CHAPTER FOUR

<u>Methodology</u>

This chapter discusses the procedures for gathering and analyzing the data. The first section tells *who* the data was collected from, providing a more general description of the research design as well a rationale for the chosen methodology. The second section offers more detail about the data gathering, including *what* data was collected and *how*. The third section discusses the way the data was analyzed, including the role of grounded theory and the use of computer software.

Research Design

Subjects were chosen from a section of Portland State University's Math 212 course, which is entitled *Foundations of Elementary Mathematics II.* Intended for prospective teachers, the math course is required for those wanting to enroll in the Graduate Teacher Education Program (GTEP). Completion of GTEP leads to issuance of an Initial Teaching License. To give a sense of the environment common to the subjects, a general overview of the content and pedagogy for the typical Math 212 course is next described. Within the context of the typical Math 212 course, specific features of the particular section used for my research are articulated.

Content and Pedagogy of Math 212

The content for Math 212 includes geometry, probability, and statistics. Instructors usually divide the ten-week course into two parts of roughly equal length, with one part dealing with geometry and the other part dealing with stochastics. The probability component includes single-stage and multistage experiments, and students investigate both theoretical and experimental probabilities. Theoretical probabilities are obtained by deriving the sample space and considering the possible outcomes of a specific event, while experimental probabilities are obtained by simulation. Students consider disjoint events as well as independent and dependent events. The statistics component includes descriptive measures of central tendency (mean, median, and mode) and spread (range, interquartile range, and standard deviation). Students also analyze data using a variety of graphs, such as boxplots, line plots, bar charts, histograms, pie graphs, and scatterplots. Themes of sampling, such as random sampling, stratified sampling, and making predictions based on sampling, are also a part of the statistics component.

The pedagogy for Math 212 varies with the instructor, but there are some common pedagogical themes. One theme is that students participate in activities, both as a class and in smaller groups. The Math 212 classroom is arranged so that students sit in groups, which can be as large as six students per group. Once a problem has been posed, or an activity given, students typically will work singly, then together in their groups, and finally share ideas with each other in a class-wide discussion. The sharing of ideas points to a second common theme, which is communication. Students are expected to communicate what they are thinking about and to ask questions of one another so as to understand each other's reasoning. A third common theme is that the teacher acts more as an inquiring guide than as a lecturer. Ideally, the teacher facilitates discussion by asking questions, encouraging other students to ask questions, and generally guiding the class to consensus where possible.

Multiple sections of Math 212 are typically offered during any given quarter. During the Winter Quarter when my research took place, classes were held two days a week for ten weeks. Each class session was scheduled to last two hours and fifty minutes. The section I chose to use for my research was led by Steve, an experienced instructor for Math 211, 212, and other courses for teachers and preservice teachers of mathematics. Steve's plan for the curriculum reflected the components described earlier, but he modified his previous practice of doing all the geometry in the first part of the course, and all the stochastics in the second part. Instead, he devoted the first four weeks to geometry, and then gave the next four weeks to stochastics, followed by a last installment of geometry. The modified schedule was done to help accommodate my research plan.

One reason I chose Steve's section is because of Steve's skill in modeling the pedagogical themes mentioned earlier. He lectures less and he poses problems, guides activities, and facilitates discussions. Precisely because his teaching style supports a social constructivist learning environment as discussed in the previous chapter, Steve's section was an appropriate choice for doing my research. A second reason for choosing his section is because Steve and I had worked well together as colleagues for more than four years, and our philosophies of teaching and learning were similar. The familiarity with each other's styles and congruence of philosophies helped make working together easy, particularly when we were co-teaching some of the lessons. A third reason is because Steve was willing to modify the sequence of geometry and probability and statistics so that my out-of-class interviews could take place before and after the stochastics portion of the class sessions.

Student Characteristics

The urban setting of Portland State University fosters a wide range of student backgrounds, and it is difficult to describe typical demographic characteristics of the students who take Math 212. Past students in my classes have ranged in age from the early twenties to the middle fifties. Some are undergraduates and some are graduate students. Some want to become teachers and occasionally some are just filling a university math requirement.

The majority of Math 212 students have taken Math 211 at PSU, a course taught in a similar style to Math 212. The content of Math 211 includes whole-number arithmetic, number theory, fractions, decimals, and ratios. Other than Math 211, most students have taken few if any post-secondary math classes. It may have been as many as twenty or more years since they have had *any* mathematics class. At the outset of Math 211 students often write a "mathography," which entails a description of their past math classes and their feelings about past math experiences. Most students describe themselves as not having been very good at math in the past, and most Math 212 students share negative memories of precollege mathematics. Some students say they feared or hated their math classes, and others say they were bored. The attitude of most beginning Math 211 students is that mathematics is a rule-oriented discipline. The role of the teacher is to reveal the rules, and the role of the students is to memorize and apply the rules.

The Math 211 experience helps to give most students a new and different vision of what “doing mathematics” entails. Because the pedagogy in Math 211 is similar to that in Math 212, most students come to expect a learning environment in Math 212 where they will be active participants and where their thinking strategies are validated. They have been enculturated to the process of problem-solving and communicating their reasoning. Although not all Math 212 students are comfortable with the learning environment, those who have come through Math 211 at least know what to expect.

A total of thirty students completed Steve’s section of the Math 212 course. A profile of the background for these students is offered with respect to the following attributes: Class level at the university, gender, when and where the students had taken Math 211, and any prior probability or statistics courses taken. These attributes are summarized in Table 1 on the next page. The student total in Table 1 is not always thirty because not all data could be gathered for every student. For example, information on the first three attributes was gleaned from the university’s information system, and the last two attributes were informed by the PreSurvey which was completed by 27 students during the first week of the course. Appendix B shows the entire PreSurvey instrument.

Class Profile			
Attribute	**Categories of Attribute**	**Number of Students**	**Total Students**
[1] Class Level at the University	Undergraduate	9	30
	Graduate	21	
[2] Gender of the Student	Male	6	30
	Female	24	
[3] Where Math 211 was Taken	At PSU	26	30
	Not at PSU	4	
[4] When Math 211 was Taken	Within Last Year	24	27
	More Than a Year Ago	3	
[5] Any Prior Probability or Statistics Courses	Yes	12	27
	No or Unsure	15	

Table 1

Almost two-thirds of the class (19 out of 30 students) were continuing on with the same instructor, Steve, for consecutive quarters. This fact is noteworthy because it helps explain the general disposition of the class with respect to their attitudes about studying probability and statistics. Of the twelve students who could recall having had prior formal education in probability or statistics, eight expressed favourable attitudes on the PreSurvey when asked "How did you feel about Probability and Statistics at that time?" For example, one student wrote that it had been an "interesting class, while another commented that it had been "fairly easy to understand." The four other students with prior formal exposure expressed unfavourable attitudes, such as: "It was my least favorite class in all of college."

