CHAPTER THREE

Theoretical and Conceptual Frameworks

Introduction

This chapter discusses both the theoretical and conceptual frameworks for my proposed study. First, in this section the distinction between these two types of frameworks is mentioned. Then, a theoretical framework for learning is described in terms of social constructivism. Finally, a conceptual framework for understanding variation is articulated.

Some in the mathematics education community talk about theoretical or conceptual frameworks as if they were interchangeable phrases, while others make a distinction (Eisenhart, 1991; Lester, 1991). Romberg (1992) cautions researchers against moving from the identification of an interesting problem straight to the design of a study without "situating their study with the work of others" (p. 56). One use for theoretical and conceptual frameworks is to help connect a study to previous research.

Eisenhart (1991) described a theoretical framework as "a structure that guides research by relying on a formal theory…constructed by using an established, coherent explanation of certain phenomena and relationships" (p. 205). She cites as an example Vygotsky's theory of socio-historical constructivism. Marshall and Rossman (1989) explain how a researcher will locate a research problem in a body of theory. The location is chosen on the basis of the researcher's own underlying assumptions, and these assumptions must be explicitly stated. Thus, theoretical frameworks can be expected to invoke a host of values and beliefs, not unique to the researcher, but shared in a common paradigm with other scholars. This is why researchers generally do not locate their work in some completely new theoretical framework, creating something that sounds compelling to them alone, but seek to identify the perspectives that align their work with other researchers.

A conceptual framework is described by Eisenhart (1991) as "a skeletal structure of justification, rather than a skeletal structure of explanation" (p. 209). This structure is based on either formal logic or experience. It consists of an argument which can incorporate differing points of view, and which culminates in the articulation of a rationale for the adoption of some ideas or concepts in favor of others. The chosen ideas or concepts serve to guide the data collection and analysis:

> "Crucially, a conceptual framework is an argument that the concepts chosen for investigation or interpretation, and any anticipated relationships among them, will be appropriate and useful, given the research problem under investigation" (p. 209).

Lester (1991) finds the distinction between justification and explanation useful for mathematics education, noting that the most persistent needs in research are to justify "why a particular question is proposed to be studied in a particular way and why certain factors (e.g., concepts, behaviors, attitudes, societal forces) are more important than others" (p. 195).

Thus, conceptual frameworks, like theoretical frameworks, also embody values and beliefs. The key difference is that theoretical frameworks are more rigidly tied to the body of theory they adhere to, while conceptual frameworks are built to be more flexible. While heavy reliance on formal theory alone may minimize potentially important information, (Becker, 1991), "conceptual frameworks facilitate more comprehensive ways of investigating a research problem" (Eisenhart, 1991, p. 211).

Theoretical Framework: Social Constructivism

The intent of this section is to be explicit about some assumptions I make in my study about the way that people learn and the nature of mathematical knowledge, and also to connect these assumptions to my choice of research questions and qualitative methodology. First, social constructivism is described as a complementary balance between two different theories of learning, one which focuses more on individual cognition and the other which focuses more on social practice. Then, an example is presented of social constructivism as it relates to the knowledge of randomness. Finally, the relationship of the theoretical framework to my research questions and methodology is discussed.

Brief Description of Social Constructivism

Cobb and Yackel (1996) discussed two theories of learning which represent different ends of a continuum. On one end is the constructivist perspective, and on the other end is the sociocultural perspective. I will discuss these perspectives and then consider social constructivism as a synthesis of the two theories.

Von Glasersfeld (1995) points to one of Jean Piaget's (1937) major works as introducing the notion of constructivism and making the term popular. Ernest says that "the term *constructivism* itself covers a panoply of theoretical positions" (Cobb, Yackel, & Wood, 1992, p. 3). Fox (1997) notes that

> "constructivism is a metaphor, and also an umbrella term, covering a family of theories of perception, memory, learning and teaching, which have in common a shared concern with human knowledge as the product of an active, constructive process" (p. 14).

Rooted in the work of Piaget, constructivism focuses on the individual building of knowledge, whereby "we construct our knowledge of our world from our perceptions and experiences, which are themselves mediated through our previous knowledge" (Simon, 1995, p. 114).

