

AN IN-SERVICE MIDDLE SCHOOL TEACHER'S  
CONTENT KNOWLEDGE OF THE  
VARIABILITY OF DATA DISTRIBUTIONS

By

Marie Pia Turini

A DISSERTATION

Submitted to  
Michigan State University  
In partial fulfillment of the requirements  
For the degree of

DOCTOR OF PHILOSOPHY

Curriculum, Instruction, and Teacher Education

2011

## ABSTRACT

### AN IN-SERVICE MIDDLE SCHOOL TEACHER'S CONTENT KNOWLEDGE OF THE VARIABILITY OF DATA DISTRIBUTIONS

By

Marie Pia Turini

Understanding of statistics is important and central to a democratic society. It helps to develop citizens, consumers, and workers that are productive and engaged critically and knowledgeably in the world around them. Statistical literacy has become primarily the responsibility of our schools. As a result, teachers need to know what and how to teach statistics. Unfortunately, few teachers have the specialized knowledge necessary to do so. While some research has identified the sense teachers make of statistics, there has been a lack of research of middle school teachers' content knowledge of variability. This study aimed to fill some of that gap by investigating a middle school teacher's content knowledge of variability in data distributions.

A qualitative study was undertaken with a middle school teacher during her academic school year. This study took a bidirectional view of studying teacher content knowledge—prior to and during teaching—including professional development sessions, lesson planning, performance tasks with an interview, and teaching a lesson. This approach differs from the typical education research on teacher knowledge that tended to focus on what teachers know either prior to or during teaching. In addition, in this study I used the construct of sensemaking,

which includes the cognitive practices of noticing, interpreting, and implementing (Weick, 1995; Drake 2006) to study teacher knowledge as a dynamic phenomenon.

The findings of this study showed that there was a dynamic connection between teacher content knowledge exhibited prior to or during teaching, and a more complete snapshot of her content knowledge was attained. Teacher knowledge exhibited in one direction—prior to teaching—was either confirmed, extended, or not made visible during teaching. And in the other direction—during teaching—new content knowledge was exhibited that was not captured before.

The results of this study also attest to the value of using the construct of sensemaking to study teacher content knowledge as a dynamic phenomenon. Findings from this study demonstrated that the construct of sensemaking aligned with the bidirectional view of studying teacher content knowledge. Findings also indicated that the fruitfulness of the teacher's cognitive sensemaking practices—notice, interpreting, and implementing—could affect the strength of the content knowledge that the teacher exhibits.

Although promising results were obtained regarding the use of this bidirectional approach and the construct of sensemaking—notice, interpreting, and implementing—to study teacher content knowledge of variability, their usefulness and applicability are yet to be investigated in other statistics topics and other academic subjects.

## ACKNOWLEDGMENTS

I acknowledge the support and love from my Savior, Jesus Christ that I received to write this dissertation. I thank Him for sending me the people that helped me through this process. I am indebted gratefully to them for their help. I thank Dr. Sandra Crespo for her unending patience and expert guidance. I thank my family for their every present understanding and love. I thank my friends—Leslie, Jason, the Sterns, Lorrie, Amy, the Tuccillos, Merna, Johnson and Linda, Jackie, and Deb—for their continued prayer and intercessions. I thank my Michigan church family—the Pruitts (both families), the Ackroyds, the Browns, and the Collogies—for their support and love. I thank Deborah Gold for suggesting I pursue a doctoral degree when I taught her fifth grade class. I thank all the other teachers for their kind words of support. I especially thank Laura for her help with proofreading, and Reine for her guidance in the formatting. I thank the Connected Mathematics staff—Glenda, Betty, Susan, Judy, Sarah, and Chris—for their guidance and help when I worked on the Project. Without everyone, this task could not have been accomplished. Thank you very much and God Bless You All!

## TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
Introduction	1
Problem	2
Theoretical Considerations	6
Chapter 2 Literature Review	11
Teacher Understanding of Variability	16
Preservice Teachers	18
Preservice Elementary Teachers	18
Preservice Middle School Teachers	22
Preservice Secondary Teachers	26
In-Service Teachers	28
In-Service Elementary Teachers	28
In-Service Secondary Teachers	28
Summary	32
Situating This Research in the Field of Statistics	32
Statistical Investigation	33
Clarification of Key Terms	36
2005 GAISE Report	38
Chapter 3 Research Design	44
Questions	45
Setting	46
The School	47
Professional Development	49
The Curriculum	52
Participants	55
Researcher's Roles	58
Data Collection	62
Chapter 4 Data Analysis	73
Theoretical Framework for Data Analysis	73
Data Analysis	78
What Is a Notice	79
Counting Notices	84
How Mae Interpreted	90
What Mae Interpreted	91
Chapter 5 A Closer Look at What Mae Noticed	93

Chapter 6 A Closer Look at Mae's Interpretations of Variability	106
Mae Used Partitioning as a Way to Interpret Variability	109
Multiple Reference Lines	109
Single Partitions or Benchmarks	110
Mae Used Measures of Center as a Way to Interpret Variability	114
Mae Chose a Classroom Problem for Her Lesson Based Upon Measures of Center	114
Mae Determined That the Purpose of Reference Lines Is to Find the Main Cluster or Modal Clump	116
Mae Used Relative Locations of Measures of Center When Discussing Variability	119
Mae Used Her Perceived Outliers and Range as a Way to Interpret Variability	125
Defining Variability Includes the Concept of Outlier	127
Mae Defines Her Perceived Outliers	128
Context Involved in Mae's Treatment of Her Perceived Outliers	130
Including Her Perceived Outliers	130
Excluding Her Perceived Outlier	134
Mae's Perceived Outliers and Range	137
Mae Used Shape as a Way to Interpret Variability	141
Shape of Whole Distribution	142
Shape of Clusters	143
Mae Constructed Definitions and Descriptions of Variability	147
Defines Variability	148
Defines/Describes Variability	149
Describes Variability	150
Chapter 7 Mae's Sensemaking of Variability Through Her Lesson Implementation	154
Measures of Center: Mae's Sensemaking During Lesson Implementation	158
Perceived Outliers: Mae's Sensemaking During Lesson Implementation	161
Defining Variability Includes the Concept of Outlier	162
Mae Defines Her Perceived Outliers	163
Context Involved in Mae's Treatment of Her Perceived Outliers	166
Including Her Perceived Outliers	167
Excluding Her Perceived Outliers	171
Range: Mae's Sensemaking During Lesson Implementation	174
Chapter 8 Discussion	181
Review and Interpretation of Results	182
Contributions From the Study	188
Sensemaking as a Construct to Study Teacher Knowledge	188
Empirically Testing Canada's (2004) Evolving Framework	194

Framework for Analyzing Middle School Teachers’ Sensemaking of Variability of Data Distributions	197
Assessing the Usefulness of the 2005 GAISE Curriculum Framework for Discussing Teacher Content Knowledge of Variability	198
Limitations	200
Implications	203
Future Research	205
Future Research for Education and Statistics and/or Mathematics Education	205
Extending the Generality of This Study	206
Extending the Domain of This Study	206
Bidirectional View of Teacher Content Knowledge	206
Sensemaking as a Construct to Study Teacher Knowledge	207
2005 GAISE Curriculum Framework	207
Studying Variability and Its Entailments	208
Summary	209
Closing	210
Appendices	213
Bibliography	223

## LIST OF TABLES

Table 2.1 Canada's (2004) Evolving Framework for Elementary Preservice Teachers' Thinking About Variation	19
Table 2.2 Item Used to Assess Teachers' Statistical Knowledge of Teaching	24
Table 2.3 Table Adapted From the 2005 GAISE Report	39
Table 2.4 Levels Where Formal Treatment of Focal Terms Are Recommended by the 2005 GAISE Curriculum Framework	41
Table 3.1 Type of Data Collection, Responsible Person, and Method of Collection	63
Table 3.2 Questions in Tabular Format With Data Collection Methods	63
Table 3.3 Professional Development: Problems Covered, Goals Accomplished, Methods Used, and Connection to Canada's (2004) Evolving Framework	68
Table 4.1 Questions, Words, and Phrases That Emerged When Selecting Mae's Notices	81
Table 4.2 Chart for Keeping Track of Mae's Notices	85
Table 4.3 Criteria for Counting Mae's Notices	86
Table 5.1 Applicable Parts of Canada's (2004) Evolving Framework	94
Table 5.2 The Results of Mae's Notices	95
Table 6.1 A Snapshot of Mae's Sensemaking and Interpretation of Variability	114
Table 6.2 A Snapshot of Mae's Sensemaking and Interpretation of Variability	125
Table 6.3 A Snapshot of Mae's Sensemaking and Interpretation of Variability	140
Table 6.4 A Snapshot of Mae's Sensemaking and Interpretation of Variability	140
Table 6.5 A Snapshot of Mae's Sensemaking and Interpretation of Variability	147
Table 6.6 A Snapshot of Mae's Sensemaking and Interpretation of Variability	153