When asked how they felt about learning probability and statistics now, only 5 of the 27 respondents expressed explicitly negative thoughts, such as LT, who wrote: "I feel very scared. It makes me nervous." Sixteen students put something that was explicitly positive, such as

DM: “Very excited, looking forward to it.”
CS: “Pretty comfortable, ready and excited”
EM: “I’m interested to learn more”

The remaining 6 students had responses that were somewhat neutral in character, as the comments by GP illustrate: “I’m open to it, but not really excited.” It seems reasonable to assume that, since 19 students did have Steve during the previous quarter, there would be some influence on the students’ attitudes and expectations. As one such student, MM, put it, “I’m excited about this class because I enjoyed 211.”

Thus, while Steve’s class was diverse with respect to age and background, there was certainly a core of students that had a common experience of having had Math 211 at PSU, and almost two-thirds of them specifically had Steve as their instructor before. This means they were familiar with the way the class would be run, and how they as students would be interacting in groups and sharing their thinking with each other. As described earlier in this chapter, the class dynamic is a large part of how the Math 211 and Math 212 courses are conducted at Portland State University. The general sense I had about the class dynamic from observing the students during the first day was one of positive expectations, and this sense is supported by the responses on the PreSurvey questions profiled above.

Overview of Research Design

The two main sources of data for my study were written instruments given to everyone in Steve’s class and individual interviews conducted with six students selected from the class as my case studies. A third source of information was class observations and videotapes made during the stochastics portion of the Math 212 course to help record the learning environment that the students experienced.

Table 2 summarizes the overall research design, the type of information gathered, who it was gathered from, and when it was gathered during the 10-week quarter. The table also shows the contexts (sampling, probability, or data and graphs) emphasized by each instrument or activity.

Overall Research Design				
When	**What**	**Contexts**	**Type of Data**	**Who**
1st week	PreSurvey	All	Written (In class)	Classwide (n=27)
2nd & 3rd week	PreInterview	All	Videotaped (Out of class)	Six cases (n=6)
5th & 6th week	Class Intervention #1 (Four Questions & Body Measurements)	Data & Graphs & Sampling	Videotaped (In class)	Classwide (Varied)
6th week	PostSurvey #1	Data & Graphs	Written (Out of class)	Classwide (n=28)
7th week	Class Intervention #2 (Known & Unknown Mixtures)	Sampling & Graphs	Videotaped (In class)	Classwide (Varied)
7th week	PostSurvey #2	Sampling & Graphs	Written (Out of class)	Classwide (n=30)
6th & 8th week	Class Intervention #3 (Cereal Boxes & River Crossing Game)	Probability & Graphs	Videotaped (In class)	Classwide (Varied)
8th week	PostSurvey #3	Probability & Graphs	Written (Out of class)	Classwide (n=29)
9th & 10th week	PostInterview	All	Videotaped (Out of class)	Six cases (n=6)

Table 2

Prior to doing any class activities in stochastics, during the first week of the academic quarter, baseline information was collected from the Math 212 students in Steve's section. The information was collected via a written survey that addressed prior mathematical experience and contained a range of questions about probability and statistics. Over the next couple of weeks, while Steve taught the geometry portion of the Math 212 course, I interviewed six students outside of regular class time. The interviews were videotaped and lasted about 45 minutes on average. The initial survey and interview were named PreSurvey and PreInterview because they occurred prior to the stochastics portion of the Math 212 course.

In the 5th week of the course, Steve made the transition from geometry to statistics and probability, and I began attending each class session. In addition to making observations, I also videotaped portions of class activities and discussions. Another colleague, Matt, also joined some class sessions to help facilitate activities and assist with videotaping.

Of the many class activities that took place over weeks 5 through 8, six activities were designed as interventions about variation. There were two activities in each of three interventions, with one intervention focused on each of the contexts, data and graphs, sampling, and probability. The three class interventions are listed in Table 2. After each one of the three interventions had been conducted in class, a take-home assignment was given, which I have called PostSurveys because they occurred after the entire intervention for that context had been completed. The PostSurveys are differentiated according to their context: For instance, the first PostSurvey corresponded to the context of data and graphs.

After the last intervention had happened in class, and the PostSurvey (Probability) had been administered, a second interview was conducted with the same six students as had been interviewed for the PreInterview. The second interview was called a PostInterview because it took place after the interventions occurred in class. Like the PreInterviews, the PostInterviews were videotaped and lasted about 45 minutes on average. All PostInterviews occurred in the last two weeks of the quarter.

Rationale for Design

The study is primarily a qualitative design. It seeks to describe conceptions held by elementary preservice teachers (EPSTs) about variation. There is a quantitative aspect to the data gathering from the survey data, but the written responses were qualitatively analyzed for the purpose of letting dominant themes emerge from the data.

In either quantitative or qualitative research paradigms, multiple methods exist. Thus, some additional justification is needed in defending the specific choices within a given paradigm. My choice was to combine classwide data with data from case studies to inform my research questions, and therefore a justification for basing a part of my methodology on case studies follows next.

<u>Case Studies:</u> Case studies are often associated with

ethnographies, grounded theories, or exploratory research. Creswell (1998) offers this definition:

> "a *case study* is an exploration of a 'bounded system' or case (or multiple cases) over time through detailed, in-depth data collection involving multiple sources of information rich in context" (p. 61).

My research is exploratory on two fronts: One, little has been known about how EPSTs think about variation, and two, frameworks for looking at EPSTs conceptions are either undeveloped or barely emergent. By looking at classwide as well as individual case data, I have added detail to an emergent framework using several sources of data. Stake (1994) states that a case study should represent "a specific, unique, bounded system" (p. 237), and the boundaries can be defined by time and place. The Math 212 course, conducted in the same location over ten weeks with the same students and instructor, represented the kind of bounded system needed to conduct a case study. How I chose the six cases is described in the next section.

There are three main reasons why case studies are useful for my research. Case studies are descriptive, they allow theory to be generated, and they use multiple techniques for gathering data. First, recall that the research questions of Chapter One were:

1) What conceptions of variation are held by EPSTs in the three contexts of data sets, sampling, and chance situations?

2) How can the conceptions of variation held EPSTs in these three contexts be characterized?
    A) What variation do EPSTs *expect*, prior to seeing the data or carrying out an experiment?
    B) How do EPSTs produce and reason about *displays* of variation?

C) How do EPSTs *interpret* variation in terms of its causes and effects?

These questions require descriptions of the conceptions of variation held by EPSTs, and ”case studies can provide a detailed description and analysis of processes or themes voiced by participants in a particular situation” (McMillan & Schumacher, 1997, p. 395). Case studies are helpful in “areas of education where little research has been conducted” (Merriam, 1998, p. 38). Since not much previous work had been done finding out EPSTs’ conceptions of variation, case studies are ideal for exploring, describing, and characterizing this knowledge. McMillan and Schumacher (1997) wrote that “case studies are appropriate for exploratory and discovery-oriented research” (p. 395).

