CHAPTER ONE

Introduction

The purpose of this study is to research elementary preservice teachers' conceptions of variation. After defining variation in the introductory section, this chapter discusses four components which motivate the study. The first component is the primacy of variation to statistics and probability (which are together sometimes referred to as stochastics), and the second is the dearth of research in understanding variation. The third component is the inclusion of stochastics in school curricula, and the fourth is the importance of teachers' subject matter knowledge. Taken together, these four components build a case for why elementary preservice teachers' conceptions of variation are a relevant and significant area to investigate in the overall picture of mathematics education. The chapter culminates with a presentation of the objectives of the study and statement of the research questions.

Variation is a term with several related forms and uses. *The Oxford Dictionary of Current English* says that variation refers to a "departure from the normal kind, amount, a standard, etc" (Thompson, 1998, p. 1018). Related in Latin to other terms like *vary, variability,* and *various,* variation implies diversity. A light rail system may promote its trains as arriving at any given stop every ten minutes, but the actual time between arrivals varies. The interval of time is not uniformly ten minutes, and the absence of uniformity indicates the presence of variability. Variation can also refer to the amount of diversity. One way to measure variation in a set of numeric data is to compute the *range,* which gives the spread of the data from the maximum to the minimum. Another way of measuring variation is to calculate the *variance* and standard *deviation* of the data set, which quantify the spread of the data about the arithmetic mean. The key element to any discussion of variation is that there are differences in the set under study, because without differences there is no variation.

Primacy of Variation

Moore (1990), in a treatise on the nature of statistical thinking, lists five core elements:

1. The omnipresence of variation in processes,
2. The need for data about processes,
3. The design of data production with variation in mind,
4. The quantification of variation,
5. The explanation of variation (p. 135).

Notice that variation is mentioned in four of the five core elements listed by Moore. A report of the joint curriculum committee of the American Statistical Association (ASA) and the Mathematical Association of America (MAA) supports not only the omnipresence of variation as one of their elements of statistical thinking, but also the elements of "measuring and modeling variation" (p. 127). The "*omnipresence of variability*" was cited as giving rise to the very need for the discipline of statistics (Cobb & Moore, 1997, p. 801, italics in original).

The idea that variability is everywhere makes sense when thinking about the world in which we live. Not only do people and their environments vary, but even repeated measurements on the same person or thing can vary (Wild & Pfannkuch, 1999). Also, "natural variation appears in the heights, reading scores, or incomes of a group of people" (Moore, 1990, p. 98). There is also a chance variation component to our world. Moore (1990) points out that one use of probability instruction is to lead students to the understanding that chance variation, as opposed to deterministic causes, explains most outcomes in our world. He writes:

> It is perhaps surprising that patterns in careful measurements or in data on many individuals can be described by the same mathematics that describes the outcomes of chance devices. Experience with variation is a first step toward recognizing the connection between statistics and probability (p. 99).

Philosophically, living in a stochasticized world implies an existence beset by variation on all sides (Davis & Hersch, 1986); mathematically, "statistics provides means for dealing with data that take into account the omnipresence of variability" (Cobb & Moore, 1997, p. 801).

In addition to the more academic examples cited above, professional statisticians also see the centrality of variation in their work, and others have framed a model of statistical thinking in which variation is the core element (Pfannkuch, 1997; Pfannkuch & Wild, 1998; Wild & Pfannkuch, 1999; Pfannkuch & Wild, 2001). In an investigation of the nature of statistical thinking from the practitioner's perspective, the first component to emerge was that "statistical thinking involves 'noticing', understanding, critically evaluating and distinguishing the types of variation" (Pfannkuch, 1997, p. 407). The theme of variation is pervasive throughout the process of any statistical enquiry. Wild and Pfannkuch (1998) interviewed professional statisticians to capture the stories their subjects wanted to tell. One subject expressed that "basically what distinguishes statistical thinking from anything else is that you accept that variation exists," while another succinctly states that "statistics is the science of variation" (Wild & Pfannkuch, 1998, p. 6). The authors posit that "this very basic element of statistical thinking, 'noticing variation and wondering why', is actually at the root of much, if not most, scientific research" (p. 7). The mindset of *noticing variation and wondering why*, when coupled with the sense that variability inheres in every facet of life, makes for a rich context in which to embed all other elements of a statistical enquiry. The mandate not only to expect variation, notice variation, but also account for the causes of variation, make up key features of Wild and Pfannkuch's (1999) model of statistical thinking.

The above examples lend credence to the tenet that variation is indeed the central feature behind statistics, and offer support for why others agree that "statisticians consider variation to be the foundation of statistical thinking, the very reason for the existence of their discipline" (Shaughnessy & Ciancetta, 2001).