Table 7.1 Phrases Mae Used to Define Her Perceived Outliers Prior to and During Teaching	164
Table 8.1 Bidirectional View of Mae's Content Knowledge of Variability When Discussing Range	186
Table 8.2 Bidirectional View of Mae's Content Knowledge of Variability	187
Table 8.3 A Snapshot of Mae's Sensemaking and Content Knowledge of Variability	191
Table 8.4 An Outline of Canada's (2004) Evolving Framework of Elementary Preservice Teachers' Thinking About Variation	195
Table 8.5 Framework for Analyzing Middle School Teachers' Sensemaking of Variability of Data Distributions	197
Table 8.6 Alignment of Mae's Content Knowledge with 2005 GAISE Levels and Bidirectional View of Teaching	199
Table 9.1 2005 GAISE Curriculum Framework	219

## LIST OF FIGURES

Figure 2.1 Concept Map Adapted From “Doing Meaningful Statistics—Central Statistical Ideas for Data Distributions” (Lappan et al., 2006, p. 5)	34
Figure 6.1 Distribution of Jalin’s Head Measurements With Suspected Outlier	128
Figure 6.2 Graphs From Performance Task 2	151
Figure 7.1 Distribution of Mae’s Class’s Head Measurements With Inner Fences for Suspected Outliers	162
Figure 8.1 Bidirectional View of Teacher Content Knowledge	184
Figure 8.2 Results of Using Bidirectional View of Teacher Content Knowledge	185
Figure 9.1 Graphs of School A and School B	217
Figure 9.2 Mae’s Class’s Head Measurements	222

## **Chapter 1**

### **Introduction**

I feel like I am just learning about fractions myself as a teacher, because I never really understood them as a child when I was in school.

Second-year sixth grade teacher

Although this quote is not about statistics, it is one that captures the reason for my study. I, too, had an epiphany while teaching that mathematics actually made sense—that it was a topic that I could reason about and come to understand. This revelation was shocking, exciting, and stimulating to me because I never considered myself a math person. I often felt excluded from the club of math knowers—sitting on the sidelines in mathematics classes wishing and, yet, fearful of being asked to fully participate. Although I maintained an 80+ average, it was from diligence and perseverance in studying the rules. It was not until my adult years that I began to realize that I was not alone in this.

Graduate school provided me the first opportunity to study a topic in the field of mathematics intently. Professor Rand Spiro impressed upon me the concepts behind the various formulas used, for example, in hypothesis testing. I began to enjoy and understand statistical concepts in ways that were deeper and more rewarding for me. I was inspired to sit in on additional statistics classes. However, my new knowledge of the field of statistics was not enough for me to share with other teachers who were charged with teaching this topic to middle school children. This motivated me to assist an author of a standards-based middle school curriculum that focused on statistics. At that time, I began to wonder about the

knowledge middle school teachers have of statistics. I did not think I was alone in needing a greater knowledge of statistics in order to teach it.

My love for and understanding of statistics increased during this time. As previously stated, I realized that many teachers do not have this knowledge—the kind that would help them engender statistical experiences that would introduce into students’ minds a mind-set, for example, of “seeing variation” (Watkins, Schaeffer, & Cobb, 2003). Teachers do not have this statistical knowledge for themselves and, therefore, most probably do not have the thinking-on-your-feet statistical knowledge needed in the classroom (Mickelson & Heaton, 2004). My personal journey in becoming a nascent knower of statistics, coupled with teachers’ gaps in this knowledge, fueled my desire to study teacher knowledge of school statistics.

This desire was somewhat quenched when I did my practicum study. In this study, I was able to investigate what a new teacher learned while teaching a statistical concept—the mean. As evidenced by the participating teacher’s posttest explanations, her conceptual understanding developed about the mean. The result of my study prompted me to read further about what teachers know about the statistics they are asked to teach. When I discovered there is little identified about what teachers know about this important topic, I decided to frame my dissertation to investigate what teachers know about the statistics they are charged to teach.

### **Problem**

Statistics is used directly in our world to model and reason about it. This renders an understanding of statistics important to developing citizens, consumers, and workers. From news reports to medical studies, citizens are surrounded by data that are needed to make decisions affecting the quality of their lives. For example, without an understanding of how samples are taken and how data are analyzed and communicated, citizens cannot effectively

participate in most of today's important political debates about the environment, health care, quality of education, and equity (Konold & Higgins, 2003). At work, people are required to understand the data that relates to their jobs. In one field, engineers are concerned with data on product quality, and in another, business people focus on costs, profits, and sales projections. This greater use of data in our information age provides a compelling reason to be statistically literate (Moore, 1990).

Many believe becoming statistically literate should begin in our schools. This puts the bulk of the responsibility onto the shoulders of teachers. The 2005 Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report is a curriculum framework that addresses pre-K–12 statistics education. The authors of the 2005 GAISE Curriculum Framework claim that statistically literate high school graduates will be comfortable handling quantitative decisions on the job and will make informed decisions about quality of life issues (Franklin et al., 2007). In addition, they state that the "...surest way to reach the necessary skill level is expanding these skills throughout the middle and high school years" (Franklin et al., 2007, p. 3). This new emphasis on statistics in pre-K–12 education makes it an available and timely topic to research, and as such confirms the importance of focusing on it as a topic of my study.

Despite statistics' importance, it only recently received emphasis in K–12 mathematics education (NCTM 1989, 2000; Shaughnessy, 1992). The 1989 and 2000 National Council of Teachers of Mathematics Standards (NCTM) influenced how it became a part of the K–12 curriculum. Prior to that time, most was taught at the college level. However, today there is a greater emphasis on statistics as evidenced by the standards-based curricula addressing it and standards, such as the 2005 New York State (NYS) Standards that ascribe 30% of their

seventh-grade curriculum to statistics and probability. Further, the Common Core State Standards for Mathematics (2010) incorporates understanding of the strand of statistics in grades 6 through 8. As previously stated, the 2005 GAISE Report presents a pre-K–12 curriculum framework for statistics education (Franklin et al., 2007). See Appendix B for a summary of this report that presents this curriculum framework for learning statistics. In 2006, The College Board also published Standards for College Success in Mathematics and Statistics focusing on alternative standards for middle school mathematics.

As a result of this greater emphasis, teachers need to know what and how to teach statistics in ways that will enable students to meet these standards. Unfortunately, few teachers have the specialized knowledge necessary to teach statistics (Shaughnessy, 1992). Understanding statistics involves a different type of thinking than other fields of mathematics. Whereas understanding most of the fields of mathematics entails deterministic thinking, statistics requires probabilistic thinking. This thinking contrasts what one usually experiences in mathematics in that it involves random or chance events that are challenging and not necessarily intuitive (Tversky & Kahneman, 1973). Thus, the teachers' knowledge gap presents a problem. How teachers understand statistics can affect the way they teach it and, therefore, students' understanding of it. In order to understand more about teacher knowledge of statistics, I chose to study teacher knowledge of the statistics they are charged to teach. My practicum findings and recent literature on teacher knowledge of statistics have influenced my dissertation study.

My practicum research showed that it is possible for a teacher to learn more about the statistics he or she is charged to teach while teaching from a standards-based curriculum. I define standards-based as curriculum written after the 1989 NCTM Standards were published.

Specifically, Mei, the teacher studied in my practicum research, learned more about the mean, the effects of variability of data on the mean, and the relationship between the shape of a distribution and the mean's location in it. This practicum study shaped my desire to learn more about the sense teachers make of the statistics they are charged to teach. With this in mind, I designed this dissertation study. While keeping the school, the curriculum, and the grade constant, I investigated the sense teachers make of the statistics that they teach. Specifically, I studied teacher knowledge of variability in data distributions, as explained below.

My focus has shifted from my practicum study to this research. It shifted from what is possible to learn to what teachers exhibit knowing. This change has been influenced by recent literature on what teachers know about the variability of data. This research includes a call from Shaughnessy (1992) for studies on teacher understanding of this important statistics topic. Statistics is a broad field with many concepts. Researchers are beginning to focus on variability as a key aspect of statistical thinking. Snee (1990) as quoted in Makar & Confrey (2004) came from a perspective of the quality control industry. He defined statistical thinking “as thought processes, which recognizes that variation is all around us and present in everything we do, all work is a series of interconnected processes, and identifying, characterizing, quantifying, controlling, and reducing variation provide opportunities for improvement” (p. 118). The focus on the variability of data in distributions as an important statistics topic is further addressed in my literature review.