Second, in addition to providing a rich description of the subjects, case studies also allow the use of descriptive data “to develop conceptual categories or to illustrate, support, or challenge theoretical assumptions held prior to the data gathering” (Merriam, 1998, p. 38). For my research, there were three aspects of understanding variation which were hypothesized as a useful construct with which to view EPSTs conceptions. The data helped define the dimensions and themes which comprised the aspects of the conceptual framework, *expecting, displaying,* and *interpreting* variation. Case studies are consistent with my study’s goal of developing a more detailed framework, grounded in the data, which is useful for describing EPSTs conceptions of variation.

Third, case studies are well-suited to multiple sources of data, especially document reviews, observations, and interviews (Stake, 1995). The use of different data sources is referred to as *triangulation*, which serves to "clarify meaning by identifying different ways the phenomenon is being seen" (Stake, 1994, p. 241). Triangulation enhances internal validity, reliability, and generalizability, which many consider to be the criteria of research soundness (Romberg, 1992). Internal validity refers to how well the findings of the research match the reality of what went on during the research process. Reliability refers to the extent in which the research findings can be replicated, but "seems to be something of a misfit when applied to qualitative research" (Merriam, 1998, p. 206). Merriam advocates the term 'consistency' in place of 'reliability', saying that "the question then is not whether findings will be found again but *whether the results are consistent with the data collected*" (p. 206, italics in original). My study was triangulated by having the three methods of data gathering include document reviews (via the surveys), observations, and interviews.

Regarding generalizability, it should be noted that whereas a case study represents a specific, unique, and bounded circumstance, the "transferability or generalizability to other settings may be problematic" (Marshall & Rossman, 1995, p. 144). Many writers contend that the notion of generalizability can only be applied to qualitative methodology in a theoretical sense. Some qualitative methodologists prefer to talk of working hypotheses, while others prefer to talk in terms of naturalistic generalizations, which "include the kind of learning that readers take from their encounters with specific case studies" (Patton, 2001, p. 583). Still, "triangulating multiple sources of data…can enhance a study's generalizability," claim Marshall and Rossman (1995, p. 144). Moreover, "especially in terms of using multiple methods of data collection and analysis, triangulation strengthens reliability as well as internal validity" (Merriam, 1998, p. 207).

Written Documents: The use of written documents as a method for collecting data in a case study is well regarded (Stake, 1995; Merriam, 1998; Patton, 2001). Written documents can supplement observational data, and "quite often, documents serve as substitutes for records of activity that the researcher could not observe directly" (Stake, 1995 p. 68). In my study, there were two main types of documents I collected: pre-activity documents (the PreSurvey) and post-activity documents (the PostSurveys for Sampling, Data & Graphs, and Probability). All of the survey instruments are listed in Appendix B. I also collected in-class work that came out of the three interventions.

Observations: The purpose of using observations as a data collection method was to record the overall class contexts in which the activities occurred. I wanted not only to capture the contributions of my six cases, but also to hear what the ideas the rest of the class shared within their small groups and in classwide discussions as they engaged in the activities.

The aim of my research was not to discern the effects of the social environment on learning. Rather, the aim was on the conceptions that the

EPSTs held, and on developing a framework that would be useful in looking at the understanding of individual cases. However, to gain perspective on what conceptions the EPSTs hold, it is critical to understand the context of the environment within which they had directed opportunities for learning. Also, the interventions used in my research provide potential groundwork for future research that could be aimed at measuring the development of a richer understanding of variation within EPSTs. Thus, observations provide a record of what happened in class, establishing a context for learning about variation in a socially mediated environment.

Observations are a common data collection technique in case studies (Stake, 1995; Merriam, 1998). Best and Kahn (1998) write that “when observation is used in qualitative research, it usually consists of *detailed notation* of behaviors, events, and the contexts surrounding the events and behaviors” (p. 253, italics in original). Patton (2001) includes the following three dimensions of concern when conducting observations: the role of the observer, the disclosure of observation, and the recording procedures. In my research, I had roles as both participant and observer. I was a participant by virtue of co-directing some of the Math 212 activities, and an observer by virtue of recording the class activities. Patton claims that

> “the participant observer employs multiple and overlapping data collection strategies, being fully engaged in experiencing the setting (participation) while at the same time observing and talking with other participants about whatever is happening” (pp. 265, 266).

Regarding disclosure, the participant observation is what Fraenkel and Wallen (2000) called *overt*, because the researcher will be identified and the cases will know they are being observed. For recording procedures, Matt and I videotaped the three class interventions. I also videotaped parts of all the other class sessions having to do with probability and statistics, and took notes after each session ended. To "minimize the errors resulting from faulty memory," caution Best and Kahn (1998), "simultaneous recording of observations is recommended" (p. 295). Videorecording of the classroom during the activities showed the context of the learning environment, the overall flow of the class, and the specific contributions of my cases to the class discussion. At multiple times during the three class interventions, the specific tables where my cases sat were videotaped, so that I could capture what they were saying to each other in small-group discussions during the activities.

Interviews: Interviewing is a common and powerful method of trying to understand other people (Fontana & Frey, 1994). Best and Kahn (1995) note that "interviews are used to gather information regarding an individual's experience and knowledge" (p. 255), and Patton (2001) says that the purpose of interviewing "is to allow us to enter into the other person's perspective. Qualitative interviewing begins with the assumption that the perspective of others is meaningful, knowable, and able to be made explicit" (p. 341).

While interview types run the gamut from being highly structured to completely unstructured, the protocol for this research is best described as semistructured (Fontana & Frey, 1994). Merriam (1998) maintains that “interviewing in qualitative investigations is more open-ended and less structured” (p. 74). The interviews for my study followed the two scripts (PreInterview & PostInterview) listed in Appendix B. The initial questions were largely open-ended, and the possibilities for the use of subsequent probes depend on the responses of the interviewees.

Although Fraenkel and Wallen (2000) suggest that “structured and semistructured interviews are best conducted toward the end of a study” (p. 510), this shouldn’t be taken as an absolute rule in qualitative research. I conducted the first round of interviews, the PreInterviews, after the PreSurveys had been collected and reviewed but before any of the class sessions in stochastics had begun. The second round of interviews, the PostInterviews, took place during the last two weeks of class, after the three interventions had taken place. For each round of interviews, there were some tasks given which were identical or to tasks given on the survey instruments. Also, the tasks for the PostInterview reflected themes that had been explored in the class sessions.

The interviews were therefore task-based, since “the subjects’ interactions [were] not merely with the interviewers, but with the task environments” (Goldin, 2000, p. 519). Goldin (2000) goes on to mention the value in task-based interviews, noting that the tasks can be adjusted in wording and content according to the results of previous research. He adds:

> Interview contingencies can be decided explicitly and modified when appropriate. In comparison with paper-and-pencil test-based methods, task-based interviews make it possible to focus research

> attention more directly on the subjects' processes of addressing mathematical tasks" (p. 520).

