instructional supports designed to support development of stochastic reasoning moved students toward a stochastic understanding of probability distribution, questions remain regarding how the lab assignment tasks changed students' thinking. Further studies should investigate how students' stochastic reasoning changes while engaging in tasks designed to support development of stochastic reasoning and investigate what role technology plays in changing students' thinking. Technology has changed professional practice in statistics, but questions remain regarding use of technology as a pedagogical tool in probability and statistics courses. Further studies should investigate how specific technologies, such as *Fathom* and other dynamic statistical environments, support development of stochastic reasoning. Further studies should also investigate how use of technological tools impacts stochastic thinking for students with varied backgrounds in mathematics.

Appendix A – Student Background Survey

1. What is your [course] discussion section number?

2. What is your academic major? (If you have a dual major, please list both majors.)

3. What is your academic minor?

4. How many college mathematics credits have you earned?

5. What is your class standing?

   Freshman Sophomore Junior Senior

6. What is your college GPA? Please check the range in which your college GPA falls.

   3.5-4.0 3.0-3.4 2.5-3.0 2.0-2.4 less than 2.0

7. What was your SAT Math score?

8. Did you earn credits for a course in probability and statistics prior to enrollment in [course]?

9. If you answered "yes" to Question 8, indicate the probability and statistics course (or equivalent) that you completed.
   - Stat 100 (or equivalent course at another college)
   - AP Statistics
   - Other college-level statistics course (not Stat 100 or AP Statistics)
   - High school probability and statistics course carrying 1 credit for high school graduation
   - None

10. What was the source of your first-semester calculus credits (or equivalent)?
    a. Math 140
    b. Math 130
    c. AP Calculus (AB)
    d. AP Calculus (BC)
    e. Other

11. If you marked "other" in Question 10, please provide the following:
    a. The name of the college/university where you took your first-semester calculus course.
    b. The name of the course.
    c. The number of credits you earned.

12. What was the source of your second-semester calculus credits (or equivalent)?
    a. Math 141
    b. Math 131
    c. AP Calculus (BC)
    d. Other

13. If you marked “other” in Question 12, please provide the following:
    a. The name of the college/university where you took your first-semester calculus course.
    b. The name of the course.
    c. The number of credits you earned.

Appendix B – Closure Questions

Stochastic-Reasoning Lab Closure Questions:

1. What is the most important thing you learned about randomness is Lab 1?
2. How does independence differ from dependence?
3. Why is an understanding of sample space important when thinking about a random variable?
4. Why is it important to understand the complementary nature of empirical and theoretical probability?
5. How is the distribution of a single sample related to the population distribution?
6. What is the most important thing you learned about sampling distribution in Lab 6?

Calculus-Review Lab Closure Questions:

1. What is an important [course] application of the calculus material you reviewed in Lab 1?
2. What question do you have about the material covered in Lab 2?
3. What material in Lab 3 do you think was helpful for you to review for this course?
4. What did you find helpful in the Lab 4 review?
5. In your opinion, when applying integration by substitution, what gives students the most trouble?
6. What did you find helpful in the Lab 6 review of integration by parts? If you did not find this review helpful, please provide feedback as to why you think the review was not helpful to you.

Appendix C – Stochastic Reasoning Lab Exemplar

Lab 3: Sample Space and Results of Flipping 4 Coins

Q1: If you flip four coins, how many heads do you think you will see?

Q2: Which of the following outcomes is most likely to result from a flip of four fair coins? ("H" represents a head showing up and "T" represents a tail showing up)?
a. HHHT b. THHT c. TTTT d. HTHT e. all four sequences are equally likely

One way to simulate flipping coins is to set up a collection with a single case representing the result of a flip of 4 coins. For the Fathom simulation, "H" will represent a coin with heads showing and "T" will represent a coin with tails showing. For example, the result "HHTT" will represent an outcome which is: heads showing on the first coin, heads showing on the second coin, tails showing on the third coin, and tails showing on the fourth coin.

Create a simulation.

Make and rename a collection

a. Open a new Fathom document. From the row of icons, pull a collection from the shelf. (Click on the Collection icon, then drag it onto the main *Fathom* white board). Rename the collection by clicking **Collection | Rename Collection** from the menu. Name the collection: "Results of Flipping 4 Coins."

Add case

b. Select the Results of Flipping 4 Coins collection. From the menu bar choose **Collection | New Cases**. Enter 1 in the box; click OK

Define case attributes

c. With the Results of Flipping 4 Coins collection selected, click on **Object | Inspect Collection** from the menu to open its inspector, or double-click to open its inspector. On the Cases tab, click on <new> and enter a new attribute named *coins4.* Enter the formula: concat(randomPick("H","T"), randomPick("H","T"), randomPick("H","T"), randomPick("H","T")). Then, click OK. Expand the collection icon to view the "Rerandomize" buttom. Click the button several times to view the simulated results of flipping 4 coins.

Sample cases

d. With the Results of Flipping 4 Coins collection selected, click on **Collection | Sample Cases** from the menu. Select the Sample of Results of Flipping 4 Coins icon, and pull a **Table** from the shelf (click on the Table icon, then drag it to the main *Fathom* white board). The table for Sample of Flip of Coin displays the simulated results of 10 virtual flips.