The purpose of this study is not to look at the cause and effect of a standards-based curriculum that intentionally supports the learning of variability and data distributions. Instead, it is designed to describe teacher sensemaking of variability captured as brief pictures or snapshots of teacher knowledge. These snapshots will be taken using three formats: written

tasks, interviews, and classroom observations at different times respectively—during and after professional development and while teaching. Because each venue provides a different picture of what teachers know, I anticipate that a more complex description of teacher knowledge of variability might emerge. I discuss more about the type of teacher knowledge I wish to study in the theoretical considerations and the literature review in this paper.

Also influencing my shift from what teachers learn to what they know is the literature on teacher knowledge. Teacher knowledge has been studied using different methods—through written assessments, in professional development settings, and in the classroom (Mickelson & Heaton, 2004). Each perspective tells a different story about teacher knowledge and zones in on a specific type, for example, content knowledge, pedagogical content knowledge (Shulman, 1987), and knowledge useful for teaching (Ball et al., 2001). In my dissertation study, I investigated teacher content knowledge of the variability in data distributions. (Mae, a pseudonym, is the teacher in this study). How I studied this knowledge was shaped by the theories on knowledge in teaching (Shulman, 1987; Ball et al, 2001), and on Weick's (1995) theoretical sensemaking model—noticing, interpreting, and implementing.

### **Theoretical Considerations**

Teacher knowledge is complex and is delineated by researchers into various types or components. Shulman (1987) categorizes teacher knowledge into seven categories:

1. content knowledge
2. general pedagogical knowledge
3. curriculum knowledge
4. pedagogical content knowledge
5. knowledge of learners

6. knowledge of educational contexts
7. knowledge of educational ends

In his taxonomy, Shulman considers content knowledge a central feature of the knowledge required for teaching. He described teachers' comprehension of content knowledge as needing to be flexible, multifaceted, and adequate to impart alternative explanations of the same concepts or principles. Thus, content knowledge needed for teaching cannot be instrumental or lacking in depth and connectedness to other topics in the terrain (Skemp, 1987; Ma, 1999). Viewing content knowledge as foundational to teacher knowledge, he also values its usefulness in making judgments and taking action.

Emerging from Shulman's typology of teacher knowledge is his concept of teachers' pedagogical content knowledge (PCK). This knowledge encompasses the transformative action teachers must take "to transform the content knowledge he or she possesses into forms that are pedagogically powerful and yet adaptive to the variations in ability and background presented by the students" (p. 15). In his model of pedagogical reasoning and action, PCK, aligns with the aspect of transformation. Embedded in this type of teacher knowledge is the knowledge a teacher uses to prepare a lesson, to represent concepts, to select an instructional format, and to adapt to students' characteristics. I agree with Shulman that comprehension of a topic alone is not sufficient, and that the usefulness of teachers' knowledge is in its value to judge and take action. Based on this, I sought to study the knowledge Shulman states is first needed—content knowledge. In this research, I investigate it in multiple sites of teachers' work: professional development, lesson planning, performance tasks, and classroom teaching.

One reason to study the content knowledge of teachers is that research is beginning to build an understanding of how students and teachers understand variability in data

distributions. This leads me to believe that it might be productive to add to this growing body of knowledge by studying what in-service teachers understand about the variability of data, that is, their content knowledge. As is discussed in the literature review, more is yet to be known about teacher understanding of this topic. This is particularly true of in-service middle school teachers. For example, Canada's (2004) Evolving Framework that characterized elementary preservice teachers' thinking about variation could be useful for studying middle-school teachers. The data analysis chapter explains how I used it in this research.

Other researchers have discussed theories on teacher knowledge. Ball et al. (2001) bring to the forefront a view of teacher knowledge that is useful in practice. When teachers are engaged in the endemic uncertainties in classroom interactions, this knowledge is visible, for example, in the mathematical decisions they make as they manage routine and nonroutine problems. These researchers call this knowledge of mathematics pedagogically useful knowledge that is seen at the level of practice. More on this type of knowledge is discussed in my literature review.

In addition to focusing on Shulman's (1987) content knowledge for teaching, I would like to use Weick's (1995) model of sensemaking: what one notices, interprets, and implements, as a guide for my observations of teacher knowledge of statistics. Weick's (1995) sensemaking model has been used to study reform implementation. One example is the research of Drake (2006). In Drake's (2006) study, teachers' stories, specifically turning-point stories, were connected to their specific practices in the context of reform. Turning-point stories were teacher stories that involved a positive change in their attitude towards mathematics at a noticeable point in their lives. They experienced a turning-point that changed both their perception of their ability in mathematics, and their perception that mathematics

could be made sense of. Those teachers with turning-point stories used their turning-point experiences as lenses to make sense of the reform policy.

I plan to extend the use of Weick's (1995) sensemaking model to teachers' sensemaking of a statistics topic they teach—variability. The three sensemaking practices of what teachers notice, interpret, and implement would be used to operationalize their knowledge of variability in data distributions. My assumptions are that what the teachers notice had meaning to them; how the teachers interpret what they notice gives insight into their knowledge; and how they implement a lesson using them together gives a fuller view of teacher knowledge. Further, Weick's (1995) and Drake's (2006) sensemaking model aligns with my bidirectional view of studying teacher content knowledge. I wish to depict a bidirectional view of teacher content knowledge that connects their sensemaking prior to teaching with their sensemaking during teaching.

Noticing variability is an important aspect of statistical thinking. In Pfannkuch's and Wild's (2000, 2004) model, one type of fundamental statistical thinking includes attention to or consideration of variation. This is subcategorized as noticing and acknowledging variation. This is not the only type of thinking inherent to statistics. However, it is a foundational one that separates it from the general types of thinking that are the hallmarks of mathematical thinking, such as looking for patterns, abstracting, generalizing, specializing, and generating and applying algorithms (Shaughnessy, 2007). Thus, studying teachers' noticing the variability in data distributions is important. It might not be a mind-set necessarily developed in a teacher who majored in mathematics.

Studying how teachers interpret the variability of data is equally essential. What teachers know about this topic serves as a base from which they can draw in order to teach. In

the following literature review, research conducted on both pre- and in-service teachers is described. A common theme across this research is that there is more to learn about teachers' interpretation of variation. My dissertation study adds to this continuing discourse. It is expected that studying the knowledge teachers manifest on this statistics topic during their classroom implementation as well as in other sites of their work, will bring a fuller view of their content knowledge of variability to the dialogue.

## **Chapter 2**

### **Literature Review**

Ball et al. (2001) have written specifically about how policy and research have defined the knowledge needed for teaching mathematics and much of this section draws from their work. Ball et al. state that teacher knowledge has been defined in three ways: (a) by documents that list what teachers should know, (b) by characteristics of teachers, and (c) by the nature of teachers' knowledge. Policy primarily uses the first way to define teacher knowledge. Research predominantly utilizes the second and third definition to investigate teacher knowledge.

Policy has created documents that both generally and specifically define what teachers need to know. General documents that call for teachers' knowledge to be deep, connected, and conceptual include, for example, those published by the Interstate New Teacher Assessment and Support Consortium and the National Board for Professional Teaching Standards. The Conference Board of Mathematical Sciences and the National Council of Teachers of Mathematics call for specific knowledge that teachers need to know in mathematics. The documents produced by these organizations are lists of what teachers need to know. They usually identify topics beyond the curriculum and stress the need for teachers to have connections among ideas.

Research embraces the second and third way to define teachers' knowledge: as characteristics of teachers and the nature of their knowledge. In using the characteristics of teachers to define teacher knowledge, researchers sought to validate that the more mathematical knowledge teachers have, the more mathematical knowledge their students will

have. Accordingly, researchers counted courses and analyzed the relationships between the extent of teachers' mathematics coursework and their students' learning. This approach entailed counting the courses taken by teachers, the credits they earned, and the degrees they attained, each of which is considered to be a representation of teachers' mathematical knowledge. In this way of defining teacher knowledge, there is little attention paid to the content, scope, or the nature of mathematics. Teachers' exposure to the mathematical content is assumed to give them the resources needed for teaching.