Thus, a semistructured, task-based interview format was the third method used for gathering data. The interviews were all videotaped so that the subjects' explanations and nonverbal communication was recorded. Some of the cases wrote on their copies of the interview scripts, and their written notes were collected when relevant. Their written notes and observations that I made during the interviews became a part of the overall data for the study, along with the transcriptions of the interviews.

In summary, sources of data for my study included observations of the classroom environment and observations during the interview setting; One classwide PreSurvey, four classwide PostSurveys, and any work provided during the Pre or PostInterviews all contributed to the written records collected for the study. Finally, 12 semistructured task-based interviews (6 Pre and 6 Post) were conducted and transcribed. Together these methods allow for a degree of triangulation, and all methods provided data contributing to the answering of the study's research questions.

Data Gathering

This section is organized chronologically: The PreSurvey was administered before the PreInterviews were conducted, and then each of the three class interventions was followed by a corresponding PostSurvey. Finally, PostInterviews were conducted.

PreSurvey

The PreSurvey was given during the second class session of the first week of the quarter (Thursday, January 8th), and was completed by 27 students. The students had known that a survey was to be administered, because I had visited the first class session and told them. Also during the first class session, I described the research project and their opportunity to be involved, and distributed the informed consent forms found in Appendix A. All students were willing to have their written work included as part of the collected data, and all eventually gave consent to be videotaped during the class sessions. I also had eleven students volunteer to be interviewed outside of class, so I used the PreSurvey responses to help determine who would make up my final six cases. The average time for completion of the PreSurvey was about 45 minutes. The structure of the nine-page PreSurvey had two parts: The first part was the first page, containing background questions to determine the attributes shared earlier in Table 1, such as what prior experiences in probability and statistics they could recall having. Also on the first page were questions about the meaning of the terms “random” and “variation.” The second part of the PreSurvey (comprising the other eight pages) held a total of nine questions, many of which had multiple parts. The specific questions and classwide responses for the PreSurvey are summarized in the next chapter, but the contexts for the questions in second part of the PreSurvey are given in Table 3.

Contexts for PreSurvey (Part Two) Questions		
Question	Brief Description	Contexts
1	Results are predicted for drawing one, several, and six handfuls of 10 candies from a jar (60 Red & 40 Yellow)	Sampling
2	Ranges are predicted for drawing six and then thirty handfuls of candies	Sampling
3	Results are predicted for drawing fifty handfuls of candies	Sampling
4	A graph is made to show what the results of fifty handfuls might look like	Sampling & Graphs
5	Test scores are shown for two different classes. The question addresses the two classes' relative performance	Data & Graphs
6	Two graphs, showing the student heights at two different school, are compared to see which shows more variability	Data & Graphs
7	Results are predicted for performing one, two, and six trials of flipping a fair coin. A single trial consists of fifty flips.	Probability
8	Chances of winning a game involving two 50-50 (Black-White) spinners are addressed.	Probability
9	Chances of winning a game involving a 50-50 and a 25-75 (Black-White) spinner are addressed.	Probability

Table 3

Aside from the background questions, all the rest of the questions on the PreSurvey were either identical or very similar to questions asked by other researchers of middle and high school students. In particular, a NSF-funded project (Shaughnessy, 2002) used written surveys with 12 different classes of middle and high school students, and many questions on the PreSurvey came from the NSF surveys, which in turn had been motivated by prior research. For example, on the PreSurvey and the NSF high school survey, two questions asked students for a description of what the terms “random” and “variation” meant to them, and those questions were similar to the ones used by other researchers (Watson, et. al., 2002; Ciancetta, 2003). A question involving the comparison of graphs showing test scores from two classes was used in the PreSurvey (Question #5) and was the same question used earlier by Watson and Moritz (1999). A question which is not identical, but parallel to those used in past research is Question #7, where I asked about trials consisting of fifty flips of a fair coin. Set in a sampling context, Question #1 uses trials consisting of pulling candies from a jar (akin to the Candy Task mentioned in Chapter 2), and was used on the NSF surveys as well as in other research (Torok & Watson, 2000; Reading & Shaughnessy, 2000; Shaughnessy et. al., 1999; Shaughnessy & Ciancetta, 2000). For the PreSurvey, I chose to use questions that had been used by other researchers to obtain an overview of middle and secondary students' thinking about variation.

PreInterview

After the class in which the PreSurvey was completed, I read through all the responses to get a general sense of how much the students had written and what they had to say. I paid particular attention to the surveys from the eleven students who had said they were willing to be interviewed. Of these eleven students, I selected six who I predicted would make good cases because I thought their responses on the PreSurvey were representative of the whole class and because they provided enough written detail to convince me that they would have no problem sharing their thoughts in an interview situation.

Having a total of six cases is largely a pragmatic decision, determined by the resources available in conducting my research. Qualitative inquiry generally involves relatively small samples, and the key in purposeful sampling is to select “information-rich cases whose study will illuminate the questions under study” (Patton, 2002, p. 230). Creswell (1998) discusses the value in purposeful sampling, saying that cases can be chosen to provide different perspectives on the phenomena under study.

Because my research design required pairs of Pre and PostInterviews for the same subjects, I actually interviewed all eleven of the volunteers over the second and third weeks of the quarter. I interviewed all eleven because although I had in mind the six cases I wanted to end up with (namely DS, EM, RL, SP, JM, and GP who are discussed in the next chapter), I could not guarantee that those six would still be enrolled or be available at the end of the quarter for the PostInterview, nor could I guarantee that the quality of their information in the interview setting would be as rich as I had expected it to be. It turned out that the six cases I had in mind from reading their PreSurveys were in fact very useful for informing my framework and helping me understand their conceptions of variation, and they did complete the postinterviews as well. The other five of the eleven volunteers also completed postinterviews, which I conducted because I had the resources and figured that their contributions gave me additional data with which to work on future research.

For the PreInterview, the subjects met with me outside of the regular classtime. A videocamera was set up to record the interview, and a separate tape recorder also was present because the cassettes were useful in transcribing the dialogue. I had a copy of the interview script (see Appendix B) and also gave a copy to the subject, who was encouraged to write on the script as desired. The PreInterview contained 13 multi-part questions. The contexts for the questions are given in Table 4. More details of the PreInterview questions and sample responses are provided in the next chapter, but my current purpose is to show how the interviews were conducted.