Create a graph

e. Pull a graph from the shelf to the Fathom whiteboard. In the table showing Sample of Results of Flipping 4 Coins, click on the attribute, *coins4,* and drag it to the empty box beneath the horizontal axis of the empty graph. A bar graph of the simulation results of 10 repetitions of flipping 4 coins will appear.

Q3: If the 4 coins were flipped 100 times, what do you think the possible outcomes would be?
- **(a)** Make a list of the possible outcomes.
- **(b)** What do you think the relative frequency for each outcome will be?

Reset this simulation

f. Reset the simulation to view the results of 100 flips of 4 coins. With the Sample of Flip of Coin collection selected, click **Object | Inspect Collection** from the menu to open its inspector. On the Sample tab, change the number of cases to 100. Click the Sample More Cases button to simulate results of 100 repetitions of flipping 4 coins. The table for Sample of Results of Flipping 4 Coins now lists the outcomes for each virtual flip. Sort the outcomes listed in the table by: (1) highlighting the *coins4* column in the table for Sample of Results of Flipping 4 Coins, (2) click **Table | Sort Descending** from the menu. The graph displays the frequencies of each outcome for 100 virtual repetitions of flipping 4 coins.

Q4: Copy and paste your screen shot onto your answer sheet, which shows the result of your simulation for 100 repetitions of flipping 4 coins. Compare your simulation results to your conjecture in **Q3**. Use the frequencies from your simulation results to determine the simulated relative frequency for each outcome that occurred for this particular simulation run.

Q5: List the possible outcomes for flipping 4 coins and determine the theoretical probability for each outcome.
Note: The probability of seeing a head in one flip of a fair coin is ½, and the probability of seeing a tail in one flip of a fair coin is ½. Use this fact along with the proposition (2.8) on page 77 of your textbook to determine the theoretical probabilities for each outcome.

Rerun the simulation.

g. Click the Sample More Case button to rerun the simulation several times. Pay attention to the graph representing the Sample of Results for Flipping 4 Coins. Observe how the frequency for each outcome varies for different simulation runs.

Q6: Explain the why the relative frequencies, which you reported in **Q4**, differ from the theoretical probabilities you determined in **Q5**.

Reset the simulation.

h. Reset the simulation to view the results of 10,000 flips of 4 coins by following the direction in step (f) and changing the number of cases on the sample tab to 10000. Be sure "Animation on" is not checked. Run the simulation.

Q7: Copy and paste your screen shot onto your answers sheet at this point in your simulation. What is the relative frequency for each outcome?

Q8: Group the outcomes you found in step (h) according to the number of heads. Use the theoretical probabilities you found in **Q5** to complete the following table. Copy and paste the completed table onto your answer sheet.

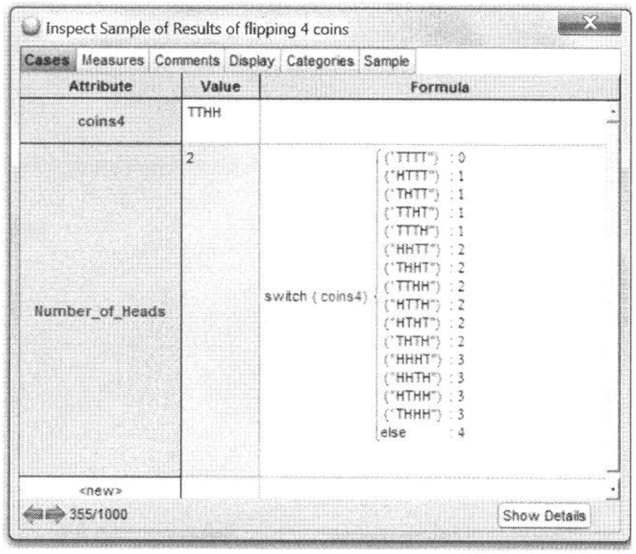
Number of Heads	Outcomes (list)	Number of Outcomes	Theoretical Probability
0			
1			
2			
3			
4			

Change how the simulation is represented.

Switch the representation

i. Combine the outcomes of your simulation to view the results in terms of the number of heads. Select the Sample of Results of Flipping 4 Coins collection and open its inspector, by clicking **Object | Inspect Collection** from the menu bar. On the Cases tab enter a new attribute, *Number_of_Heads*. Use the switch function to define *Number_of_Heads* by changing the representation of each the outcome.

See the example on the next page



Note: Be sure to put parenthesis around each outcome in the formula, i.e. ("TTTT"). To enter a new line on the left hand side of the switch function, click **control>enter**.

Graph the new attribute

j. Graph the empirical distribution for the Number of Heads observed for 10,000 repetitions of flipping 4 coins. Pull a new graph from the shelf. Open the inspector for Sample of Results of Flipping 4 Coins collection. Go to the cases tab. Click on the attribute, *Number_of_Heads,* and drag it to the empty box beneath the horizontal axis of the empty graph. A dot plot of the simulation results of 10,000 repetitions of flipping 4 coins will appear. Change the dot plot to a histogram by clicking in the box on upper right hand corner of the graph.