According to Ball et al. (2001), the two major works in this approach are Begle's meta-analysis of studies in 1979 and Monk's longitudinal study of youth in 1994. In 1979 Begle conducted the National Longitudinal Study of Mathematical Abilities to examine factors that affect mathematics learning. His analysis showed that the relationship between the number of courses teachers had taken past calculus and student performance produced positive main effects on students' achievement in only 10% of the cases, and negative main effects in 8%. Ball et al. (2001) speculate that there are two possible explanations of this. One, it could be based on the compression of knowledge that accompanies advanced mathematical work. This might interfere with the unpacking of content that teachers need to do. Second, more coursework in mathematics might be accompanied by more experience with conventional approaches to teaching mathematics. In addition, Begle (1979) examined the relationship of teacher course work and their students' performance, which revealed that a major or minor in mathematics yielded positive main effects in 9% of the cases, and negative main effects in 4%. It is important to note that the greatest number of positive main effects was produced from the analysis of the relationship between the number of credits in mathematics methods courses and student performance—23% with only 5% negative effects. On the basis of his findings, Begle

(1979) claimed that advanced mathematical understanding contributed little to teacher effectiveness.

Monk (1994) reached similar conclusions to Begle's (1979) in that the number of mathematics courses teachers take makes a difference, but only to a point. He analyzed data from the Longitudinal Study of American Youth. Monk measured teacher preparation by a survey, which documented mathematics and science teachers' overall educational level and years of teaching experience. He found that five courses in mathematics, independent of the specific content covered, was the threshold beyond which few effects accrue. In addition, as with earlier studies, he found that there was no effect on student performance based upon whether a teacher had majored in mathematics. Finally, as in Begle's (1979) analysis, Monk uncovered significant effects for courses in undergraduate mathematics pedagogy: that they contributed more to student gain than courses in undergraduate mathematics.

Based on these studies that focus on teachers' coursework as a proxy for knowledge, knowing how many mathematics courses a teacher has taken does not enable us to predict whether the teacher would be able, for example, to untangle the complexities in teaching multiplication with decimals to fifth graders (Ball et al, 2001). While focusing on teachers' mathematical qualifications affords certain information, it does not reveal the nature of teachers' mathematical knowledge. This gap is filled by the third approach to defining mathematical knowledge for teaching, which entails a closer analysis of the nature of teachers' mathematical knowledge.

The third approach to define teacher knowledge builds on the second approach in that it acknowledges the importance of the content of teachers' knowledge. However, it also prominently includes a qualitative focus on the nature of teachers' knowledge. Using this

method, researchers probed closely at the mathematical knowledge of teachers on specific mathematics topics rather than measuring second order indicators of knowledge. The studies using this approach differed from the second approach in that many of them are qualitative and used interviews to explore teachers' knowledge rather than surveys used in the first approach.

A host of these studies on specific mathematical topics opened up the idea of mathematical knowledge for teaching. This work focused on both the substantive knowledge of teaching and the knowledge of mathematics. Researchers working in this approach often use methods that probe teachers' knowledge and that situate questions in and around those that might arise in teaching. As discussed in the theoretical section, one of the most significant contributions with this closer focus on teacher knowledge has been a new conception of subject matter knowledge for teachers called *pedagogical content matter knowledge*. This is a special kind of teacher knowledge that links content with aspects of teaching and learning (Shulman, 1986).

One example is the work of Liping Ma. Ma (1999) did a comparative study of elementary school teachers' mathematical knowledge. She compared Asian and American elementary teachers' answers to interview tasks. These tasks were taken from the elementary mathematics curriculum. They included for example finding the area of a rectangle and dividing fractions. The results of her study showed that the American teachers lacked the profound understanding of fundamental mathematics that is necessary to teach the elementary grades. She based her findings solely on the teachers' performance on these interview tasks. She did not study their teaching.

Inasmuch as this approach has been enlightening regarding teacher knowledge, Ball et al. (2001) believe that it leaves gaps because there is a distance between these studies of

teacher knowledge and teaching itself. These researchers sought to resolve these gaps by creating an alternative approach to studying teachers' mathematical knowledge for teaching. They looked toward teacher practice. They support a view of teacher knowledge that focuses on whether and how teachers are able to use mathematical knowledge in the course of their work. These researchers look to define teachers' mathematics knowledge as that which is pedagogically functional—what they know, how they know it, and what they are able to mobilize (p. 451). Thus, they believe that the knowledge of mathematics necessary for teaching needs to be redefined from one about teachers and what teachers know to one about teaching and what it takes to teach.

These different perspectives on studying teacher knowledge offer important insights. However, they might not offer the most complete picture of what teachers know. They do not include what is displayed in the moments of teaching. The knowledge that is displayed before teaching might not portray the knowledge that is displayed when teaching. They might be different, similar, or complementary to each other. Both are important in understanding teacher knowledge.

I wish to combine the view of teacher knowledge prior to teaching with the view of teacher knowledge while teaching. This bidirectional view would be in contrast to the unidirectional view of teacher knowledge—the direction that what one knows before teaching points towards what one would know in teaching. My bidirectional view would indicate a twofold approach to studying teacher knowledge. It could depict what a teacher knows before teaching and what a teacher knows while teaching. I believe seeing teacher knowledge in this way could present a more complete picture of what a teacher knows about a topic.

In an attempt to use this bidirectional view of teacher knowledge, I needed a way to operationalize the large construct of teacher knowledge. I saw that teachers' sensemaking about a topic could be a proxy for the knowledge they had about it. In this vein, I sought the help of a sensemaking model. Weick's (1995) model of sensemaking speaks of individuals making sense of something through three sensemaking practices—noticing, interpreting, and implementing. Drake (2006) used Weick's (1995) model to study teachers' sensemaking of educational reform. (See Chapter 4 for further discussion on Weick's (1995) and Drake's (2006) work with sensemaking.) In my dissertation research, I used the sensemaking model to study teacher knowledge of an important statistics topic—variability in data distributions.

Using the sensemaking model (Weick, 1995; Drake, 2006), I plan to look at a teacher's noticing and interpreting prior to teaching and also at her lesson implementation. To get a bidirectional view of what she knows, the one focus will be on teacher's sensemaking (noticing and interpreting) prior to teaching—professional development, lesson planning, and performance tasks with interview—and the other focus will be on her lesson implementation. Together these two views could portray a more complete picture of what a teacher knows.

In short, a bidirectional view of teachers' knowledge as seen through their sensemaking practices prior to and during teaching will portray a more complete picture of what they know. In this way I chose to study teacher content knowledge of the variability in data distributions. To inform my work I did a review of the literature on teachers' knowledge of variability in data distributions. This research is discussed in the following section.

### **Teacher Understanding of Variability**

In the *Handbook of Research on Mathematics Teaching and Learning*, Shaughnessy (1992, 2007) summarized the research and development of teachers' understanding of statistics

and gave recommendations for future research. In the first edition of the handbook, he suggested that classroom teachers' conceptions of probability and statistics be studied. In the handbook's second edition, he acknowledged the progress that has been made in this area. The literature review that follows summarizes these studies. Two of the issues in future research that he calls for are pertinent to my study. First, he calls for clarifying the term *distribution*. This word, used frequently by statisticians and statistics educators, might be used to refer to a single distribution of data, to a sampling distribution of statistics, or to a probability distribution. In future research he states it would be beneficial to be more precise in the use of the term *distribution*. Second, he recommends more research is needed on teachers' conceptions of statistics. He claims that teachers have the same difficulty with statistical concepts as the students, and suggests that research needs to find ways to aid teachers in developing their statistical knowledge and thinking. This confirms my desire to understand teacher statistical knowledge perhaps to direct future teacher development.

Shaughnessy (2007) also outlined some implications from research for the teaching of statistics. He recommends that:

- Variability is emphasized as one of the primary issues in statistical thinking and statistical analysis.
- Comparisons of data sets should be introduced much earlier with students, prior to formal statistics.
- Students' intuitive notions of center and variability should be built upon.
- The role of proportional reasoning in the connections between populations and samples should be more explicit.

I agree with Shaughnessy's statement that teachers have the same difficulty with statistical concepts as students and his above-outlined implications. Based on this, I chose to study teacher knowledge of variability in data distributions through comparing data distributions that would include their intuitive notions. A more detailed review follows of the literature on teachers that influenced my dissertation study. These studies are categorized into pre- and in-service teachers and further by their level of teaching—elementary, middle, and secondary—in order to highlight who were studied and in what context they were studied. My research seeks to fill the gap in studying teachers. It investigates in-service middle school teachers. Their knowledge of variability is yet to be studied in multiple aspects of their work. This includes their professional development, lesson planning, performance tasks, and their classroom teaching. As the following literature review indicates, for the most part their knowledge of variability has been left unexamined.