Understanding Variation: A Paucity of Research

Although variation is central to stochastics, there is relatively little research on people's understanding of this concept. This is not surprising, since as of the late 1980s a review of the literature showed that far more research had been done in the area of probability than in statistics (Garfield & Ahlgren, 1988). Several researchers had attempted to describe stages in the development of probabilistic thinking (Piaget, 1975; Falk, 1983; Fischbein & Gazit, 1983; Green, 1983). Others had focused on intuitive reasoning about probability, revealing not only the kinds of misjudgments people make but also suggesting explanations for these errors (Kahneman & Tversky, 1982; Konold, 1983). Only recently has the attention of researchers turned directly to concept of variation.

In delivering a keynote address to the Mathematics Education Research Group of Australasia (MERGA), Shaughnessy (1997) noted that "although there have been investigations into students' concepts and beliefs about 'averages', there does not seem to be a similar tradition of research into students' ideas about variability or spread" (p. 5). That address may well have served as a catalyst for researchers to uncover the specific ways in which people thought about variation in different contexts, or at the very least it gave a voice to what other researchers had been noticing as well.

Since that MERGA address, research specifically about conceptions of variation has been slowly emerging, amidst other calls for more research in this area. In a study about data distributions, Mellissinos (1999) comments that "student notions of variability is considered a needed area of research" (p. 1). Concepts of variation, graphicacy, and centers are all factors relating to the research of Watson and Moritz (1999), who remark that "very little research has explored children's strategies involved in comparing data sets" (p. 146). Similarly, Watson and Moritz (2000a) note that little research has been done on students' cognition of sampling situations, for which variation is an integral component. These researchers have consistently called for more research on the understanding of variability in the context of comparing data sets and sampling. The current situation is summarized nicely by Torok and Watson (2000), who wrote that "an appreciation of variation is central to statistical thinking, but very little research has focused directly on students' understanding of variation" (p. 147).

The research which does exist on the concept of variation was mostly aimed at students in grades 3-12. Reading and Shaughnessy (2001) used subjects from grades 4-12, and Watson et. al. (2002) used subjects from grades 3, 5, 7, and 9. However, while needed research has only recently been and is still being conducted on *school students'* understanding of variation, research has provided few exploratory results on the conceptions of variation held by *teachers*. As others have noted, research on understanding variation is still in its nascency (Watson, Kelly, Callingham, Shaughnessy, 2002; Torok & Watson, 2000; Jones et. al., 2000).

Curricular Inclusion of Stochastics: A Brief History

Stochastics has not always been a vital part of the school curriculum in the U.S. In fact, up until the last decade or so it could safely be said that calls for including stochastics in the American school curriculum had fallen on deaf ears. Despite the encouragement of the 1923 report *The Reorganization of Mathematics in Secondary Education* to include some stochastics among the usual fare of algebra, geometry, and trigonometry, subsequent papers such as the 1938 report *Mathematics in General Education* and the 1940 report *The Place of Mathematics in Secondary Education* placed less emphasis on the importance of stochastics (Bidwell & Clason, 1970). By the era of New Math, the secondary curriculum piloted by the University of Illinois in 1958 comprised eleven units, none of which specifically addressed stochastics (NCTM, 1970). It is therefore no wonder that in 1959, when the College Entrance Examination Board proposed curriculum including a unit introducing probability with statistical applications, they were able to applaud themselves for "one of the more novel suggestions of the Commission" (Bidwell & Clason, 1970, p. 703). Later, the 1963 Cambridge report *Goals for School Mathematics* emphasized an "elementary feeling for probability and statistics" (Cambridge Conference on School Mathematics, 1963, p. 9), but by the time this emphasis was reiterated in the 1975

NACOME report, the actual representation of stochastics in the curriculum was still meager (Gawronski & McLeod, 1980). This brief historical review helps bolster the claim of "the traditional complete absence of stochastics from the school curriculum" (Shaughnessy, 1992, p. 467), and lets us more fully appreciate this situation.

The National Council of Teachers of Mathematics called for increased teaching and learning in stochastics in their 1980 publication *An Agenda for Action*, but it was the NCTM *Standards* of 1989 which gave stature to the place of stochastics among other curricular strands in the United States. (NCTM, 1980; NCTM, 1989). This place has been affirmed in the subsequent release of the *Principles and Standards for School Mathematics* (NCTM, 2000). Also helping drive up interest in stochastics has been the 1997 addition of statistics to the list of Advanced Placement exams. The syllabus for this exam includes data exploration, study design, probability distributions through simulation, and inference. The growth of participation in this exam has been steep. For example, while about 7,600 high school students around the world took the first AP Statistics exam in 1997, that number had risen to over 65,000 by 2004.