Compare the representations of the simulation

Q9: Copy and paste your screen shot onto your answers sheet at this point in your simulation. Note the graph for the frequency of *Number_of_Heads* for Sample of Results of Flipping 4 Coins collection, which shows the result of your simulation for 10,000 repetitions of flipping 4 coins in terms of the number of heads. Compare this graph with the graph of the Frequency of *coins4* (the outcomes) for Sample of Results of Flipping 4 Coins collection. Click on the histogram bar representing 1 Head. The histogram bar should change to the color red and the bars representing the outcomes which comprise 1 Head on the Frequency of *coins4* graph will also change to the color red.

Q10: Using your understanding of combinatorics, explain how the number of outcomes (corresponding to each histogram bar representing the frequency of *Number_of_Heads*) is related to the number of ways you can arrange the following:
(a) 4 Tails (i.e. no Heads)?
(b) 1 Head and 3 Tails?
(c) 2 Heads and 2 Tails?
(d) 3Heads and 1 Tail
(e) 4 Heads?

Compare the simulation results with theory

Compare the empirical results of the simulation with the theoretical probabilities.

k. Select the Sample of Flip of Coin collection and open its inspector, by clicking **Object | Inspect Collection** from the menu bar. On the measures tab enter the following measures along with an associated formula shown below:

Measure	Formula
no_heads	count(coins4 = "TTTT")
one_head	count(coins4 = "HTTT") + count(coins4 = "THTT") + count(coins4 = "TTHT") + count(coins4 = "TTTH")
two_heads	count(coins4 = "HHTT") + count(coins4 = "THHT") + count(coins4 = "TTHH") + count(coins4 = "HTTH") + count(coins4 = "HTHT") + count(coins4 = "THTH")
three_heads	count(coins4 = "HHHT") + count(coins4 = "HHTH") + count(coins4 = "HTHH") + count(coins4 = "THHH")
four_heads	count(coins4 = "HHHH")
Number_of_flips	count( )
Proportion_0Heads	no_heads ÷ Number_of_flips
Proportion_1Head	one_head ÷ Number_of_flips
Proportion_2Heads	two_heads ÷ Number_of_flips
Proportion_3Heads	three_heads ÷ Number_of_flips
Proportion_4Heads	four_heads ÷ Number_of_flips

The long-run relative frequencies for this probability experiment (the simulation you just ran) provide empirical probabilities for observing a particular number of Heads.

Enter measures

Q11: From the Measures tab on the inspector for the Sample of Flip of Coin collection enter values for the empirical probabilities in the following table. Also, enter the theoretical probabilities. Copy and paste the table onto your answer sheet. Are the values for the empirical probabilities reasonable? Explain.

	Empirical probability result from simulation	Theoretical probability
No Heads		
1 Head and 3 Tails		
2 Heads and 2 Tails		
3 Heads and 1 Tail		
4 Heads		

Q12: The sample space for the probability experiment of flipping 4 coins is {4 Tails, 1 Head and 3 Tails, 2 Heads and 2 Tails, 3 Heads and 1 Tail, 4 Heads}. When asked to give the sample space for a probability experiment where 4 coins are flipped, a student in a previous probability and statistics class wrote: {TTTT, HTTT, HHTT, HHHT, HHHH}. Explain why this representation might be problematic and could be indicative of a misunderstanding.

Q13: Explain why the outcomes listed in **Q2** can be equally likely and yet when flipping 4 coins, the probability of seeing two Heads and 2 Tails is greater than the probability of seeing 4 Tails.

Appendix D – Calculus Review Lab Exemplar

Lab 3: Review of Definite Integrals and Integral as Area under a Curve

1. (a) Graph the region bounded by: $0 \le f(x) \le$ x and $2 \le x \le 6$

   (b) Write an integral, which represents the area of the region, and integrate to find the area of the region.

2. (a) Graph the region bounded by: $0 \le f(x) \le e^{-x}$ and $0 \le x \le 4$

   (b) Write an integral, which represents the area of the region, and integrate to find the area of the region.

Evaluate each of the following. Be sure to show your work.

3. $\int_{1}^{4} \frac{x^2}{5} dx$

4. $\int_{0}^{6} [3-(x-6)^2] dx$

5. $\int_{-1}^{2} \frac{1}{4}(6-x^3) dx$

6. $\int_{0}^{1} e^{-2x} dx$

7. $\int_{1/2}^{2} \left(\frac{4}{9x^2}\right) dx$

8. $\int_{-1}^{1} 0.15(2-x^4) dx$

9. $\int_{0}^{1} 42x^6(1-x) dx$

Appendix E – Stochastic Conceptual Assessment

Show all your work and explain your reasoning.

1. A town has two hospitals. On the average, there are 45 babies delivered each day in the larger hospital. The smaller hospital has about 15 births each day. Fifty percent of all babies born in the town are boys. In one year each hospital recorded those days in which the number of boys born was 60% or more of the total deliveries for that day in that hospital.

Is it more likely that the larger hospital recorded more such days, that the smaller hospital did, or that the two hospitals roughly recorded the same number of such days? Explain your reasoning.