Piaget's contention was that "learning is subordinated to development" (1964, p. 184), and he posited four hierarchical stages of cognitive development. The transition through these stages depends on four factors: maturation, experience, social transmission, and equilibration. Piaget linked maturation to embryogenesis, which "concerns the development of the body, but it concerns as well the development of the nervous system and the development of mental functions" (Piaget, 1964, p. 176). Equilibration was described as a self-balancing of two intrinsic polar yet complementary processes or behaviors, assimilation and accommodation (Fosnot, 1996; Herscovics, 1996). Assimilation enables the learner to "fit new knowledge into his or her existing cognitive structure", while accommodation comprises the learner's "reorganization and expansion of such a cognitive structure" (Herscovics, 1996, p. 351). Mental structures develop via a cycle of disequibrilation, accommodation, assimilation, and equilibration within the individual learner (Janvier, 1996).

Piagetian constructivism, to use Janvier's (1996) term, is "seen as an individual cognitive activity that involves the internal reorganization of mental schema" (Teppo, 1997, p. 3). Others call this merely constructivism, but concerned that some people were missing the epistemological implications behind Piaget's ideas, von Glasersfeld (1998) was prompted to "add the qualifier *radical* to constructivism" (p. 23). The epistemological implications of radical constructivism, according to von Glasersfeld (1995) are:

> "1
> - Knowledge is not passively received either through the senses or by way of communication;
> - knowledge is actively built up by the cognizing subject.
>
> 2
> - The function of cognition is adaptive, in the biological sense of the term, tending towards fit or viability;
> - cognition serves the subject's organization of the experiential world, not the discovery of an objective ontological reality" (p. 51).

As Simon (1995) wrote, "the radical constructivist position focuses on the individual's construction, thus taking a cognitive or psychological perspective" (p.16).

On the other end of the theoretical continuum is the sociocultural perspective of learning, which "…views learning as the enculturation of an individual into a community of practice, and the focus of inquiry is placed on the individual's participation in social practice" (Teppo, 1997, p. 3). The sociocultural perspective owes much to the work of Vygotsky, whose emphasis was on language and the social world (Cobb, 1994; Vygotsky, 1986). Vygotsky claimed that cognitive development resulted from social interaction and education, by means of language. "Knowledge is socially constructed. Vygotsky emphasizes the part played by social activity and cultural practices as sources of thinking" (Harvard, 1997, p. 40). A central idea expressed in Vygotsky's writings is a person's increasing capacity for self-regulation. His notion of the zone of proximal development was used to describe "the place at which a child's empirically rich but disorganized spontaneous concepts 'meet' the systematicity and logic of adult reasoning" (Kozulin, 1986, p. xxxv). Learners realize their capacity for self-regulation by being guided in their participation by others who are more capable, resulting in a level of performance which the learners would otherwise have been unable to reach on their own (Harvard, 1997). Cobb (2000) notes that the sociocultural perspective tends to “elevate the social process above the psychological” (p. 309).

Synthesizing the constructivist and sociocultural perspectives, Cobb, Yackel and Wood (1992) invoked the label *social constructivism* to emphasize both the role of the individual and the social environment in learning. In their opinion,

> “the question of whether mathematics is essentially cognitive or whether it is essentially sociological or cultural in nature is considered to be irrelevant. We have instead suggested that it is useful to see mathematics as both a cognitive activity constrained by social and cultural processes and as a social and cultural phenomenon that is constituted by a community of actively cognizing individuals” (p. 32).

Referring to these authors’ work, Teppo wrote that “their *social constructivist* or ‘emergent’ framework” implies that the development of individual and the development of social meaning cannot exist independently of each other (1997, pp. 3-4). What this means is that mathematical learning is both an individual process and as a process influenced by social mediation. Based on my experience in the classroom as both a teacher and a student, I agree with Cobb (1994) that “mathematical learning should be viewed as both a process of active individual construction and a process of enculturation into the mathematical practices of wider society” (p. 13).

Social Constructivism and Mathematical Knowledge

Some people view all of mathematics as a set of absolute truths. Learning is therefore a matter of how well a student knows or does not know these truths. If all mathematical knowledge is a body of infallible truths, then the issue of a student personally or socially constructing their own understanding becomes problematic. What if the student constructs the wrong knowledge? A social constructivist perspective towards learning implies that knowledge is not a fixed, universal entity, but something that is negotiated publicly (Ernest, 1991). To give an example of mathematical knowledge that embodies the epistemological stance of social constructivism, consider the notion of randomness. As will be discussed, randomness is not a rigidly established body of mathematical knowledge.