Contexts for PreInterview Questions

Question	Brief Description	Contexts
1	Results are predicted for drawing one, several, and six handfuls of candies from a jar. One trial is a handful of 10 candies.	Sampling
2	Lists are shown for different outcomes of six trials. Subjects are asked to comment on the likelihood of each list occurring.	Sampling
3	The supposed results of 30 trials are shown. Subjects are asked about the likelihood of the results being real or fake.	Sampling & Graphs
4	The supposed results of 300 trials are shown. Subjects are asked about the likelihood of the results being real or fake.	Sampling & Graphs
5	Three graphs show different ways of portraying the same data set. Subjects are asked how the graphs differ.	Data & Graphs
6	A set of 21 measurements for the duration of a train ride is given. Subjects are asked for reasons why the results are not identical.	Data & Graphs
7	The 21 measurements from Q6 are graphed in two different ways. Subjects are asked to compare the two graphs.	Data & Graphs
8	Two different graphs are shown: Wait-times for eastbound and westbound trains. Subjects are asked to compare the graphs.	Data & Graphs
9	One trial is defined as sixty tosses of a fair die. Results are predicted for one trial.	Probability
10	Supposed results are shown for four trials with the die, and subjects are asked which results seem real or fake.	Probability
11	Results from repeated trials with the die are predicted.	Probability
12	A 2:1 (White:Black) spinner is used, and results are discussed in terms of what seems surprising to the subjects.	Probability
13	For the spinner in Q12, one trial is defined as sixty spins. Two graphs showing supposed results from 20 trials are compared	Probability & Graphs

Table 4

The first question on the PreInterview was identical to the first question on the PreSurvey (akin to the Candy Task) for two reasons. First, I wanted to see how their verbal responses compared with what they had written, and to see if they had anything more to add to their earlier explanations. Second, I wanted a familiar context to ease them into the interviewing mode. The smoothness of the transitions from question to question were important because experience has shown me that sometimes the scenario described in a question can be confusing to students who had never actually done such activities. For example, telling someone to imagine drawing 6 handfuls of 10 candies each, with replacement, from a jar containing 100 candies in it (40 of which are yellow and 60 of which are red), can seem like an overwhelming amount of different numbers to keep track of for one question. Because the PreInterview took place before the stochastics portion of Math 212, I did not want to assume that the questions would automatically make sense to the subjects. Therefore, just as the PreSurvey was scaffolded to include one trial, then several, then six, so too was the PreInterview modeled to move gradually to ever larger numbers of trials.

The PreInterview script contained specific questions that I used with each subject, but the protocol also allowed me flexibility to follow the subjects' train of thought. Thus, each interview contained common questions as well as unique input from the different cases.

Class Intervention #1

All of the class interventions are described in greater detail in Appendix C, and in this chapter they are only briefly explicated. The two activities comprising the Class Intervention for the context of data and graphs were called “Four Questions” and “Body Measurements”.

The “Four Questions” activity was chosen for two reasons. One reason is because Steve and I had each used versions of the activity with other Math 212 classes, and were therefore experienced in how it went and what it offered. The second reason is because it offered a good opportunity to discuss both average and spread in data sets. Steve therefore started the class exploration of statistics in the fifth week by having the entire class gather data from one another in response to four questions:

> How many pets do you have?
> How many years have you lived in Portland (to nearest half-year) ?
> How many people are in your household?
> How much change (in coins) do you have today?

After graphing the data in different ways, the class had a discussion about levels of detail provided by each type of graph and about what were “typical” values for a Math 212 student or for the whole class. The tension between centers and spread of data was one theme to emerge from the discussion over graphs from the “Four Questions” activity.

The second activity in the class intervention focusing on the context of data and graphs was “Body Measurements”, which was also selected for the reason that it was a well-rehearsed activity for Steve and me. More importantly, a similar activity was to be used for the NSF-sponsored project (mentioned earlier) that looked at conceptions of variation in middle and high school classes. As in “Four Questions”, we gathered class data for “Body Measurements”: Everyone’s own armspan, height, handspan, head circumference, and pulse rate per minute was recorded. Also, all students in class measured Matt's armspan, to show what data from a repeated-measurements experiment looked like. Again, we had a class discussion about the data and graphs, this time focusing more on causes of variation.

PostSurvey – Data & Graphs

The take-home assignment PostSurvey (Data & Graphs) was given at the end of the final class session of week 6, and collected the following week. Students had from Thursday to Tuesday to work on the PostSurvey, and they were encouraged to work together on the assignment. The PostSurveys were graded by Steve only as "done" or "not done" , similar to other writing assignments given in class. Table 5 summarizes the questions asked on the PostSurvey for the context of data and graphs, and the classwide results are summarized in the next chapter.

Summary of PostSurvey (Data & Graphs) Questions

Question	Brief Description
1a	Bar charts are given showing the 30-year average monthly rainfall for Portland and Columbus. Students discuss differences and causes in the rainfall patterns.
1b	Boxplots are given for the same data sets used in Q1a. Again students discuss differences, and also are asked which city they think is rainier, and why.
1c	Assuming that the average June rainfall in Columbus is 4 inches, students are asked to draw a graph showing what each day's rainfall in June might look like.
2a	Dotplots and boxplots are given showing annual traffic death rates for two regions in America, the South and Northeast. Students are asked to compare the rates.
2d	Students are asked to think of factors that might explain the differences in rates between the two regions.

Table 5

I had several reasons for choosing the PostSurvey (Data & Graphs) questions that I did. One was to examine reasoning while students compared or evaluated the different graphs. For example, would students refer more to centers or to spreads? Also, I wanted see what kinds of causes for variation they could come up with on their own, and in Q1c, I was curious about their ability to reason from the given average of 4 inches of rain to generate a reasonable graph showing appropriate variation for the rainfall during a typical June in Columbus. Finally, for my six cases and the others who would be interviewed a second time, I wanted additional evidence in reasoning from histograms, dotplots, and boxplots.

Class Intervention #2

In the seventh week of class, the two activities “Known Mixture” and “Unknown Mixture” were done with Steve’s students. Matt and I had done this activity at six schools as a part of the NSF-sponsored project. We had seen how effective the activities were in drawing attention to variation.

For the Known Mixture, we started with a general discussion of what were samples, who uses samples, and what samples were good for. Then the following scenario was given as a part of a handout (see Figure 14):

Scenario for Known Mixture Activity
The band at Johnson Middle School has 100 members, 70 females and 30 males. To plan this year’s field trip, the band wants to put together a committee of 10 band members. To be fair, they decide to choose the committee members by putting the names of all the band members in a hat and then they randomly draw out 10 names

Figure 14

The class discussed initial expectations for this scenario, especially focusing on what would happen if the random draw of 10 names were to be repeated thirty times. After students talked about predictions for drawing thirty samples each of size ten, we simulated this activity using chips in a jar. Actual data was gathered and graphed. Then we had a discussion about how the graphs of the predicted data compared to one another, how the graphs of the actual data compared to one another, and also how the predicted graphs compared to the actual graphs.