### **Preservice Teachers**

**Preservice Elementary Teachers.** Canada (2004) studied how elementary preservice teachers (EPSTs) expected, displayed, and interpreted variation. He studied these within three statistical contexts: repeated sampling, data distributions, and probability outcomes. He used pre- and posttests with classroom interventions comprised of hands-on activities, computer simulations, and discussions with multiple opportunities to attend to variation. His results showed that there was overall improvement regarding what the preservice teachers expected and why with regards to variation. In general, they demonstrated stronger intuitions about variation. Their predictions were more realistic, their expectation of variation more balanced, and their descriptions of distributions were more rich and robust in their recognition of variation and distribution as important concepts. Canada's (2004) Evolving Framework

characterized elementary preservice teachers' thinking about variation. See Table 2.1 for an outline of his framework.

Evolving Framework	
[1] Expecting Variation	
A] Describing What is Expected	
i) Concerning Expected Value	
ii) Concerning Repeated Values	
iii) Concerning Range or Extremes	
B] Describing Why (Reasons for Expectations)	
i) Involves Possibility or Likelihood	
ii) Involves Experiential Reasoning	
iii) Involves Proportional Reasoning	
iv) Involves Distributional Reasoning	
[2] Displaying Variation	
A] Producing Graphs	
i) Technical Details	
ii) Characteristics of the Distribution	
B] Evaluating and Comparing Graphs	
i) Focus on Average	
ii) Focus on Range or Extremes	
iii) Focus on Shape	
iv) Focus on Spread	
C] Making Conclusions about Graphs	
i) Emphasizing Decisions in Context	
ii) Emphasizing Consistency or Reliability	
iii) Emphasizing Level of Detail & Usefulness	
[3] Interpreting Variation	
A] Causes and Effects of Variation	
i) Definitions & Descriptions	

*Table 2.1 Canada's (2004) Evolving Framework for Elementary Preservice Teachers' Thinking About Variation*

Table 2.1 (cont'd)

ii) Examples
B] Influencing Expectations and Variation
i) Naturally Occurring Causes
ii) Physically Induced Causes
C] Effects of Variation
i) Effects on Perception
ii) Effects on Decisions
D] Influencing Expectations and Variation
i) Quantities in Sampling
ii) Number of Samples

Table 2.1 Canada's (2004) *Evolving Framework for Elementary Preservice Teachers' Thinking About Variation*

Canada's (2004) *Evolving Framework* provides a lens through which three different *aspects* of EPSTs' understanding of variation can be viewed. In his work the three aspects address how EPSTs reason in terms of *expecting*, *displaying*, and *interpreting* variation. (See Table 2.1.) Each aspect has *dimensions* that are lettered A–D on the table. Following this, each dimension has *themes* that Canada (2004) focused upon in analyzing his EPSTs' thinking about variation. These themes are numbered i–iv on the table.

His three main aspects aligned with the type of tasks and activities he provided: repeated sampling, data distributions, and probability outcomes. Since my research solely involved data distributions, I thought his two dimensions of Displaying and Interpreting variation would be applicable. (See Table 2.1.) The following explains in more detail these parts of his framework.

Of particular interest to this study is the second aspect of reasoning about variation: Displaying. Canada (2004) found that there were three dimensions to this aspect of his EPSTs' thinking about variation that emerged. They are:

1. producing graphs
2. evaluating and comparing graphs
3. making conclusions about graphs

The activities in my research predominantly addressed the second dimension of Evaluating and Comparing Graphs. Within this dimension Canada (2004) looked for four themes in his EPSTs' thinking about variation—(a) *average*, (b) *range or extremes* (which Mae called outliers), (c) *shape*, and (d) *spread*. These themes are considered relevant to this research because I used problems that involve comparing and analyzing data distributions.

Another aspect of his EPSTs' thinking about variation seemed to pertain to my research—Canada's third aspect: Interpreting Variation. This is based on my research focusing on knowledge through interpreting. In particular, two dimensions of this aspect could apply to my work. The first dimension, *causes and effects of variation* includes the theme of *defining and describing*. This theme seemed to be helpful because through defining and describing, I could see how the teacher interprets variability. The second applicable dimension is that of *influencing expectations and variation*. Of particular interest to my study are its themes of *naturally occurring* and *physically induced* causes of variation. These themes could apply because the measurement data used in my study lend themselves to discussing the causes of variation—natural or induced.

In contrast to Canada's (2004) elementary preservice teachers, mine are in-service and in the middle school, but Canada's (2004) Evolving Framework is nonetheless useful. Based

on how it appeared helpful, I used it to guide my data analysis. See Chapter 4 for further discussion on its applicability, and Chapter 8 for how his framework proved helpful in characterizing in-service middle school teachers' content knowledge of variability.

**Preservice Middle School Teachers.** Sorto (2004) asked what are important aspects of statistical knowledge necessary for teaching at the middle school level in the United States, and what do preservice teachers know about the various aspects of statistical knowledge of teaching? Her research focused on the following aspects in the domain of statistical knowledge:

- reading, interpreting, and inferring data using graphical displays such as histograms, line plots, stem and leaf plots, and tables
- recognition, description, and use of shapes of data distributions
- development and use of measures of center and spread

For each aspect Sorto (2004) measured different levels of performance. For graphical displays, three levels were considered: extracting information from the data, finding relationships in data, and moving beyond the data (Friel et al., 2001). For distribution, measures of center, and spread, the three levels of performance or cognitive outcomes are statistical literacy, reasoning, and thinking (delMas, 2002; Garfield, 2002). In Sorto's (2004) study, statistical literacy referred to reasoning with statistical ideas when asked why or how results are produced. For example, knowing what type of data leads to a particular graph or statistical measure, knowing what factors influence the shape of a distribution, selecting the appropriate measure of center, or interpreting what these measures reveal about the data. The third level, statistical thinking refers to the application of students' understanding to real-world problems, to critique and evaluate the design and conclusion of studies, or to generalize

knowledge obtained from classroom examples to new and somewhat novel situations. For measures and distribution, this might mean using them to make predictions and inferences about the group to which the data pertain.

As one might expect, Sorto's (2004) results showed that preservice teachers performed better in the domain of pure statistical knowledge (65.72%) than the domain where they had to apply this knowledge to teaching (45.14%). She also found that none of these teachers showed mastery or near mastery level of correctness for all items in either domain of knowledge. For pure statistical knowledge, Sorto (2004) found that prospective teachers performed better at the lowest level of performance—statistical literacy that involved mainly extracting information from a graph and recognition, identification, or computation. At the higher levels, statistical reasoning and thinking, they did progressively worse.

One item that was very telling measured the ability of prospective teachers to identify errors in student responses. See Table 2.2. In solving this problem the teachers mistook the frequencies as data observations and calculated the median.

One middle school class generated data about their pets shown below. Students were talking about the data and one said:

“The mode is dogs, the median is duck, and the range is 1 to 7.”

If you think the student is right, explain why.

If you think the student is wrong, identify the mistake(s).

Pet	Frequency
Bird	2
Cat	4
Cow	2
Dog	7
Duck	1
Fish	2
Goat	1
Horse	3
Rabbit	3

*Table 2.2 Item Used to Assess Teachers' Statistical Knowledge of Teaching*

In general, Sorto (2004) found that preservice teachers know that it is possible to have many data sets with the same mean. However, only a small percent can justify this statement with an argument that relies on both the algorithm and the concept of the mean as a balance point. A smaller percentage of preservice teachers are able to create a distribution with a specific mean that is not a whole number. The majority of these prospective teachers are bothered by having a nonwhole number as an average. Prospective teachers do not think of what the measures of center and of spread tell you about the data when trying to find them. Instead they reach for a procedural method to get an answer. They could neither estimate the

mean of a data set without reaching for the algorithm, nor convince a child that there could be several data sets with the same mean, or explain what an average of 3.5 people represents.

The errors in thinking that these preservice teachers exhibited, whether in answering questions focused on their content knowledge, or on their ability to evaluate student work, depended upon their content knowledge. While this might be obvious in the first type of question, the second type that targets teachers' pedagogical content knowledge also draws upon teacher's content knowledge. For example, the knowledge the teacher had about variability of categorical data affected her ability to judge a student's statement about it. These findings confirm my assumption that the statistical knowledge of preservice teachers might be calculation driven.

Although Canada (2004) and Sorto (2004) studied preservice teachers, there are some differences in their studies. Whereas Canada (2004) investigated preservice elementary teachers' thinking about variation, Sorto (2004) studied preservice middle school teachers' understanding of statistical topics that included measures of variation, such as the range. In addition, Canada's (2004) study was geared toward developing an understanding of variation in his EPSTs. This was in contrast to Sorto (2004) who did not specifically have her subjects study the area of statistics as part of her research. This is not to say that statistical concepts were not a part of her subjects' required coursework. Sorto's (2004) research is one study conducted on middle-school preservice teachers. My research adds to the understanding of teacher knowledge of variability of data by studying in-service middle-school teachers. In contrast to these studies, Makar and Confrey (2005) chose to study preservice secondary teachers. Their research design included interview tasks that focused on variability of data distributions and that took place before and after a fifteen-week preservice course.