Not only is randomness “a complicated concept that is describable in many ways” (May, 1997, p. 222), but there remains a “fundamental question of the real nature of randomness” (Rial, 1998, p. 482). In fact, the exact wording “What is random?” is the title of more than one article, at least one book, and the question is addressed by several other authors as well (Kac, 1983; May, 1997; Beltrami, 1999; Pagels, 1982; Kolata, 1986; Bennett, 1998). Knowledge of what is random can seem tenuous and is very much dependent on context. Some authors advocate leaving the notion of randomness to the realm of the undefined, much the same way that the notions of point or line are left undefined in geometry. Pagels (1982) commented that “mathematicians have never succeeded in giving a precise definition of randomness” (p. 85) and went on to say, “Mathematicians don’t know what random is!” (p. 87).

While a concise definition seems intractable, aspects of the concept such as uncertainty, unpredictability, and lack of local pattern or causes do provide useful lenses with which to describe randomness. For instance, in considering the randomness of a coin toss, Ford (1983) claims that “coin tosses are universally presumed completely random despite their obvious underlying determinism” (p. 40). His opinion is that the laws of physics should be able to discern a priori what side will face upward. For Ford, coin tosses are not completely random. As another example, consider which sequence looks more random: Sequence A (01101100) or Sequence B (01010101). Seife (1997) claims that any mathematician could tell you that Sequence A is more random, “but none could tell you just how much more random” (p. 532). In fact there are recent tests based on information theory which have been designed to measure randomness, and Seife (1997) points to conclusions that let theorists claim that $\sqrt{2}$ is more random than **e**, though both are less random than π. However, Batanero and Serrano (1999) admit that “in a test of randomness there is always a small probability that the sequence is not random in spite of having passed the randomness tests” (pp. 560-561). Whether or not finite sequences are in fact random can never be definitively answered (Steinbring, 1990), a perspective which leads Kac (1983) to state that, from an operational point of view, “the concept of randomness is so elusive as to cease to be viable” (p. 406). Beltrami (1999) suggests that “in response to the persistent question ‘Is it random?’ the answer must now be ‘probably, but I’m not sure’” (p. 109).

The point of the above examples is that determining what counts as random proves to be more of a human, social construct and less of an absolute, objective knowledge. Individuals build on their own personal conceptions of randomness by interacting with others in a process of social negotiation to arrive at some consensus of what they mean by the concept. Social constructivism allows the meaning of randomness to change over time as people change. Neyland (1995) writes that

> “Mathematics concepts and proofs can never be regarded as beyond revision and correction; they may require renegotiation as standards of rigour change or new meanings emerge…Mathematics is a dialogue between people exploring mathematical problems, and it must be viewed in its historical and social context” (p. 143).

What we know as random as today, we may not know as random tomorrow (Wild & Pfannkuch, 1999). If there ever was a mathematical concept which lends itself to a social constructivist perspective of learning, surely that concept is randomness.

Relationship of Social Constructivism to Research

As a theoretical framework, social constructivism has two main consequences for my study. First, the nature of my research questions needs to be faithful to the underlying tenets of social constructivism. Secondly, my methodology for answering those questions also needs to reflect the social constructivist perspective. These two consequences are next discussed.

The nature of my research questions allows for multiple perspectives of variation to emerge. If there were only one right way to look at or understand variation, then an appropriate research question might be, “Do the subjects have the correct view of variation?” A test could be constructed to find out whether or not subjects had the correct knowledge, but such an approach violates the epistemological tenets of social constructivism. Variation is like randomness in the sense that it exists in many forms, and it can be described in many ways. Even for the concept like the standard deviation of a sample, although statisticians know that the unbiased estimator is obtained via a divisor of (n-1) rather than (n), Moore (1991) observed that "There are arguments in favor of both n and n-1. Since the difference is quite small when n is of moderate size, we use n for simplicity" (p. 222). Also, there is much more to variation than just the standard deviation (e.g. Torok & Watson, 2000), and thus the intent of this research is to find out *what* EPSTs understand, not to find out whether or not they have a purportedly "correct" knowledge.

My research questions fit the descriptive intent of the study. The two principle questions align with a social constructivist perspective because they aim to uncover the knowledge that the subjects themselves hold. The five supporting questions refer to explicit aspects of understanding variation which are hypothesized as useful for examining conceptions of variation. These five aspects are discussed later in this chapter as a part of my conceptual framework. On the one hand, my study does look at the knowledge that individual EPSTs hold, but on the other hand the methodology provides contexts for learning that are explicitly social. The subjects will participate in classroom activities that require them to work together in small groups, discussing their ideas and responding to questions. By capturing the discussion of the groups of subjects, I am attending to the social component of knowledge construction. I want to find out what the subjects say to each other in the classroom setting as they work together, which is why the small group discussion will be recorded. I also want to find out what the subjects say on their own, which is why individual written responses will be collected, and why individual interviews will be conducted.