We then made a transition into the second activity in this intervention, which was the Unknown Mixture. It was made clear that even though we had known what was in the earlier jars, samples still had varied. Now we had larger jars, each containing 1000 chips of yellow and green, and the mixtures were known to be the same in each jar. However, the exact mixture was not known to the class (it was actually 550 yellows and 450 greens). The students were, in their groups, to decide what sample size they wanted to use (we imposed an upped limit of size twenty for all groups) and how many samples they wanted to draw. Then they were to carry out their plans, do the sampling, graph the results, and make some conjecture about the true mixture in the jar. After the simulation was carried out, we had a class discussion about the different choices made in sampling, the class results, and we tried to forge a class consensus about what was the true mixture.

(PostSurvey-Sampling)

The take-home assignment PostSurvey (Sampling) was given at the end of the final class session of week 7, and collected the following week. Table 6 summarizes the questions asked on the PostSurvey for the context of sampling, and the classwide results are summarized in the next chapter.

Summary of PostSurvey (Sampling) Questions	
Question	Brief Description
1a	A boxplot shows the results of 20 samples of size 10 drawn from the smaller candy jar (60 Red & 40 Yellow). Individual sample results are then inferred.
1b	Since the minimum result on Q1a shows 3 Reds in a handful of 10, the number of trials needed to get a 0 or 1 Red are predicted.
2	Results are predicted for drawing one, several, and six samples of 100 candies from a larger jar (600 Red & 400 Yellow)
3	Ranges are predicted for drawing thirty and then three hundred samples of 100 candies
4	Results are predicted for drawing fifty samples of 100 candies

Table 6

I asked Q1a because it got at the idea explored in class, how boxplots can obscure variation. I was also curious to investigate their ability to work with boxplots, especially working backwards and predicting what the results might be which contributed to the boxplot (as opposed to taking the data and creating the boxplot as is typically done). Q1b was added to draw on our class experience with the ProbSim software and the numbers of samples needed to get extreme results. Questions 2, 3, and 4 all were parallel to questions previously asked in the PreSurvey and PreInterview for the Small Jar (60 Red & 40 Yellow), only now the sampling was done from the Large Jar (600 Red & 400 Yellow). I was curious to see how answers for the Large Jar compared to what students had put for the Small Jar, and I also wanted to have my six cases thinking about the Large Jar because I built additional questions on that sampling scenario in the PostInterview.

Class Intervention#3

There were two activities that made up this intervention, “Cereal Boxes” and “The River Crossing Game,” and these were chosen specifically because of the probability aspects involved in the activities. Namely, Cereal Boxes relies on the use of spinners and River Crossing on the use of dice as random generators, and these two activities were the main ones done in Math 212 involving spinners or dice.

Cereal Boxes actually took place in the first class session of week 2, just before we gathered data for Body Measurements. As explained earlier, there was considerable overlap in the contexts, and Cereal Boxes is a good example of this overlap. Cereal Boxes is sample-until scenario, supposing that any of five different stickers can be obtained with each box of cereal opened, and that the five stickers have equal chances of being obtained. The question is, about how many boxes would be opened to obtain all five stickers, and the situation can be simulated by using a five-region spinner. Cereal Boxes brings together probability, sampling, and data and graphs in a way that also highlights variation, and that is why I chose to use the activity.

The second activity for this intervention, the River Crossing Game, involved finding the sum of two dice. Credit for this activity goes to the *Math and the Mind’s Eye* curriculum (Shaughnessy & Arcidiacono, 1993). Using two players, each player got 12 chips to place on their side of a “river”, along spaces marked 1 through 12. After configuring their chips in an initial arrangement (see Figure 15 for an example of two players' initial arrangements), players took turns tossing a pair of dice. If either player had any chips on the space showing the total for the dice, one chip could "cross the river" and be removed from the board. The winning player was the first one to remove all the chips on his or her side.

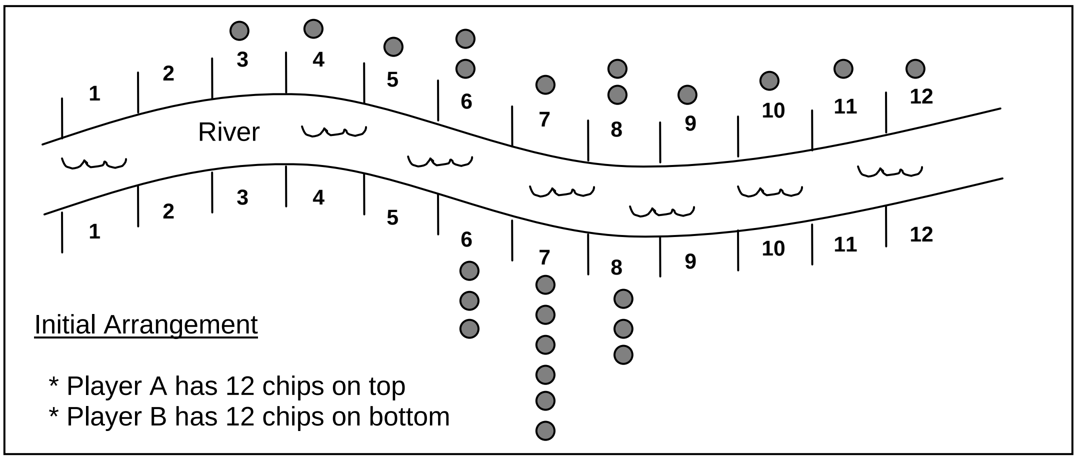

Figure 15

For instance, in Figure 15, if the dice resulted in a sum of 10, Player A on top could remove one chip. If the dice showed 8, Player A and B could each remove one chip. As with Cereal Boxes, in the River Crossing Game we made predictions, gathered and graphed data, and discussed results.

(PostSurvey-Probability)

The take-home assignment PostSurvey (Probability) was given at the end of the final class session of week 8, and collected the following week. Table 7 summarizes the questions asked on the PostSurvey for the context of probability, and the classwide results are summarized in the next chapter. Question 1 was crafted to be similar to the sampling questions on both the

PreSurvey and also the PreInterview and the PostSurvey (Sampling), only this time the context was probability. Instead of grabbing handfuls of candies, the students were asked to consider spins at a 50:50 (White:Black) spinner as comprising one trial. Also, Question 1 was similar to the PreSurvey question in probability that used as its trial 50 flips of a fair coin.

Summary of PostSurvey (Probability) Questions	
Question	Brief Description
1	The number of blacks resulting from 50 spins at a ½ -white & ½ -black spinner make up a single trial. Results are predicted for doing one, two, and six trials.
2	Ranges are predicted for doing thirty and then three hundred trials with the spinner.
3	A graph is made to show what the results of fifty trials at the spinner might look like

Table 7

Question 2 was similar to what was asked on the PreSurvey and PostSurvey (Sampling) in the sampling context, and the graph of fifty trials for question 3 was similar to what was asked for fifty handfuls at the Small Jar in the PreSurvey. I was curious to see how the probability questions with the flips of a coin (in the PreSurvey) compared to the spinner environment (in the PostSurvey), and also how the responses for the context of probability compared with those responses for the sampling context. Finally, for my six cases I wanted added familiarity with the spinner scenario, since I based several additional questions in the PostInterview on trials at the same spinner.