2. Anthony works at a theater, taking tickets for one movie per night at a theater that holds 250 people. The town has 30,000 people. He estimates that he knows 300 of them by name. Anthony noticed that he often saw at least two people he knew.

Assume that people are not coming to the theater because they know Anthony and there is nothing special about the type of movie. Is it in fact unusual that at least two people Anthony knows attend the movie? Give an explanation of your reasoning and be sure to address issues of randomness and distribution.

Appendix F – Confidence Interval Assessment

1. Two different samples will be taken form the same population of test scores where the population mean and standard deviation are unknown. The first sample will have 25 data values, and the second sample will have 64 data values. A 95% confidence interval will be constructed for each sample to estimate the population mean. Which confidence interval would you expect to have greater precision (a smaller width) for estimating the population mean?
   a. I expect the confidence interval based on the sample of 64 data values to be more precise.
   b. I expect the confidence interval based on the sample of 25 data values to be more precise.
   c. I expect both confidence intervals to have the same precision.

2. A 95% confidence interval is computed to estimate the mean household income for a city. Which of the following values will definitely be within the limits of this confidence interval?
   a. The population mean
   b. The sample mean
   c. The standard deviation of the sample mean
   d. None of the above

3. Each of the 110 students is a statistics class selects a different random sample of 35 Quiz scores from a population of 5000 scores they are given. Using their data, each student constructs a 90% confidence interval for μ the average Quiz score of the 5000 students. Which of the following conclusions is correct?
   a. About 10% of the sample means will not be included in the confidence intervals.
   b. About 90% of the confidence intervals will contain μ.
   c. It is probable that 90% of the confidence intervals will be identical.
   d. About 10% of the raw scores in the sampled will not be found in these confidence intervals

4. A 95% confidence interval for the mean reading achievement score for a population of third grade students is (43, 49). The margin of error of this interval is:
   a. 3
   b. 5
   c. 6

5. Justin and Hayley conducted a mission to a new planet, Planet X, to study arm length. They took a random sample of 100 Planet X residents and calculated a 95% confidence interval for the mean arm length. What does the 95% confidence interval for arm length tell us in this case? Select the best answer.
    a. I am 95% confident that this interval includes the sample mean arm length.
    b. I am confident that most (95%) of all Planet X residents will have an arm length within this interval.
    c. I am 95% confident that most Planet X residents will have arm lengths within this interval.
    d. I am 95% confident that this interval includes the population mean arm length.

6. Suppose that a random sample of 41 state college students is asked to measure the length of their right foot in centimeters. A 95% confidence interval for the mean foot length for students at this university turns out to be (21.709, 25.091). If instead a 90% confidence interval was calculated, how would it differ from the 95% confidence interval?
    a. The 90% confidence interval would be narrower.
    b. The 90% confidence interval would be wider.
    c. The 90% confidence interval would be the same as the 95% confidence interval .

7. A pollster took a random sample of 100 students from a large university and computed a 95% confidence interval to estimate the percentage of students who were planning to vote in the upcoming election. The pollster felt that the confidence interval was too wide to provide a precise estimate of the population parameter. What could the pollster have done to produce a narrower confidence interval that would produce a more precise estimate of the percentage of all university students who plan to vote in the upcoming election?
    a. Increase the confidence level to 99%.
    b. Increase the sample size to 150.
    c. Both a and b.
    d. None of the above.

8. A newspaper article states with 95% confidence that 55% to 65% of all high school students in the United States claim they could get a hand gun if they wanted one. This confidence interval is based on a poll of 2000 high school students in Detroit. How would you interpret the confidence interval from this newspaper article?
    a. 95% of large urban cities in the United States have 55% to 65% high school students who could get a hand gun.
    b. If we took many samples of high school students from different urban cities, 95% of the samples would have between 55% and 65% high school students who could get a hand gun.

c. We can be 95% confident that between 55% and 65% of all United States high school students could get a hand gun.
d. You cannot use this confidence interval to generalize to all teenagers in the United States because of the way the sample was taken.

9. The Gallup poll (August 23, 2002) reported that 53% of Americans said they would favor sending American troops to the Persian Gulf area in an attempt to remove Hussein from power. The poll also reported that the "margin of error" for this poll was 4%. What does the 4% margin of error indicate?
   a. There is a 4% chance that the estimate of 53% is wrong.
   b. The percent of Americans who are in favor is probably higher than 53% and closer to 57%.
   c. The percent of Americans who are in favor is estimated to be between 49% and 57%.

10. Suppose two researchers want to estimate the proportion of American college students who favor abolishing the penny. They both want to have about the same margin of error to estimate this proportion. However, Researcher 1 wants to estimate with 99% confidence and Researcher 2 wants to estimate with 95% confidence. Which researcher would need more students for her study in order to obtain the desired error margin?
    a. Researcher 1.
    b. Researcher 2.
    c. Both would need the same number of subjects.
    d. It is impossible to obtain the same margin of error with two different confidence levels.