Qualitative methods offer ways to capture rich detail into the thinking of subjects, and are particularly apropos of social constructivism. Ernest (1997) wrote that "the emergence of constructivism in research in mathematics education has foregrounded a new set of research emphases that are central to the qualitative research paradigm" (p. 31). Taking a social constructivist stance towards knowledge and learning leads to methods allowing for an investigation of what knowledge is held by the subjects and how that knowledge is learned, an investigation which includes "or subjects' perceptions, purposes, premises, and ways of working things out" (Noddings, 1990, p. 14).

Patton (2001) offers a useful summary of the themes of qualitative inquiry, and he organizes these themes into three categories: design strategies, data collection and fieldwork strategies, and analysis strategies. Within the first category, he mentions that the "cases for study…are selected because they are 'information rich' and illuminative." Regarding data collection, qualitative data such as observations, interviews, and careful document review can lead to "detailed, thick description" regarding "people's personal perspectives" (Patton, 2001, p. 40). Lastly, in terms of data analysis, qualitative methods allow for "immersion in the details and specific of the data to discover important patterns, themes and interrelationships," and the analysis may end "with a creative synthesis" (p. 41). The themes of qualitative inquiry fit with the social constructivist framework for my study.

The remainder of this chapter articulates the conceptual framework for understanding variation.

Conceptual Framework: Understanding Variation

Just as variation is at the heart of a statistical investigation, so too is the understanding of variation at the heart of this study. In looking at elementary preservice teachers' conceptions, it is helpful to have a conceptual framework which helps organize the aspects of an understanding of variation. Eisenhart (1991) gave a useful description of the purpose of conceptual frameworks in guiding research, and Miles and Huberman (1994) support Eisenhart's view by saying that

> "A conceptual framework explains, either graphically or in narrative form, the main things to be studied – the key factors, constructs, or variables – and the presumed relationships among them. Frameworks

> can be rudimentary or elaborate, theory-driven or commonsensical, descriptive or casual" (p. 18).

The conceptual framework for this study will be a synthesis of the ideas promoted by other studies that have looked at variation along with ideas based on my own experiences studying and teaching stochastics. There are two themes in the framework. The first theme pertains to five different aspects of understanding variation: expecting, noticing, displaying, describing, and interpreting. The second theme has to do with three contexts in which variation can be exposed and studied: data sets, sampling, and probability situations.

Aspects of Understanding Variation

The five aspects of understanding variation are not suggested as the only correct way to look at people's conceptions. Instead, these aspects make up a working hypothesis of themes that may emerge from student responses. The function of these aspects is to offer a potential lens through which the EPSTs conceptions of variation can be viewed. In addition to describing these aspects, some hypothesized interactions between them are suggested.

<u>Expecting Variation:</u> The expectation of variation is suggested as a way for students to demonstrate their intuitive, experientially or mathematically based reasoning in situations for which variation is inherent.

For example, when drawing repeated samples of the same size from a population, it is useful to think ahead of time about the kinds of variation that are reasonable to expect. Particularly illuminating are the ends of the spectrum. On one end, a person may expect their samples to look very much alike due to a small amount of variation from sample to sample. On the other end, a person may expect huge disparities between samples owing to large amounts of expected variation. In a pilot for this study, a class of elementary preservice teachers investigated sampling variation using bags of M&Ms. I brought in enough bags of M&Ms so that every pair of students could share one bag. These bags were the regular 1.69-ounce kinds which are usually sold individually in stores, and which tend to have between 50 and 60 pieces of candy in a mix of colors. The question posed to the students was, “What’s inside a typical bag of M&Ms, both in terms of total candies and in terms of color distribution?” Some students, who had no knowledge of the nature of the way these candies are packaged, anticipated no variation. They felt the bags should not only have the same number of candies, but the same color distribution. Other students expected that the totals and color distributions in each bag would be different, but were unsure of the amount of variation.

As another example, consider chance situation whereby the data is a compilation of the outcomes of repeated events. The expectation of variation can be tied to the research on probabilistic thinking. Reliance on the outcome approach or on proportional reasoning could result in minimal attention to variation because subjects might focus on a single number to represent their best guess. A spinner which is three-fourths black and only one-fourth white might be expected to produce mostly black outcomes, but what variation would students expect? That is, in 20 spins would students expect all the outcomes to be black?