PostInterview

The PostInterviews took place during the 9th and 10th weeks of the quarter, after Steve had gone back to teaching geometry. I followed a very similar protocol with the PostInterviews as I had with the PreInterviews, videotaping each interview as well as making a separate audiorecording for the transcribing process. The PostInterview contained 13 multi-part questions, and the contexts for the questions are given in Table 8. More details of the PostInterview questions and sample responses are provided in the next chapter.

Contexts for PostInterview Questions		
Question	Brief Description	Contexts
1	Results are predicted for drawing one, several, and six handfuls of candies from a large jar. One trial is a handful of 100 candies.	Sampling
2	Lists are shown for different outcomes of six trials. Subjects are asked to comment on the likelihood of each list occurring.	Sampling
3	The supposed results of 30 trials are shown. Subjects are asked about the likelihood of the results being real or fake.	Sampling & Graphs
4	The supposed results of 300 trials are shown. Subjects are asked about the likelihood of the results being real or fake.	Sampling & Graphs
5	Two graphs show supposed results of forty trials at the small and also at the large jar. Subjects are asked if graphs are real or fake.	Sampling & Graphs
6	A set of 20 measurements for the weight of a muffin is given. Subjects are asked for reasons why the results are not identical.	Data & Graphs
7	The 20 measurements from Q6 are graphed in two different ways. Subjects are asked to compare the two graphs.	Data & Graphs
8	35 different muffins from the West End bakery are shown. Subjects are asked how much their (36^{th}) muffin might weigh.	Data & Graphs
9	Two different graphs are shown: Muffin weights for East and West End bakeries. Subjects are asked to compare the graphs.	Data & Graphs
10	Results are predicted for doing one, two, and six trials at the 50:50 spinner. One trials is fifty spins.	Probability
11	Lists are shown for different outcomes of six trials. Subjects are asked to comment on the likelihood of each list occurring.	Probability
12	A graph shows the supposed results of twenty trials. Arguments from other people about the results are discussed by subjects.	Probability & Graphs
13	Two graphs show supposed results of two classes doing thirty trials at the spinner. Subjects are asked if graphs are real or fake.	Probability & Graphs

Table 8

Again, as in the PreInterview, I chose the very first question on the PostInterview to be identical to a question the subjects had already seen, since it had been asked in the PostSurvey (Sampling). Also, the first four questions were isomorphic to those on the PreInterview, except that they used the Large Jar (600 Red & 400 Yellow) as opposed to the Small Jar (60 Red & 40 Yellow). Question 5 involved both Small and Large jars, the only question in any of the instruments to do so directly. Questions 6, 7, and 9 were asked in part because of their similarity to the MAX train ride questions on the PreInterview. Lastly, question 10 was identical to the first question on the PostSurvey (Probability), and set the subjects up for a transition to the last few questions that also involved spinners.

It was mentioned earlier how the activities in the interventions were designed to get at variation. As can be inferred from Table 8 , each activity in an intervention had a direct tie-in to some task or tasks on the PostInterview questions. For instance, the intervention on data and graphs included different types of graphs and the amounts of variation they showed. Body Measurements directly got at the ideas behind repeated measurements, as did the muffin weight questions on the PostInterview. The Known and Unknown mixtures let students actually draw chips from a container to experience the idea of drawing candies from the Large and Small Jars. Cereal Boxes and the River Crossing Game let students use the traditional random generators of spinner and dice to get a sense of what was likely in a probability context. Finally, there is a big difference on the PostInterview compared to the PreInterview: PostInterview questions 8, 9, and 12 all included boxplots as well as either dotplots or histograms, but boxplots were not covered on the

PreInterview. Thus, several of the tasks involving graphs had two types of graphs, again relying on the experience gained in the class interventions.

Data Analysis

This section first presents the role of grounded theory in the data analysis process, both in general terms and also as it applied to my study. Subsequently, an brief outline is given of how the data were analyzed, with the details in the subsequent chapters. The role of computer software in facilitating the process of data analysis for my study is also discussed.

Role of Grounded Theory

The general principles for the data analysis are derived from grounded theory, applied both to the survey and the interview data. The aim in the analysis was to describe and categorize the data, resulting in a tentative theory of what are the key features of the subjects' conceptions of variation. Strauss and Corbin (1994) said that "grounded theory is a *general methodology* for developing theory that is grounded in data systematically gathered and analyzed" (p. 273, italics in original). The techniques prescribed fit the aim of the analysis needed for this study.

The ultimate aim of most qualitative studies, claim Huberman and Miles (1994) is "to describe and explain (at some level) a pattern of relationships, which can be done only with a set of conceptually specified analytic categories" (p. 431). Grounded theory begins by describing and building categories, the dimensions of which are defined by their conceptual properties.

Emergent tentative hypotheses suggest links between categories and properties (Patton, 2001; Merriam, 1998).

The development of these categories, properties, and hypotheses occurs through a method of constant data comparison, “a process whereby the data gradually evolve into a core of emerging theory” (Merriam, 1998, p. 191). More specifically, data analysis begins even while the data is being collected, possibly in the form of memos regarding the researcher’s conjecture for possible categories. Then, the data is iteratively compared to the emerging categories, while the categories are refined in light of reviewing the data (Strauss & Corbin, 1994). “This process of taking information from data collection and comparing it to emerging categories is called the *constant comparative* method of data analysis” (Creswell, 1998, p. 57, italics in original). Patton (2001) calls this comparative analysis “a central feature of grounded theory development” (p. 490).

Because my study was concerned with EPSTs conceptions, the techniques of grounded theory were of particular appeal because “concepts are the basic units of analysis in the grounded theory method” (Strauss & Corbin, 1990, p. 63). The authors continue by stating that “conceptualizing our data becomes the first step in analysis” (p. 63). Open coding is described as the process of identifying the concepts and discovering their properties. A detailed line-by-line analysis, while labor intensive, is extremely generative and especially useful at the outset of a study. The establishing of categories has much in common with the clustering technique of Miles and Huberman

(1994), in which the aim was to “understand a phenomenon better by *grouping* and then *conceptualizing* objects that have similar patterns or characteristics” (p. 249). An example of the use of this technique in the study of understanding variation is provided by Torok and Watson (2000). Adding to the power of open coding is axial coding, defined as “the process of relating categories to their subcategories, termed ‘axial’ because coding occurs around the axis of the category, linking categories of the level of properties and dimensions” (Strauss & Corbin, 1998, p. 123). Microanalysis is the combined approach of open and axial coding, using a line-by-line analysis, to “generate initial categories (with their properties and dimensions) and to suggest relationships among categories” (p. 57).