11. A random sample of 25 college statistics textbook prices is obtained and the mean price is computed. To determine the probability of finding a more extreme mean than the one obtained from this random sample, you would need to refer to:
    a. The population distribution of all college statistic textbook prices
    b. The distribution of prices for this sample of college statistics textbooks
    c. The sampling distribution of textbook prices for all samples of 25 textbooks

12. Imagine you have a huge jar of candies that are a generic version of M&Ms. We know that 40% of the candies in the jar are brown. Imagine you create a sample by randomly pulling 20 candies out of the jar. If you repeated this 10 times to create 10 samples, each with 20 candies, about how many browns would you expect to find in each of the 10 samples?
    a. Each sample would have exactly 8 brown candies.
    b. Most of the samples would have 0 to 8 brown candies.
    c. Most of the samples would have 8 to 20 brown candies.
    d. Most of the samples would have 6 to 10 brown candies.
    e. You are just as likely to get any count of brown candies between 0 and 20.

References

Abrahamson, D. (2007). *The real world as a trick question: Undergraduate statistics majors' construction-based modeling of probability.* Paper presented at the American Education Research Association, Chicago, IL.

Abrahamson, D. (2009a). *Appropriate tools: On grounding mathematical procedures in perceptual intuitions*. Paper presented at the American Education Research Association, San Diego, CA.

Abrahamson, D. (2009b). Embodied design: Constructing means for constructing meaning. *Educational Studies in Mathematics*, *70*, 27-47. doi: 10.1007/s10649-008-9137-1

Abrahamson, D. (2009c). Orchestrating semiotic leaps from tacit to cultural quantitative reasoning: The case of anticipating experimental outcomes of a quasi-binomial random generator. *Cognition and Instruction*, *27*(3), 175-224.

Abrahamson, D. (2009d). A student's synthesis of tacit and mathematical knowledge as a researcher's lens on bridging learning theory. *International Electronic Journal of Mathematics Education*, *4*(3), 195-226. Retrieved from http://www.iejme.com/

Abrahamson, D., & Wilensky, U. (2007). Learning axes and bridging tools in a technology-based design for statistics. *International Journal of Computers for Mathematical Learning*, *12*(1), 23-55. doi: 10.1007/s10758-007-9110-6

Adler, J., & Davis, Z. (2006). Opening another black box: Researching mathematics for teaching in mathematics teacher education. *Journal for Research in Mathematics Education*, *37*(4), 270-296.

Artigue, M., Batanero, C., & Kent, P. (2007). Mathematics thinking and learning at the post-secondary level. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age Publishing.

Bakker, A., & Gravenmeijer, K. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Barragues, J., Guisasola, J., & Morais, A. (2006). Chance and probability: What do they mean to university engineering students? *International Journal of Mathematics Education in Science and Technology, 37*(8), 883-900. doi: 10.1080/00207390600818948

Barragues, J., Guisasola, J., & Morais, A. (2007). Which event is more likely? Estimating probability by university engineering students. *Recherches en Didactique des Mathematiques, 27*(1), 45-76.

Batanero, C. (2006). *The challenges of teaching statistical inference*. Paper presented at the Jornado de Classificacao e Anelise de Dados Univeridade Lusiada Lisboa, Lisbon, Portugal.

Batanero, C., Biehler, R., Maxara, C., Engel, J., & Vogl, M. (2005). *Using simulation to bridge teachers' content and pedagogical knowledge in probability.* Paper presented at the International Commission on Mathematics Instruction, Lindoia, Brazil.

Batanero, C., Godino, J., & Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education, 12*(1). Retrieved from http://www.amstat.org/publications/jse/

Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In G. Jones (Ed.), *Exploring probability in school*. New York: Springer.

Batanero, C., & Serrano, L. (1999). The meaning of randomness for secondary school students. *Journal for Research in Mathematics Education, 30*, 558-567.

Batanero, C., Tauber, L., & Sanchez, V. (2004). Students' reasoning about the normal distribution. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Ben-Zvi, D., & Garfield, J. (2004). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Biehler, R. (1994). Probabilistic thinking, statistical reasoning, and the search for causes: Do we need a probabilistic revolution after we have taught data analysis? In J. Garfield & G. Burrill (Eds.), *Research papers from the Fourth International Conference on Teaching Statistics, Marrakech 1994*. Minneapolis, MN: University of Minnesota.

Biehler, R. (1996). *Students' difficulties in practicing computer-suppported data analysis: Some hypothetical generalizations from results of two exploratory studies.* Paper presented at the International Conference on Teaching Statistics, Spain.

Biehler, R. (1999). Learning to think statistically and to cope with variation. *International Statistical Review*, 67(3), 259-262.

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In K. Clements (Ed.), *Third International Handbook of Mathematics Education* (pp. 643-689). New York: Springer.

Biggs, J., & Collis, K. (1991). Chapter 5: Multimodel learning and the quality of intelligent behavior. In H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Borgen, K., & Manu, S. (2002). What do students really understand? *The Journal of Mathematical Behavior, 21*(2), 151-165.

Brannen, J., & Moss, G. (2012). Critical Issues in Designing Mixed Methods Policy Research. *American Behavioral Scientist, 56*(6), 789-801.

Bransford, J., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Bush, W., & Greer, A. (2001). *Mathematics assessment: A practical handbook, grades 9-12*. Reston, VA: National Council for Teachers of Mathematics.