**Preservice Secondary Teachers.** Makar and Confrey (2005) studied prospective secondary mathematics and science teachers' articulation of the notions of variations in their own words. This is a qualitative study that describes how preservice teachers articulated ideas of variation as they compared two distributions of data in terms of the relative improvement in test scores. Through interviews with tasks before and after a fifteen-week preservice course, the researchers documented the different types of language used to express variation. They found that these teachers used both standard and nonstandard language to express rich views of variation. Standard statistical language included *proportion* or *number improved*, *mean*, *maximum/minimum*, *sample size*, *outliers*, *range*, and *shape* (e.g., skewed, bell shaped). Categories of nonstandard language emerged in two overlapping areas—variation and distribution. For variation the nonstandard language included terms such as *spread*, *clustered*, *clumped*, *grouped*, *bunched*, *gathered*, *spread out*, *evenly distributed*, *scattered*, or *dispersed*. For distribution it involved low-middle-high clumps called *triads*, *modal clumps* (middle portion of the distribution; Konold et al., 2002), and *distribution chunks* (e.g., handful of students who improved the most). The authors refer to these locutions as “variation talk.” Of particular interest is that their preservice teachers' nonstandard language revealed strong relations between expressions of variation and expressions of distribution. The authors conjectured that nonstandard statistical language might naturally integrate these concepts, or that the respondents tended to link variation and distribution so inseparably that when they noticed one in a graph they noticed the other. On the other hand, if they did *not* perceive one, they did not perceive the other.

A further analysis of the variation talk revealed past participles (such as clustered, clumped, grouped, bunched, gathered, spread out, evenly distributed, scattered, or dispersed),

that implied attention to the shape as a pattern of variability. This is in contrast to the conventional statistical terms like *range* or *standard deviation* that are measures. The other three types of variability talk included aspects of variability that involved partitioning the distribution to examine a subset or chunk of the data, for example, low-middle-high, modal clump (Konold et al., 2002), or other meaningful clump. As a result of their study, the authors see that there are more than just the two perspectives of distribution that are usually discussed in literature: single points and aggregate. They showed that a third perspective arose—partial distributions or “mini-aggregates”—that deserves further study.

This study is significant because the variation talk used by their participants is similar to the language used in other studies of learners’ conception of variation and distribution (Bakker et al. 2004; Canada, 2004; Hammerman & Rubin, 2004).

Further, Makar’s and Confrey’s (2005) conjecture that nonstandard statistical language might naturally integrate the concepts of variation and distribution—based on their teachers tending to perceive these concepts in graphs together (or not together)—guided my decision to use data distributions in the interview tasks. In addition, their conjecture influenced my decision to use a curriculum during professional development that focuses upon variability in data distributions. I expected that comparing and describing data distributions would be a natural setting for teachers’ knowledge of variability to be expressed. Their study has also made me aware of the value of nonstandard language that teachers might use in making sense of the variation of data when analyzing and comparing distributions. Finally, their suggestion to research how teachers understand concepts of variation and distribution confirmed the need to study in-service teachers. In this dissertation study, I sought to add to the research on how

teachers understand the concepts of variation and distribution by studying in-service teachers both during professional development and during their teaching.

### **In-Service Teachers**

**In-Service Elementary Teachers.** Mickelson and Heaton (2004) did a qualitative research of a third-grade teachers' statistical reasoning about data and distribution in the applied context of a classroom-based statistical investigation. They explore the complexity of teaching and learning statistics and offer insight into the role and interplay of a teacher's statistical knowledge and context. They illuminated the problems a third-grade teacher had in thinking on her feet statistically. During a lesson the teacher's statistical reasoning played a central role in orchestrating her class's investigation. Mickelson and Heaton (2004) showed that when the knowledge of variability is learned as a disconnected isolated entity in professional development, the transfer of that knowledge to an applied setting is problematic and sporadic. Of particular interest to my study is the authors' statement, "The importance of connecting the teaching and learning of statistics for teachers to their specific purposes for the study of specific K–6 content and in particular classroom contexts cannot be understated" (p.21). This statement influenced my decision to use a middle-school curriculum in my professional development sessions. I believe it applies to 7<sup>th</sup> to 12<sup>th</sup> grade education as well. In this way, the statistics concepts involved in my professional development sessions are specified for the work the teacher can do in her seventh-grade classroom.

**In-Service Secondary Teachers.** Hammerman and Rubin (2004) studied secondary teachers' reasoning in the presence of variability in a professional development setting. They found that people work with the variability of data to make it more manageable and comprehensible. This is especially true when using software aspects that make the data easier

to view and manipulate. These researchers outlined steps for understanding how their teachers reduced variability in their analysis:

1. analyzed the data in bins (similar to a histogram)
2. used circle graphs to look at ratios, independent of counts—allowing them to look at the pattern of ratios
3. focused on the general pattern of ratios ignoring small fluctuations, i.e., looking at the signal in the data and attempting to eliminate the noise
4. compared general trends across protocols

Hammerman's and Rubin's (2004) work illuminated the tension between reducing variability to deal with the complexity of data with the risks of making claims that would not be true if all the data were included in the analysis. They also discussed strategies applied to manage the complexity of data using the software tool:

1. analyzing data in bins
2. using cut points such as the mean or median
3. using slices of data, which seemed to alleviate the tension between using bins to reduce variability and the number of points being attended to, and expanding the scope of the data being considered to reach a comfortable minimum
4. binning in the context of covariation

Some of their noteworthy findings included:

- When teachers had the opportunity to view and manipulate individual data points in a data set using computer software, they became less comfortable with the measures of center that they knew as containing information about all

of the data.

- They came to believe that the teachers lacked a perspective on data that would help them see the mean as a representative value. This stance regards any distribution as due to a noisy process (Konold & Pollatsek, 2002) and, thus, conceptualizes the mean as a relatively stable measure of the signal within the noise.
- Teacher-participants occasionally focused solely on variability, especially when central tendency seemed unimportant.

Makar and Confrey (2004) also studied secondary teachers. They studied their statistical reasoning after a six-month professional development that focused on building conceptual understanding and experience with powerful statistical ideas. Inferential statistical concepts were discussed. However the majority of more advanced topics (e.g., t-tests, confidence intervals, null hypothesis, and p-values) were experienced through computer simulations on a conceptual not a formal level. The professional development experiences involved comparing two sampling distributions of student results on state tests. They were curious to see how teachers viewed difference in measures of the distributions, i.e., whether small differences in quantitative measures indicated a tolerance for variation. Their methodology included qualitative analysis of four interviews wherein the teachers described the relative performance of the distributions of male and female students. Pre- and posttest quantitative analysis of statistical content knowledge provided triangulation.

Makar's and Confrey's (2004) results illuminated teacher understanding of variation in the following ways: (a) within distribution—variability of data, (b) between distributions—variability of measures, and (c) how they distinguished between these two types of variation.

Makar and Confrey (2004) found that all teachers clearly recognized variation within a single distribution, and that they compared range and standard deviation between the two distributions. They found that some respondents had a deterministic view of the descriptive measures while others indicated some tolerance for variability in the mean. Their research raises an important issue about variation—which variation are we referring to when we compare two distributions—variation within each group or between groups thus considering variation within the measures themselves? They found that comparing distributions to be a fruitful arena for expanding teachers' understanding of distribution and conceptions of variability. This finding influenced my decision to use comparing distributions as part of my interview tasks.

These three studies focused on in-service teachers and involved professional development that included the variability of data distributions. However, these studies differed in the level of teachers they researched and somewhat in the context in which they studied them. Whereas the research done by Hammerman and Rubin (2004) and Makar and Confrey (2004) took place only in the professional development setting, Mickelson and Heaton's (2004) research was conducted in the classroom. The context in which a study of teacher knowledge takes place is significant as there is a difference in what teachers' show they know when they are teaching (Ball et al., 2001). Studying what teachers' know about the variability of data as exhibited in their teaching is important as this is when the students interact with this statistical concept. What teachers' show they know in the classroom might affect what students' learn. In this dissertation study, I chose to study teacher content knowledge of variability in data distributions as they are teaching in the classroom and during professional development. Specifically, I studied middle school teachers.