Expecting variation to occur is fundamental to an overall understanding of the concept, and it is one aspect that will be researched in my study. Before conducting any experiments or attempting a task, and prior to looking at data relevant to the situation, it should be asked, “What variation is expected?”

<u>Noticing Variation:</u> Expecting variation is considered before experiments are conducted and before data is examined. Noticing variation occurs after data becomes available. In looking at the data, do the subjects notice the variation which is present?

Again the M&M investigation offers a good example. Students each prepared a bar chart showing the color distribution of their bag (as a percentage of the total candies in the bag), and all the bar charts were lined up on the chalkboard. The distributions were quite varied. For example, some bags had 20 % red candies and some had 35 % red. Some students didn’t seem to notice the variation among bags, and instead seemed intent on trying to find out which color was most dominant across all bags. Other students clearly indicated they could see that the percentages were jumping up and down all across the chalkboard, and that the bags were quite different from one another.

In another example taken from a pilot for this study, students were asked to gather data on the amount of money in coins that we had each brought to class. Only two people out of thirty-two total had exactly the same amount of money, while the range went from zero to over ten dollars in change. As students looked at the results, I was curious to see what comments would emerge regarding the spread of the data. I taught two sections of the course, and later in the week I was able to bring together the results of both sections. Some students did seem to notice not just that the means of the two data sets were different, but also that the spreads of the data from the means were also markedly different.

Noticing variation is an important aspect for understanding variation, and I want to find out if students are aware of the variation which is present in data. Is the variation apparent to students, or do they need to have it pointed out to them? Do they lack the language to talk about variation or do they simply not notice it? I conjecture a possible interdependence between the aspects of expecting and noticing. If students expect variation to occur, then perhaps they would be apt to notice variation that is or is not present in data. Even if students have no initial sense of expectation, it could happen that the phenomenon of actual variation is so striking that they may notice the variation that occurs. Once data is shown, students will have a chance to comment on any variation that they notice.

Displaying Variation: After raw data is gathered or presented, there are different ways of reducing the data. Some graphical displays obscure variation among data more than others. I want to know if students can discriminate between graphs which highlight and graphs which minimize variation.

For example, boxplots do a good job of showing the range and the interquartile range, but do not show variation of the data within the quartiles. I showed a class the results from the M&M investigation, and pointed out a boxplot representing the percentages obtained for red candies in each bag. I also graphed the same data in a dot plot. The two graphs are shown below in Figure 10.

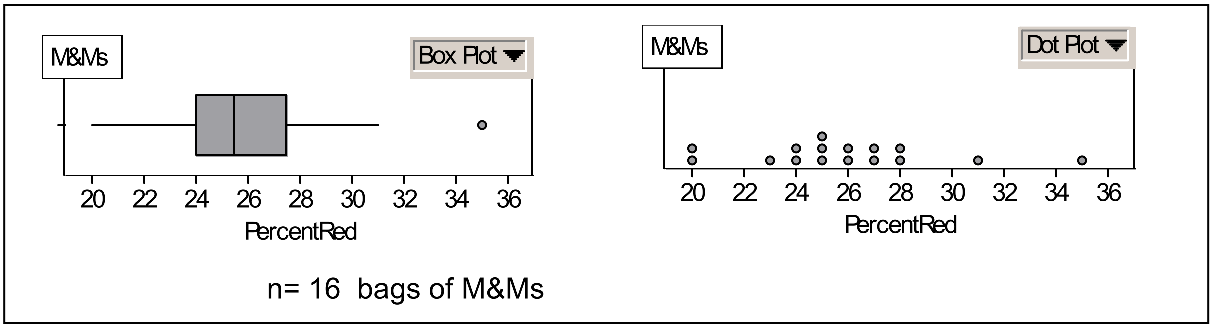

Fig. 10

The graphs shown in Figure 10 represent actual data, and make for a good discussion about what gets obscured or emphasized in different types of presentation. First of all, the boxplot readily shows the median (25.5) and identifies the middle 50% of the data as falling between 24 and 27.5, neither of which is as quickly obtainable from the dot plot. That is, in the dot plot some counting and calculation must be done to figure out where the quartiles are. Secondly, the mode is apparent in the dot plot, and the mean can be calculated from information in the graph, but but both mode and mean are not obtainable from the boxplot. Thirdly, the range is quickly seen in both graphs. Lastly, the actual distribution of the data – the way values are spread out and vary from one another – is obscured in the boxplot but not in the dot plot. One cannot tell from the boxplot where the gaps are (such as no values of 21, 22, 29, etcetera), nor can one tell the frequency for each value.