However, the rise to prominence of stochastics in the school curriculum is not constrained to the United States alone. Not long before the release of the 1989 *Standards*, Garfield and Ahlgren (1988) cited promising new curricular materials being developed not only in America, but also in the United Kingdom. Furthermore, in reference to developing mathematically literate world citizens, these authors noted the “vigorously growing movement to introduce elements of statistics and probability into the secondary school curriculum, and even the elementary school curriculum” (p. 44). Others affirm the international trend, and point to examples in Spain, Australia, and New Zealand, in addition to America and the United Kingdom (Batanero, Godino, Valecillos, Green, & Holmes, 1994; Shaughnessy, Garfield, & Greer, 1996; Mellissinos, Ford, & McLeod, 1997; Watson & Moritz, 2000a). In England and Wales, for example, the recommendation for stochastics includes "collecting, representing, and interpreting data" (Department for Education, 1995, p. 10). The curriculum in New Zealand calls for "statistical investigations within a range of meaningful contexts" (Ministry of Education, 1992, p. 186). In reference to notions of sampling and making inferences, *A National Statement on Mathematics for Australian Schools* suggests that "the groundwork should be laid in the early years of schooling in the context of data handling and chance activities" (Australian Education Council, 1991, p. 64). It does indeed seem clear that "topics in data handling have begun to play a more prominent role in the mathematics curricula in many countries" (Shaughnessy et. al., 1996, p. 205).

Looking closer at the PSSM (NCTM, 2000), the following recommendations are made in its Data Analysis and Probability strand:

> Instructional programs from prekindergarten through grade 12 should enable all students to –
>
> - Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them;
> - Select and use appropriate statistical methods to analyze data;
> - Develop and evaluate inferences and predictions that are based on data;
> - Understand and apply basic concepts of probability (p.48).

It is worth noting that both aspects of stochastics – probability and statistics – are intertwined in the same curricular strand, and this illustrates the way in which concepts of chance are considered integral to a holistic perspective of data handling.

While some may suggest that data analysis take place in an environment wholly separate from probability, the position taken in this study and by the NCTM advocates the synthesis of concepts of data and chance in the school curriculum. Under this position, "in the context of data analysis, or statistics, probability can be thought of as the study of potential patterns in outcomes that have not yet been observed" (Scheaffer, 2002, p. 6). That the full meaning of data analysis should incorporate both aspects of statistics and probability is a perspective shared by many others as well (Jones, Thornton, Langrall, Mooney, Perry, & Putt, 2000; Torok & Watson, 2000; Shaughnessy et. al., 1996).

Importance of Teachers' Knowledge

Since stochastics continues to be emphasized in the school curricula, and since variation is a vital element of stochastics, it makes sense to wonder what teachers know about variation. Quinn (1997) identified stochastics as one of the three "problematic areas of preservice elementary education" (p. 112), along with geometry and rational numbers. Lajoie and Romberg (1998) agree that stochastics may be as new a topic for teachers as for children, and that "teachers must be provided with appropriate preservice and inservice training that will give them the knowledge base they need to feel comfortable teaching about data and chance" (p. xv).

Implicit in stressing the importance of teachers' knowledge is the belief that teachers themselves are important to the enterprise of learning. While this may seem intuitive, it depends on what we mean by "knowledge." The claim of the 1966 Coleman Report was that "teachers, or more accurately variations among teachers, do not make a difference in school achievement" (Shulman, 1988, p. 10). The process -

product research program showed that “teachers did make a difference” (p. 10), but the focus of that program was more on teachers’ classroom behavior, not on their knowledge. Much of the past research “showed little or no statistical relationship between teacher knowledge and student achievement” (Grossman, Wilson, & Shulman, 1989, p. 25). One reason for these results is that teachers’ knowledge was measured in very limited ways, by looking at the number of classes taken or a teacher’s grade point average. This impoverished view suggested that teacher knowledge wasn’t vital to student learning, even though this connection may appear to be a matter of common sense (Lockwood, 1998; Fennema & Franke, 1992). In more recent times, a report by the National Commission on Teaching and America’s Future cited growing research that “what teachers know and do is one of the most important influences in what students learn” (Darling-Hammond, 1998, p. 6).

Certainly there are many dimensions to the enterprise of teaching, such as teachers’ beliefs, knowledge, attitudes, skills, and classroom behavior (Shulman, 1988; Borko et. al., 1992). While teachers may make modifications to their practice along these dimensions throughout their careers, for many of them the initial experience of teacher training provides groundwork in these areas. Foremost among the goals for a teacher education program is that preservice teachers begin to gain the components of knowledge which research suggests are important for teaching (Cooney, 1994). Perhaps because teacher knowledge is an incredibly complex issue (Lehrer & Franke, 1992), it has been studied in terms of different types or components of knowledge. Two components which research has begun to delineate are subject matter knowledge and pedagogical content knowledge.