## **Summary**

In summary, in the area of teacher understanding of variability in data distributions, a small amount of research on teachers' understanding of variability has contributed to a decent amount of findings. These studies influenced my dissertation study. The findings of researchers discussed in the literature review have influenced and refined both who I chose to study and how I studied them. My research is motivated by the subjects, who for the most part are missing in the current research, and by the context in which current studies were conducted. My study differs in that rather than studying teachers in training or in a course, I studied middle school teachers in the context of their school lives. My findings might not differ significantly from that of other researchers who studied elementary and secondary teachers. However, it is intended that my findings will contribute to the current body of research by investigating an important group of teachers, middle school in-service teachers, in a context that is real to their lives as teachers.

To guide my analysis, I used Canada's (2004) Evolving Framework of EPSTs' thinking on variation and Makar's and Confrey's (2005) research on teachers' "variation talk." In addition, I used a curriculum framework for pre-K—12 statistics education, the 2005 GAISE Report (Franklin et al., 2007), and definitions from formal statisticians as references in my analysis. The following section explains where my dissertation study is situated in the field of statistics, and how the 2005 GAISE Curriculum Framework (Franklin et al., 2007) and statisticians' definitions of key terms about variability can help interpret my findings.

### **Situating This Research in the Field of Statistics**

The terrain of knowledge in the field of statistics is large. In this section, the parts of this terrain that relate to the present dissertation study are discussed. For clarity, this

section also includes definitions of important terms used throughout this dissertation, and situates them within the 2005 GAISE Curriculum Framework (Franklin et al., 2007). The purpose of this section is to provide the reader with a brief tour of the statistics ideas involved in the study of variability that will aid in the reading of the findings of this dissertation study.

### **Statistical Investigation**

Statistical problem solving is an investigative process. Throughout this process, variability is a main focus. This statistical investigation involves four parts:

1. posing a question
2. collecting the data
3. analyzing the data
4. interpreting the results

The focus of this research is on the third part of the process, analyzing the distribution(s) of data. The following concept map in Figure 2.1 outlines the ways in which data distributions can be analyzed. The concept map is adapted from the Connected Mathematics Project's (CMP) text, *Data Distributions*, which was written for teachers of middle school (Lappan et al., 2006). Some of the terminology used in that text might not be that of an expert statistician. Therefore, the terms in the concept map that are used in this dissertation study are further defined to represent the more precise and standard definitions acceptable to statisticians.

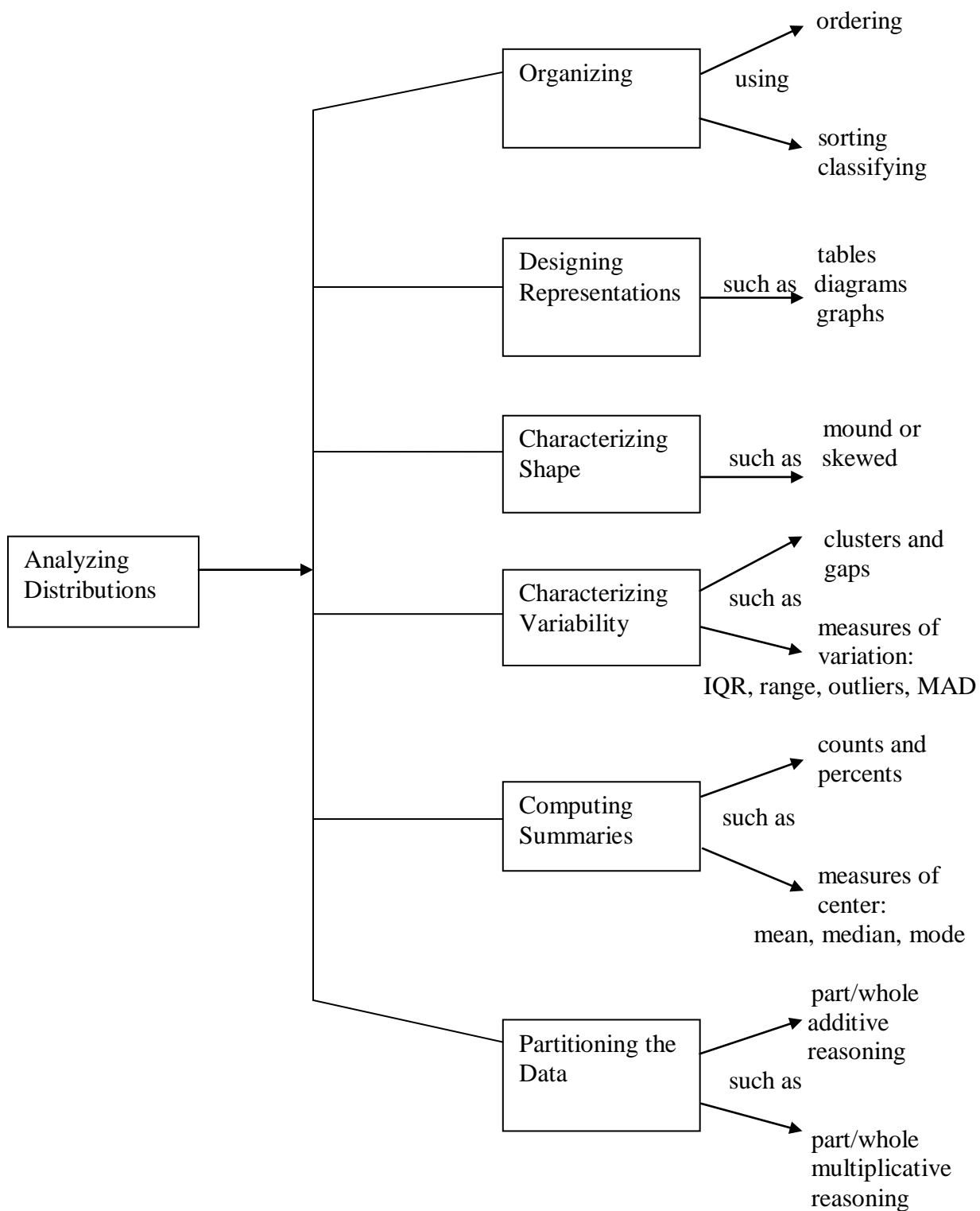


Figure 2.1  
Concept Map Adapted From “Doing Meaningful  
Statistics—Central Statistical Ideas for Data  
Distributions” (Lappan et al., 2006, p. 5)

First is an explanation of the concept map viewed as a list. Similar to the map, the list shows the ways in which data can be analyzed. They include:

1. Organizing the data  
Using: ordering, sorting/classifying
2. Designing representations  
Such as: tables, diagrams and/or graphs
3. Characterizing shape  
Such as: mound shaped or skewed
4. Characterizing variability  
Such as: clusters and gaps  
measures of variation:  
quartiles, including interquartile range, MAD  
range  
outliers
5. Computing summaries  
Such as: counts or percents  
measures of center:  
mean  
median  
mode
6. Partitioning the data  
Such as: part-whole (additive) reasoning  
relative frequency (multiplicative) reasoning

This list and the concept map in Figure 2.1 depict parts of the process of analyzing distributions. You can be engaged with all or some of the parts of the process in your analysis. One possible scenario of the process is described in the following.

Most likely in the beginning of analyzing the distribution, the data is organized. Once that is accomplished, a suitable way to represent the data is decided. Representing the data might include a table and its related graph or a diagram. From there the distribution is characterized in no special order either by its shape, variability, or by computing summaries that include the following measures of center: mean, median, or

mode. Another way the distribution is analyzed is by partitioning. When the distribution is partitioned into sections, either additive or multiplicative reasoning is used. For example, the percent of data can be described in these sections. Notwithstanding all of these potential parts of the analysis, the key idea in analyzing data is accounting for variability with the use of distributions (Franklin et al., 2007, p. 12).

### **Clarification of Key Terms**

For the purpose of clarification in this dissertation study, some of the terms in this concept map are explained. They are key terms in analyzing data. One term is *distribution*. Another is *variability* with its measures: range, interquartile range (IQR), variance, standard deviation, Mean Absolute Deviation (MAD), and outlier. The definitions of these terms are gleaned from statistics education texts ranging from college level to the CMP text, *Data Distributions* (Lappan et al, 2006) used in this dissertation study, and websites of mathematics/statistics terminology.

A distribution of a data set tells us what values a variable of interest takes and how often it takes these values. The overall distribution of a data set can be described. Graphs can be used to help clarify the distribution of a data set. Unlike individual cases, distributions have properties, such as measures of center, (i.e., mean, median, mode), variability (e.g., outliers, range), or shape (e.g., clumps, gaps). Distributions can be described by measures of center and variability (Lappan et al., 2006. p. 89).