In analyzing my data, both within-case and cross-case analyses were done. Within-case analysis “means that the researcher identifies themes within a single case” (Creswell, 1998, p. 252), and provides a detailed description for that case (Huberman & Miles, 1994). For my study, since variation was looked at in several distinct contexts, it is important to look for themes occurring within each individual case before comparing across cases. Cross-case analysis “involves examining themes across cases to discern themes that are common to all cases” (Creswell, 1998, p. 250), and it typically follows within-case analysis when multiple cases are studies. Some authors warn of the complexities of cross-case analysis (Huberman & Miles, 1994; Miles & Huberman, 1994), but Merriam (1998) says that “ultimately, cross-case analysis differs little from analysis in a single qualitative case study” (p.

195). The result of the analysis is a comprehensive set of "categories, themes, or typologies that conceptualize the data from all the cases" (p. 195).

While discussing multiple cases, Creswell (1998) brings up the topic of how many cases is enough, cautioning against a rationale for improved generalizability, "a term that holds little meaning for most qualitative researchers" (p. 63). Rather, a more fitting rationale for the tenets of grounded theory is a desire for saturation, meaning that point at which new cases are not adding any new information to the establishment of categories (Creswell, 1998).

Outline of Analysis

I have referred to my six interview subjects as my six cases, because I do focus on the thinking of those individuals in Chapter 6. However, in addition to using the interview data from the six cases, I also used classwide data from the written surveys, summarized in the next chapter (Chapter 5). In looking at the class as a whole, the class does fit the notion of a separate "case" when using the language of grounded theory, although I just refer to the class or the classwide data. I outline how I used the classwide data as well as the interview data in developing the emergent framework that came from the analysis.

Recall that my initial conceptual framework from Chapter 3 held the following aspects and dimensions as a potentially useful lens for looking at EPSTs conceptions of variation:

[1] Expecting Variation
- A] Describing What is Expected
- B] Describing Why (Reasons for Expectations)

[2] Displaying Variation
  A] Producing Graphs
  B] Comparing Graphs
  C] Making Conclusions about Graphs
[3] Interpreting Variation
  A] Defining Variation
  B] Causes of Variation
  C] Effects of Variation
  D] Influencing Expectations and Variation

The classwide data was used to see if the responses fell into the above aspects and dimensions, and also what themes might emerge for each dimension. For example, in describing *what* was expected, at least 27 students writing on several questions addressed that dimension, and I was able to find some common themes which added depth to the dimensions of the intial framework. Thus, by the end of the analysis of the classwide survey data in Chapter 5, I had a revised framework which I then carried into the analysis of the interview data. The revised framework is shown at the end of Chapter 5 because it began to emerge from the analysis in that chapter.

Using the revised (emergent) framework, I then looked at the interview data for the PreInterview in the first part of Chapter 6, and considered how all six of the cases responded. The goal was to confirm and add further depth to the themes I had seen emerge from the classwide data. Finally, to show the usefulness of the revised framework for looking at individual thinking, I turned to the PostInterview data and focused on a subset of PostInterview questions, writing about each of the cases in turn for the second part of Chapter 6. Creswell (1998) notes that

"When multiple cases are chosen, a typical format is to first provide

> a detailed description of each case and themes within the case, called a *within-case analysis*, followed by a thematic analysis across the cases, called a *cross-case analysis*" (p. 63).

The comparison across cases, with respect to the framework of analysis, occurs in the last chapter.

Role of Computer Software

A last methodological component to be discussed is the role of computer software in my data analysis. The creative yet systematic process of theory building, which in my study took the form of fleshing out a conceptual framework about variation, was aided enormously by the NUD•IST software.

Grounded theory techniques allow for the inclusion of a wide scope of data, such as the written responses, transcribed interviews, and observational notes. As the process of theory development moves through cycles of constant comparison, memos suggesting the continual conjecturing and refinement of categories and concepts also become a part of the data. Thus, data management becomes a crucial issue in using grounded theory. Richards (1994) boldly states that "all researchers working in the qualitative mode will clearly be helped by some computer software" (p. 105). The use of qualitative data analysis software facilitates not only the management of data, but "it can offer leaps in productivity for those adept at it" (Patton, 2001, p. 447).

The software used in this study was NUD•IST (Non-numerical Unstructured Data Indexing, Searching, and Theorizing), a theory-building program that aids in data storage, coding, retrieval, and category comparisons and linking (Richards and Richards, 1994; Richards, 1994; Patton, 2001).

NUD•IST is well-suited for the analysis techniques of grounded theory, although it cannot be emphasized enough that software only assists in the process – software does *not* analyse data for the researcher (Patton, 2001; Creswell, 1998). The flexibility in coding, categorizing, and revising made the development of theory a dynamic and reflexive process. The single biggest obstacle was in learning what NUD•IST could do for the analysis, and how to get the program to do what I wanted. Once the various ways of categorizing, indexing, and coding were learned – the logistics of the program as well as the potential – NUD•IST became an invaluable tool.

Some researchers have used NUD•IST to do automatic coding. Watson and Moritz (2000a) mention how, in their research, “command files automatically assigned responses to categories, which were manually checked for accuracy” (p. 11). My study used the more laborious yet potentially illuminative task of line-by-line coding from the outset, chiefly because my framework was emergent and not rigidly defined at the initial phase of analysis. I did some automatic coding on some key words associated with descriptions of variation, but my primary reliance on NUD•IST was to help identify the themes within the dimensions of the framework.

Conclusion

Subjects were chosen from the Math 212 class, where the pedagogical style in the class and in its prerequisite, Math 211, encourages students to communicate their thinking both verbally and in writing. Subjects were familiar with and able to openly share their thinking, a necessary condition for this study. Within the array of possible methods in the qualitative research paradigm, a modified case study design was chosen as suitable for answering my research questions. In addition to interview data gathered from six cases, classwide data was gathered via a set of written documents (surveys), and the study was also augmented by in-class observations.

The three data gathering methods of observation, document review, and interviews all were connected to the three class interventions. The three interventions (made up of two activities each) corresponded to the three contexts for looking at conceptions of variation: variation in data sets, variation in sampling, and variation in probability situations.

After administering a PreSurvey to the whole class, PreInterviews were held outside of class time. Then, after each class intervention was conducted, a take-home PostSurvey was distributed. Finally, PostInterviews were completed.

To analyze the data, the use of grounded theory techniques allowed for the characterization of the subjects' understanding to naturally emerge within the conceptual framework in the shape of distinct but linked themes. The rudimentary structure hypothesized in the conceptual framework was validated and extended by looking at all the classwide data. Case data from the PreInterview added further detail to the framework, which was then used to look at individual thinking on a subset of the PostInterview data. The data analysis was facilitated by the use of the NUD•IST computer software. The final result was a rich description and overall characterization of the conceptions of variation held by elementary preservice teachers.