Shulman (1988) gave a definition of subject matter knowledge as “that comprehension of the subject appropriate to a content specialist in that domain” (p.26), a comprehension which includes “the key facts, concepts, principles, and explanatory framework of a discipline” (Borko et. al., 1992, p.195). The importance of subject matter knowledge is echoed in the *Professional Standards for Teaching Mathematics* (NCTM, 1991), which includes a section addressing the professional development of teachers. The professional development section, Standard 2 entitled “Knowing Mathematics and School Mathematics”, states that “teachers of mathematics should develop their knowledge of the content and discourse of mathematics, including mathematical concepts, procedures, and the connections among them...” (p. 132). This description seems very much in line with the idea that subject matter knowledge is about knowledge *of* mathematics and knowledge *about* mathematics (Simon, 1993). Standard 4, “Knowing Mathematics Pedagogy,” speaks about “teachers’ knowledge and ability to use and evaluate...ways to represent mathematics concepts and procedures...” (NCTM, 1991, p. 151), and this relates to pedagogical content knowledge. This idea of representing mathematics is mentioned by Shulman (1986), who includes in his definition of pedagogical content knowledge the use of analogies, examples, illustrations, and demonstrations, "in a word, the ways of representing and formulating the subject that make it comprehensible to others" (p. 9). Knowing how to accomplish this representation – how to select appropriate tasks, ask good questions, and assess what students understand – is seen as the heart of pedagogical content knowledge (McDiarmid, Ball, & Anderson, 1989; Borko et. al., 1992; Tirosh, 2000).

My study is primarily concerned with the subject matter knowledge of teachers regarding the topic of statistical variation. While pedagogical content knowledge is also important, the research on how children learn about variation is still sparse. Hence, it is difficult to know just how to help teachers work with their pupils when research is still painting the picture of how students think about this topic. It is also seems problematic to address pedagogical content knowledge when even the teachers' own subject matter knowledge is unclear. In the case of elementary preservice teachers, it is unknown what conceptions of variation they have.

Concerning the recommendations of the PSSM (NCTM, 2000) for the Data Analysis and Probability strand for grades 3-5, expectations include: The design of investigations; the consideration of the effects of different methods of data collection; the comparisons of data distributions; and the proposal and justification of predictions and conclusions based on data. Variation is an inherent concept within each of these expectations in probability and statistics, at the elementary level. The research on student thinking, presented in the next chapter, is beginning to inform us of what elementary students can understand about variation. The big question is: What about the subject matter knowledge of the prospective teachers of these schoolchildren? Do preservice teachers participate in experiences where they themselves can develop an understanding and appreciation of variation? If teachers are expected to gain some requisite knowledge at their training institutions, then the colleges and universities should be a place teachers can learn about variation. More than just traditional math courses are called for, and Simon (1993) suggests that " in order to break the cycle of teachers with weak conceptual backgrounds providing conceptually impoverished instruction, preservice mathematics courses will need to prepare prospective teachers more adequately" (p. 252).

Objectives of the Study

Within stochastics variation is the dominant characteristic of statistical thinking. While research on probability and statistics has produced findings in a number of areas which I'll bring up during the literature review in the next chapter, research specific to the concept of variation has only recently surfaced. Moreover, the research on variation which has come to light has predominantly been conducted with precollege students. Missing from the literature is an idea of how preservice or inservice elementary teachers reason about variation.

Thus, one objective of this study is to develop a framework to characterize the conceptions about variation held by elementary preservice teachers (EPSTs). A second objective is to compare EPTS' conceptions of variation before and after an instructional intervention focusing on variation. A third objective is to investigate types of tasks that might be useful to uncover EPSTs' thinking about variation. My research questions directly reflect these objectives of the study:

1. What are the components of a conceptual framework that help characterize EPSTs' thinking about variation?
2. How do EPSTs' conceptions of variation before an instructional intervention compare to those conceptions after the intervention?
3. What tasks are useful for examining EPSTs' conceptions of variation in the contexts of sampling, data & graphs, and probability?

The next four chapters show how this study unfolded to address the above research questions. Chapter Two includes a description of the previous research on conceptions of variation that has been done in the following three contexts: Variation in data sets, variation in sampling, and variation in chance situations. Also, Chapter Two discusses an initial conceptual framework that was inspired by prior research and a pilot study. Chapter Three provides the methodology used to gather and analyze the data, and also contains a detailed description of how the initial conceptual framework developed into a richer, evolving framework. The meaning of each element in the evolving framework is discussed as part of the results in Chapter Four. Also in Chapter Four, the evolving framework is used to compare EPSTs' conceptions on various tasks from both before and after the instructional interventions. A summary of results is given in Chapter Five, as are implications for teaching and recommendations for future research.