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Cobb, G. (1992). Teaching statistics. In L. A. Steen (Ed.), *Heeding the call for change*. Washington, DC: The Mathematical Association of America.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, *23*(7), 13-20.

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Corter, J., & Zahner, D. (2007). Use of external visual representations in probability problem solving. *Statistics Education Research Journal, 6*(1), 22-50. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ6(1)_Corter_Zahner.pdf

David, F. (1998). *Gods, games, and gambling: A history of probability and statistical ideas*. Mineola, NY: Dover Publications.

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal, 4*(1), 55-82. Retrieved from http://192.146.242.100/pub/netdocs/11/files/allegati/SERJ4(1)_delMas_Liu-2.2005.pdf

Devore, J. (2004). *Probability and statistics for engineering and the sciences* (6th ed.). Belmont, CA: Brooks/Cole-Thompson Learning.

Ernest, P. (1999). Forms of knowledge in mathematics and mathematics education: Philosophical and rhetorical perspectives. *Educational Studies in Mathematics, 38*, 67-83.

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgement. *Psychological Review, 104*, 301-318.

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

Fischbein, E. (1987). *Intuition in science and mathematics*. Dordrecht, The Netherlands: D. Reidel Publishing Company.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, *28*(1), 96-105.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, M., & Perry, M. (2005). A curriculum framework for PreK-12 statistics edcuation. Prepared by Guidelines for Assessment and Instruction in Statistics Education (GAISE) Retrieved from http://www.stat.amstat.org/education/gaise/

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*(1), 25-34. Retrieved from https://www.stat.auckland.ac.nz/~iase/publications/isr/95.Garfield.pdf

Garfield, J., Aliaga, M., Cobb, G., Cuff, C., Gould, R., Lock, R., & Witmer, J. (2005). Guidelines for assessment and instruction in statistics education (GAISE) college report. Retrieved from http://www.amstat.org/Education/gaise/GAISECollege.htm

Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. Statistics Education Research Journal, 4(1), 92-99. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. New York: Springer.

Garfield, J., delMas, R., & Chance, B. (2007). Using students' informal notions of variability to develop an understanding of formal measures of variability. In M. Lovett & P. Shah (Eds.), *Thinking with data*. New York: Lawrence Erlbaum Associates.

Gelman, R., & Greeno, J. (1989). On the nature of competence: Principles for understanding in a domain. In L. Resnick (Ed.), *Knowing, learning, and instruction* (pp. 125-186). Hillsdale, New Jersey: Lawrence Erlbaum.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British journal of psychology, 49*(3), 182-194.

Gibson, J. J. (1994). The visual perception of objective motion and subjective movement. *Psychological Review, 101*(2), 318-323.

Giuliano, M., Nemirovsky, I., Concari, S., Perez, S., Alvarez, M., & Sacerdoti, A. (2006). *Conceptions about probability and accuracy in Argentine students who start a career in engineering*. Paper presented at the Seventh International Conference on Teaching Statistics (ICOTS-7), Salvador, Brazil.

Green, David R. (1997). Recognizing randomness. *Teaching Statistics, 19,* 36-42.

Greeno, J. (1978). Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist, 12*(3), 262-283.

Greeno, J. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education, 22*(3), 170-218.

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

Hancock, C., Kaput, J., & Goldsmith, L. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337-364.

Hammerman, J., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal, 3*(2), 17-41. Retrieved from http://iase-web.org/documents/SERJ/SERJ3(2)_Hammerman_Rubin.pdf

Heid, M. K., Perkinson, D., Peters, S., & Fratto, C. (2005). *Making and managing distinctions: the case of sampling distributions.* Paper presented at the 27th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.

Heitele, D. (1975). An epistemological view of fundamental stochastic ideas. *Educational Studies in Mathematics, 6*, 187-205.

Hernandez, B., Huerta, A., & Batanero, C. (2006). *An exploratory study of students' difficulties with random variables.* Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil.

Hirsch, L., & O'Donnell, A. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education*, *9*(2). Retrieved from http://www.amstat.org/publications/jse/v9n2/hirsch.html

Inzunsa, S. (2008). *Probability calculus and connections between empirical and theoretical distributions through computer simulation*. Paper presented at the

Eleventh International Congress on Mathematics Education 2008 - Topic Study Group: Research and development in the teaching and learning of probability, Monterrey, Mexico.

Ireland, S., & Watson, J. M. (2009). Building a connection between experimental and theoretical aspects of probability. *International Electronic Journal of Mathematics Education, 4*(3), 230-260.

Ives, S. (2007). *Preservice teachers' knowledge of randomness and teaching probability*. Paper presented at the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Stateline (Lake Tahoe), NV.

Jones, G., Langrall, C., & Mooney, E. (2007). Research in probability. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age Publishing.

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

Kapadia, R., & Borovcnik, M. (Eds.). (1991). *Chance encounters: Probability in education*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Konold, C., & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 139-215). Reston, VA: National Council of Teachers of Mathematics.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education, 2*(1).

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289.

Korolov, L., & Sinai, Y. (2007). *Theory of probability and random processes* (2nd ed.). Heidelberg: Springer.