Histograms can obscure variation, depending on how the data is grouped. In graphing how much change students had brought to class, I asked groups of students decide for themselves how to graph the data. Most groups chose histograms, but the groups used different interval widths. Some groups used intervals from $1.00-$1.25, and other used intervals from $1.00-$1.50. The graphs (see Fig. 11) told different stories about the variation present in the data, and some students commented on these differences.

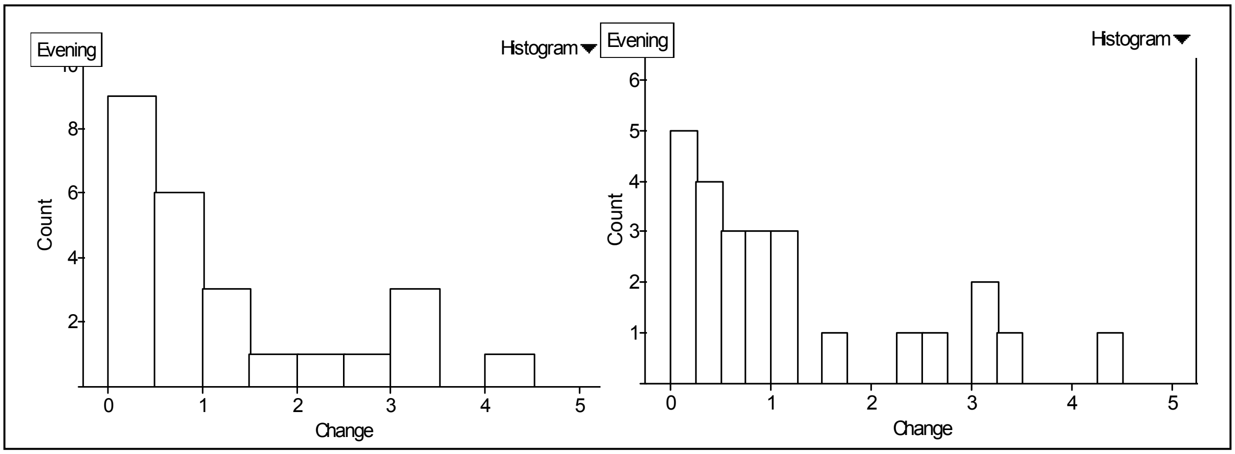

Fig. 11

For example, the larger interval widths disguise the fact that no students had between $1.75 and $2.25 in change.

In the pilot for this study, I showed students an assortment of computer-generated graphs (similar to Figure 11) for the same set of contrived data. The students were able to articulate the differences they saw in the way the variation was displayed. I also asked students which of the graphs I had presented would they choose to display data in order to emphasize the variation they saw in the data. The reason I showed them the graphs as opposed to having them make their own graphs is because I wanted the students to concentrate more on what the graphs were telling them as opposed to wrestling with the challenges of creating accurate graphs.

Displaying variation is closely tied to noticing variation. Research on graphicacy and distributions suggests that information about centers and spreads is provided by graphs, but different graphs provide different degrees of information.

<u>Describing Variation:</u> For this aspect of understanding variation students' descriptions might be either quantitative or qualitative.

Quantitative descriptions of variation that I have seen students use included measures of range, interquartile range, and standard deviation. For example, when comparing two contrived data sets with 25 test scores in each set (see Figure 12), some students commented that the sets had identical ranges. The two data sets also had identical means, medians and interquartile ranges. Although the standard deviations (which are different) were provided to students, not many students commented on this statistic.

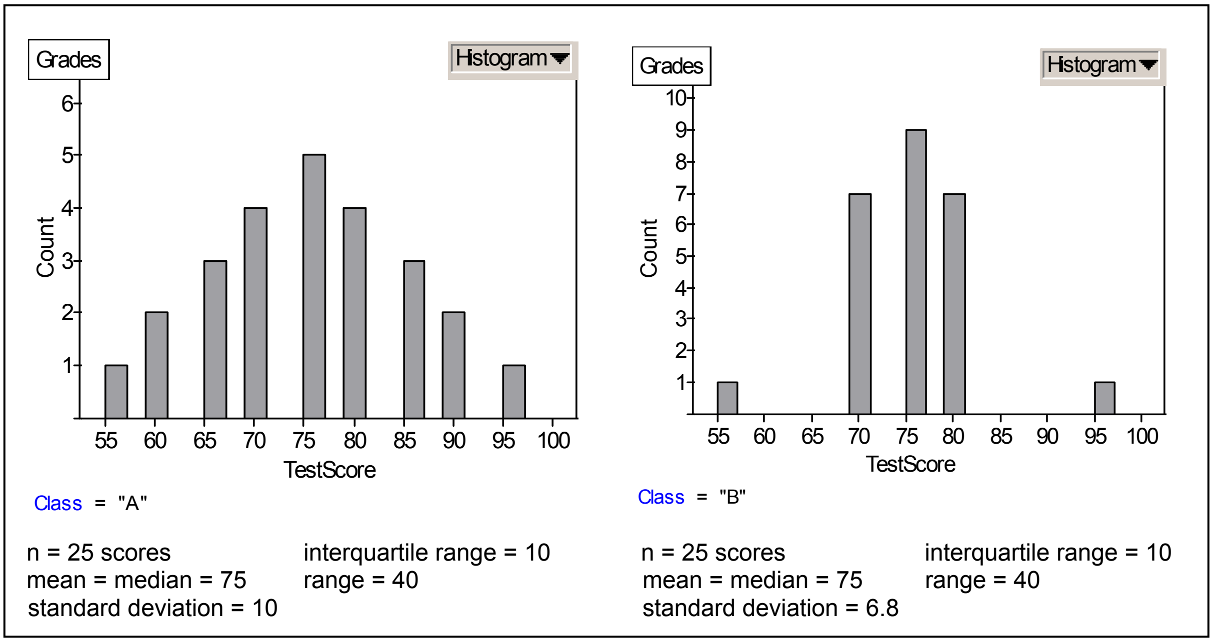

Fig. 12

Research on EPSTs conceptions of variation can be meaningful independent of the subjects' understanding of the standard deviation. Torok and Watson (2000) conducted a study of conceptions of variation using sixteen students (four each from grades 4, 6, 8, and 10). The students responded to questions about sampling and weather data in terms of what they thought was reasonable. Torok and Watson comment that "this study successfully explored students' understanding of variation without ever employing the phrase 'standard deviation'" (200, p. 166).

Qualitative descriptions involve attend to the language that students use as they talk or write about the data. For example, students may say that the data for Class A in Figure 11 is spread more evenly, or that the data for Class B is clumped together near the center. Other researchers have assigned qualitative descriptors to classify student responses about variation, such as the NARROW-REASONABLE-HIGH coding scheme (Shaughnessy et. al., 1999; Reading & Shaughnessy, 2000; Shaughnessy & Ciancetta, 2001).

It seems reasonable to hypothesize interdependence between whether students notice variation, how the variation is displayed, and the way the variation is described. It may be that people will engage in the process of noticing variation, followed by a description of what they are attending to in the graphical display.

Interpreting Variation: Two themes are grouped together in this aspect, the causes and the effects of variation.

By asking students about the causes of variation, students conjecture and reason about the source of the variation. For example, in the M&M investigation the students discussed not only why the bags were so different from one another, but also why so few of the bags were representative of the true color proportions that the company claimed to produce. Is there variation among the samples because the company only cares about the weight of the package and not the color mix of each bag? Is there more variation when looking at “fun-size” bags (n= about 24) than “regular” bags (n= about 55) because of the smaller sample size? When conducting a statistical investigation, and in looking at data, it seems useful to wonder about where the variation is coming from and why it is present (Wild & Pfannkuch, 1999).

Another interesting line of questioning about causes of variation came from a probability experiment in the pilot to this study. Students repeatedly threw two dice and computed the sums of the numbers facing upward. Theoretically, the probability of obtaining a sum of seven is 1/6. Different groups of students threw dice many times (between 30 and 60), but for each group the percentage of their total throws which yielded a sum of seven strayed either above or below the expected 16.7 %. Students offered reasons for why the experimental percentages did not match the theoretically expected percentage, which included the idea that the dice were uncontrollable, or that luck was not on their side. There was also some class awareness that the variation in this experiment was due to randomness.

Thinking about the causes of variation can lead to questions about the effects of variation. After doing probability experiments with my class, discussion turned to what the class would predict if some new student joined the class and threw the dice 30 times. Although the probability of obtaining a sum of seven is 1/6, many students argued that 30 throws would not yield exactly 5 sums of seven. Questions of confidence, such as asking "How confident are you that 30 throws would have at least 3 sums of seven?", are questions that probe students' thinking about the effects of variation. If there is a considerable amount of variation, as there can be when looking at "fun-size" bags of M&Ms, then extreme results (such having no Red candies in the bag) seem less surprising.