Kuhn, M., Hoppe, U., Lingnau, A., & Wichmann, A. (2006). Computational modelling and simulation fostering new approaches in learning probability. *Innovations in Education and Teaching International*, *43*(2), 183-194.

Kuntze, S., Engel, J., Martignon, L., & Gundlach, M. (2010). *Aspects of statistical literacy between competensy measures and indicators for conceptual knowledge: Empirical research in the project "RIKO-STAT"*. Paper presented at the 8th International Conference on Teaching Statistics, Ljubljana, Slovenia. Retrieved from http://icots.net/8/cd/pdfs/contributed/ICOTS8_C159_KUNTZE.pdf

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Education Research Journal, 1*, 29-63.

Lee, H. S., Angotti, R., & Tarr, J. (2010). Making comparisions between observed data and expected outcomes: Students' informal hypothesis testing with probability simulation tools. Statistics Education Research Journal, 9(1), 68-96. Retrieved from https://www.stat.auckland.ac.nz/~iase/serj/SERJ9(1)_Lee.pdf

Liu, Y., & Thompson, P. (2007). Teachers' Understandings of Probability. *Cognition & Instruction, 25*(2/3), 113-160.

Lunsford, M., Rowell, G., & Goodson-Espy, T. (2006). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. *Journal of Statistics Education, 14*(3). Retrieved from http://www.amstat.org/publications/JSE/v14n3/lunsford.html

MacGillivray, H. (2006). *Using data, student experiences, and collaboration in developing probabilistic reasoning at the introductory tertiery level.* Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil.

Makar, K., & Confrey, J. (2004). Secondary teachers' statistical reasoning in comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Martin, L. (2008). Folding back and the dynamical growth of mathematical understanding: Elaborating the Pirie–Kieren Theory. *The Journal of Mathematical Behavior*, *27*(1), 64-85.

Maxara, C., & Biehler, R. (2010). *Students' understanding and reasoning about sample size and the law of large numbers after a computer-intensive introductory course on stochastics*. Paper presented at the Eighth International Conference on Teaching Statistics, Ljubljana, Slovenia. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_3C2_MAXARA.pdf

Meletiou-Mavrotheris, M. (2003). Technological tools in the introductory statistics classroom: Effects on student understanding of inferential statistics. *International Journal of Computers for Mathematical Learning, 8*, 265-297.

Meletiou-Mavrotheris, M., & Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. Statistics Education Research Journal, 1(2), 22-37. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj

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

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*(2), 123-137.

Moore, D., & Cobb, G. (2000). Statistics and mathematics: Tension and cooperation. *The American Mathematical Monthly, 107*(7), 615-630.

Noll, J. (2011). Graduate teaching assistants' statistical content knowledge of sampling. *Statistics Education Research Journal, 10*(2), 48-74. Retrieved from http://iase-web.org/documents/SERJ/SERJ10(2)_Noll.pdf

Noll, J., & Shaughnessy, J. M. (2012). Aspects of students' reasoning about variation in empirical sampling distributions. *Journal for Research in Mathematics Education, 43*(5), 509-556.

Patton, M. (1990). *Qualitative Evaluation and Research Methods.* London: Sage Publishing.

Peters, S. (2011). Robust understanding of statistical variation. *Statistics Education Research Journal, 10*(1), 52-88. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ10(1)_Peters.pdf

Pfannkuch, M. (2005). Probability and statistical inference: How can teachers enable learners to make the connection? In G. Jones (Ed.), *Exploring probability in school* (pp. 267-294). New York: Springer.

Pfannkuch, M., & Reading, C. (2006). Reasoning about distribution: A complex process. *Statistics Education Research Journal*, *5*(2), 4-9. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2).pdf#page=7

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children* (L. Leake, Burrell, P., & Fischbein, H., trans.). New York: W.W. Norton & Company.

Pirie, S., & Kieren, T. (1994). Growth in mathematical understanding: How can we characterize it and how can we represent it? *Educational Studies in Mathematics, 26*, 165-190.

Pratt, D., Johnston-Wilder, P., Ainley, J., & Mason, J. (2008). Local and global thinking in statistical inference. Statistics Education Research Journal, 7(2), 107-129. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj

Prodromou, T. (2012). Students' construction of meanings about the co-ordination of the two epistemological perspectives on distribution. *International Journal of Statistics and Probability, 1*(2). doi: 10.5539/ijsp.v1n2p283

Prodromou, T., & Pratt, D. (2006). The role of causality in the co-ordination of two perspectives on distribution within a virtual simulation. *Statistics Education Research Journal, 5*(2), 69-88. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj

Reaburn, R. (2011). *Students' understanding of statistical inference: Implications for teaching* (Doctoral dissertation, University of Tasmania). Retrieved from http://eprints.utas.edu.au/12499/1/Whole.pdf

Reading, C., & Shaughnessy, Micheal. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Ross, S. (2006). *A first course in probability* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Roth, W. (2001). Situating cognition. *The Journal of the Learning Sciences, 10*, 27-61.

Rubel, L. (2007). Middle school and high school students' probabilistic reasoning on coin tasks. *Journal for Research in Mathematics Education*, *38*(5), 531-556.