What then does variation imply about the data? How confident can one be in the trustworthiness of data when the variation is not what was expected? Some tasks used in previous research on variation have asked students whether the purported results were actually obtained from doing the experiment or simply fabricated (Watson, Kelly, Callingham, & Shaughnessy, 2002). It would be interesting to hear elementary preservice teachers' reactions to similar tasks, and find how they interpret variation that seems unreasonable. The interpretation of variation may be related to the expectation of variation. That is, perhaps a person's expectation influences their interpretation and vice versa.

Contexts for Understanding Variation

The five aspects of variation described above are offered as potential areas in which an understanding of variation can be organized. This study will consider three contexts for looking at variation. The contexts are based on the kinds of tasks that other researchers used for the studies discussed in the literature review, including variation in data sets, variation in sampling, and variation in chance situations. Some examples from the pilot to this study will clarify the distinctions between these contexts.

<u>Variation in Data Sets:</u> In this context, data can be gathered or provided without

explicit ties to either sampling or probability situations. The focus is more on where the data came from, what it means, and how best to describe the data. For example, when the class investigated how much money in coins we each had brought, the initial point of the exercise was to gather and display data. We did similar exercises by gathering data on body measurements, such as arm span and wrist circumferences. Once the data were available, then we displayed it and discussed notions of center and spread.

Later in class, instead of gathering data, I provided data sets which were either real or contrived, with the purpose of exploring students' interpretation of the data. Data similar to that in Figure 11 above was used to talk about whether one class did better on a test than the other class. In addition to comparing different data sets, I also want to explore different displays for the same data set, using displays which mask or highlight the variability in the data.

<u>Variation in Sampling:</u> The M&M investigation is a good example of how samples vary in the way they reflect the population from which they were drawn. Opening many "fun-size" bags shows the effect of taking repeated samples of approximately the same size. Even though the population of M&M candies is reported to be 20% red, a sample of size ten may not have exactly two red. The proportion of reds in the sample may vary widely, but increasing the sample size reduces this variation. For example, ten samples of size twenty have red percentages that vary more than ten samples of size two thousand.

For students who have not done the M&M investigation in a previous class, the population color proportions are unknown. Thus, the sample data is the only information through which an inference about the population can be made. By aggregating data from the individual bags, the effect of an increased sample size on variation between samples can be seen. For instance, four samples of 100 M&Ms usually look more similar to each other than sixteen samples of 25 M&Ms. An example using actual class data is presented in the next chapter.

<u>Variation in Chance Situations:</u> The probability experiment involving the sum of two dice is an example of this context. The data gathered is completely governed by chance. With the dice

experiment, many students knew that the chance of obtaining a sum of seven is 1/6. What is interesting in this context is to provide opportunities for students to move away from answers that rely on an expected value for one outcome, and to move towards the anticipated results of many outcomes in which random variation is sure to play a part. How extreme would the data from tossing the dice have to be before students questioned the fairness of the dice or the legitimacy of the data?

Another in-class probability experiment was done for which none of the students knew the theoretical probability. The task was to write down how many spins of a five-spinner (a disk partitioned into five sectors of equal area) it took until the pointer landed on each of the five numbered sectors at least once. Each student repeated the experiment fifteen times. Since they did not know the expected value for the number of spins, students could only use their experimental data in making a prediction. The outcomes of only fifteen experimental trials tend to be quite varied.

Conclusion

The theoretical framework I have adopted for my study holds that knowledge is something personally constructed by the learner and mediated by social interaction. There are different aspects of variation, just as there are different aspects of other mathematical topics such as randomness. The theoretical framework of social constructivism leads to research questions which ask for a description of EPSTs conceptions of variation, rather than questions about the correctness of their knowledge of variation. The framework also suggests that qualitative methods are appropriate for investigating my research questions. The specific choice of research design, which involves case studies, is discussed in the next chapter.

My conceptual framework offers a lens for viewing EPSTs conceptions of variation. The lens comprises the five aspects of understanding variation: expecting, noticing, displaying, describing, and interpreting. Three contexts for exploring variation, chosen because of their roots in previous research and because of my pilot experience with these contexts, are variation in sampling, variation in chance situations, and variation in data sets. The conceptual framework involving the five aspects is hypothesized as useful for describing variation. It is a deliberately flexible framework in the sense that it allows for a variety of types of responses about variation, yet adheres to the ideas generated in previous research. Some interactions are hypothesized among the five aspects of understanding variation, and other aspects different from these five may emerge from the data as well.