Rubin, A., Bruce, B., & Tenney, Y. (1990). *Learning about sampling: Trouble at the core of statistics.* Paper presented at the Proceedings of the Third International Conference on Teaching Statistics, Dunedin, New Zealand.

Saldhana, L., & Thompson, P. (2006). *Investigating statistical unusualness in the context of a resampling activity: Students exploring connections between sampling distributions and statistical inference*. Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil. Retrieved from https://www.stat.auckland.ac.nz/~iase/publications/17/6A3_SALD.pdf

Saldhana, L., & Thompson, P. (2007). Exploring connections between sampling distributions and statistical inference: An analysis of students' engagement and thinking in the context of instruction involving repeating sampling. *International*

Electronic Journal of Mathematics Education, 2(3). Retrieved from www.iejme.com/032007/d032009.pdf

Sanchez, E., & Inzunsa, S. (2006). *Meanings' construction about sampling distributions in a dynamic statistics environment*. Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil.

Schwartz, D., & Bransford, J. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475-522.

Schwartz, D., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production and statistics instruction. *Cognition & Instruction, 22*(2), 129-184. Retrieved from http://aaalab.stanford.edu/papers/CI2202pp129-184.pdf

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematical thinking and learning* (pp. 465-494). New York: McMillan.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age Publishing.

Shaughnessy, J. M., Ciancetta, M., & Best, K. (2004). *Students' attention to variability when comparing distributions.* Paper presented at the Research Presession of the 82nd Annual Meeting of the National Council of Teachers of Mathematics Philadelphia, PA.

Shaughnessy, J. M., & Cianetta, M. (2002). *Students' understanding of variability in a probability environment.* Paper presented at the Sixth International Conference on Teaching Statistics, Cape Town, South Africa.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data Handling. In A. Bishop, Clements, K., Keitel, C., Kilpatrick, J., & Laborde (Eds.), *International handbook of mathematics education* (pp. 205-237). Dordrecht, The Netherlands: Kluwer Academic publishers.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*(2), 114-145.

Simon, M. (2006). Key developmental understandings in mathematics: A direction for investigating and establishing learning goals. *Mathematical Thinking and Learning, 8*(4), 359-371.

Simon, M. (2013). *Developing theory for design of mathematical task sequences: Conceptual learning as abstraction.* Paper presented at the ICMI Study 22: Task Design in Mathematics Education.

Skemp, R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching, 77*, 20-26.

Slauson, L. (2008). *Students' conceptual understanding of variability*. (Doctoral dissertation, The Ohio State University) Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/dissertations/dissertations.php

Smith, Toni. (2008). *An investigation into student understanding of statistical hypothesis testing*. (Doctoral dissertation, University of Maryland, College Park). Retreived from Proquest. (3324893)

Spillane, J. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition & Instruction, 18*(2), 141-179.

Stanovich, Keith. (2011). *Rationality and the reflective mind*: Oxford University Press.

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.

Steinbring, H. (1989). Routine and meaning in the mathematics classroom. *For the Learning of Mathematics, 9*(1), 24-33. Retrieved from http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-21335/Routine_and_meaning.pdf

Stigler, S. . (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge, MA: The Belknap Press of Harvard University Press.

Stijn, V., Castro Sotos, A., Onghena, P., & Vershcaffel, L. (2007). *Students' reasoning about sampling distributions before and after sampling distribution activity*. Paper presented at the 56th Session of the International Statistical Institute, Lisbon, Portugal.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publishing.

Thompson, P., Liu, Y., & Saldhana, L. (2007). Intricacies of statistical inference and teachers' understanding of them. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 207-232). New York: Lawrence Erlbaum Associates.

Tijms, H. (2007). *Understanding probability*. Cambridge, United Kingdom: Cambridge University Press.

van Oers, B. (2000). The appropriation of mathematical symbols: A psychosemiotic approach to mathematics learning *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 133-176).

Von Glasersfeld, E. (1991). *Radical constructivism in mathematics education* (Vol. 7): Springer.

Von Glasersfeld, E. (1996). Aspects of radical constructivism and its educational recommendations *Theories of mathematical learning* (pp. 307-314).

Vygotsky, L. (1978). *Mind in society*. Cambridge, Massachusetts: Harvard University Press.

Vygotsky, L. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge, MA: The MIT Press.

Watson, J. (2000). Preservice mathematics teachers' understanding of sampling: intuition or mathematics. *Mathematics Teacher Education and Development, 2*, 121-135. Retrieved from http://www.merga.net.au/documents/MTED_2_watson.pdf

Watson, J. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

West, R.,Toplak, M., & Stanovich, K. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology, 100*(4), 930.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal, 5*(2), 10-26. Retrieved from http://iase-web.org/documents/SERJ/SERJ5(2)_Wild.pdf

Wild, C., & Pfannkuch, M. (1999). Statistical thinking and empirical inquiry. *International Statistical Review, 67*(3), 223-248. Retrieved from https://www.stat.auckland.ac.nz/~iase/publications/isr/99.Wild.Pfannkuch.pdf

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics, 33*, 171-202.

Yackel, E., & Cobb, P. (1996). Sociomath norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458-477.

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education, 16*(2). Retrieved from http://www.amstat.org/publications/jse/V16n2/zieffler.pdf