*Variability* is the key idea in analysis of data and is the main focus of this dissertation study. When discussing variability, statisticians use the terms *variability*, *variation*, and *dispersion* interchangeably. To a statistician these terms are considered

synonymous. *Variability, dispersion, or variation* in a variable provides a measure of how far away from the center the data tend to be.

One measure of variability or dispersion is the *range*. The range is the difference between the most extreme observations; or more specifically the difference between the largest and smallest data points. Another measure of variability is the *interquartile range* (IQR) the difference between the upper and lower quartile.

Another measure of variability is the *Mean Absolute Deviation* (MAD). It is the absolute distance of each data value from the mean. It is an indicator of spread based on all the data and provides a measure of absolute variation in the data from the mean. It also serves as a precursor to standard deviation (Franklin et al., 2007). The formula used to find the MAD is:

$$\frac{\text{Total Distance From the Mean for All Values}}{\text{Number of Data Values}}$$

There are two other commonly used measures of dispersion. These are the *variance* and the *standard deviation*. The variance of a set of observations is the average squared deviation of the data points from their mean. The variance is determined by:

$$\frac{\text{Total of the Squares of the Deviations of the Observations From Their Mean}}{\text{Total Number of Observations}}$$

The standard deviation of a set of observations is the positive square root of the variance of the set (Aczel, 1996, p. 18–9).

An *outlier* is another key concept when describing distributions. While the other aspects used to describe distributions (shape, center, and variability) focus on the overall patterns in the data, outliers describe deviations from the pattern (Reading & Reid, 2006). The *Data Distributions* text defines an outlier as an unusually high or low data value in a

distribution (Lappan et al., 2006, p. 92). A more precise definition of an outlier is an observation that is numerically distant from the rest of the data. Grubbs (1969, quoted in “Outlier,” n.d.) defined it as:

An outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs.

One method to define an outlier is an observation that is outside the range of the following formula, if  $Q_1$  and  $Q_3$  are the lower and upper quartiles respectively and  $k$  is some constant:  $[Q_1 - k(IQR), Q_3 + k(IQR)]$ .

Aczel (1996) describes the point that is a distance of  $1.5(IQR)$  above the upper quartile as the upper inner fence, and the lower inner fence as  $1.5(IQR)$  below the lower quartile. These are guidelines for suspected outliers. Possible outliers might be at the outer fences—distances of  $3(IQR)$  above or below the upper or lower quartile.

## **2005 GAISE Report**

The 2005 GAISE Report is a coherent curriculum framework for statistics education for grades pre-K–12 (Franklin et al., 2007). It breaks down learning statistics into three levels: Level A, Level B, and Level C. It connects these levels to the statistical investigation process. All four parts of the statistical problem solving or investigative process are developed at all three levels, but the depth of understanding and sophistication of methods increases across these levels.

For the third part of the statistical investigation process, Analyzing the data, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) depicts a certain depth of understanding and sophistication of methods across each of the Levels—A, B, and C. These are shown in the table that is adapted from the report. Also included in Table 2.3 are

the depth of understanding and sophistication of methods across the levels for the nature of variability and the focus on variability.

<b>Process component</b>	<b>Level A</b> Univariate categorical data analysis	<b>Level B</b> Univariate numerical data analysis	<b>Level C</b> Bivariate data analysis
Analyze data	<p>Use particular properties of <i>distributions</i> in the context of a specific sample</p> <p>Display variability within a group</p> <p>Compare: -individual to individual  -individual to group</p> <p>Beginning awareness of group to group</p> <p>Describe a distribution</p> <p>Use tools for exploring distributions, including</p> <ul style="list-style-type: none"> <li>• bar graph</li> <li>• dotplot</li> <li>• stem and leaf plot</li> <li>• scatterplot</li> <li>• tables (using counts)</li> <li>• mean</li> <li>• median</li> <li>• mode</li> <li>• range</li> </ul>	<p>Learn to use particular properties of <i>distributions</i> as tools of analysis</p> <p>Quantify variability within a group</p> <p>Compare group to group in displays</p> <p>Compare two or more distributions using graphical displays and numerical summaries</p> <p>Use more sophisticated tools for summarizing and comparing distributions, including</p> <ul style="list-style-type: none"> <li>• histograms</li> <li>• IQR and MAD</li> <li>• five-number summaries and boxplots</li> </ul>	<p>Understand and <i>use distributions</i> in analysis as a global concept</p> <p>Measure variability: -within a group  -between groups</p> <p>Compare group to group using displays and measures of variability</p> <p>Identify appropriate ways to summarize numerical or categorical data using tables, graphical displays, and numerical summary statistics (including outlier analysis)</p>

Table 2.3 Table Adapted From the 2005 GAISE Report

Table 2.3 (cont'd)

	Observe association between two variables	Acknowledge sampling error  Some quantification of association; simple models for association	Describe and quantify sampling error  Quantification of association; fitting of models for association
Nature of variability	Measurement variability  Natural variability  Induced variability	Sampling variability	Chance variability
Focus on variability	Variability within a group	Variability within a group and variability between groups  Covariability	Variability in model fitting

Table 2.3 Table Adapted From the 2005 GAISE Report

Source: Franklin, C., Kadar, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria, VA: American Statistical Association.

Although the 2005 GAISE Curriculum Framework is written for grades pre-K–12, its levels are not meant to be grade based. It states:

Statistical education should be viewed as a developmental process. To meet the proposed goals, this report provides a framework for statistical education over three levels. If the goal were to produce a mature practicing statistician, there certainly would be several levels beyond these. There is no attempt to tie these levels to specific grade levels...they are based on development in statistical literacy, not age. (p.13)

The levels are based upon experience. For example, a learner with little to no statistical experience will most likely need the statistical experiences of Level A. On the other hand, a high school graduate with the appropriate statistical experiences could operate at a depth of understanding and the sophistication of methods at Level C. In addition, the work from the various levels assumes and develops the concepts from the lower levels.

The 2005 GAISE Report (Franklin et al., 2007) has placed these key terms in their curriculum framework. Specifically for the learner, the formal treatment of these key terms is recommended to begin at the various GAISE Levels. For example, if you were to plan statistical experiences for a new learner, the key terms, *distribution*, *variability*, and *range* would be appropriate to treat in a formal way at Level A. Similarly, according to the 2005 GAISE Curriculum Framework, a learner with the foundational experiences of Level A would be ready to formally engage with the key terms, *Mean Absolute Deviation (MAD)*, and *interquartile range* at Level B. Finally, at Level C, the 2005 GAISE Curriculum Framework recommends for the learner to experience formal treatment of the key terms, *standard deviation*, *variance*, and *outliers*. Table 2.4 lists these key terms and the levels where the 2005 GAISE Curriculum Framework recommends that the learner begin to formally experience them (Franklin, et al. 2007).

<b>Term</b>	<b>Level A</b>	<b>Level B</b>	<b>Level C</b>
Distribution	X		
Variability	X		
Range	X		
Mean Absolute Deviation		X	
Interquartile range		X	
Standard deviation			X
Variance			X
Outlier			X

*Table 2.4 Levels Where Formal Treatment of Focal Terms Are Recommended by the 2005 GAISE Curriculum Framework.*

*Source:* Franklin, C., Kadar, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria, VA: American Statistical Association.

This table does not indicate that these terms would not be encountered sooner than the indicated level. It indicates where they would be first treated formally. For example, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) recommends formal outlier

analysis at Level C and more informal experiences with outliers at Level A and B. At Level A, it states that, “An understanding of error versus natural variability will help students interpret whether an outlier is a legitimate data value that is unusual or whether the outlier is due to a recording error” (Franklin et al., 2007, p. 33). The report also states that Level B learners might encounter outliers sooner than Level C when using statistical software or graphing calculators (Franklin et al., 2007, p. 48). At the same time, another key term, *variance*, is not mentioned in the report, but is assumed to be at Level C along with standard deviation.

Finally, according to the 2005 GAISE Curriculum Framework (Franklin et al., 2007) the key term *variability* would be formally introduced at Level A. The authors of the 2005 GAISE Curriculum Framework state that, “At Level A, it is imperative that students begin to understand the concept of variability. As students move from Level A to Level B to Level C, it is important to always keep at the forefront that *understanding variability is the essence of developing data sense*” (Franklin et al., 2007, p. 3, emphasis added in original). Along with this, the 2005 GAISE Curriculum Framework recommends that different aspects of the nature and focus of variability be formally treated at different levels. For example, regarding the nature of variability, at Level A the 2005 GAISE Curriculum Framework recommends that the learner has experiences with measurement variability, natural variability, and induced variability. At Level B, it recommends that the learner has experiences with sampling variability, and at Level C that the learner has experiences with chance variability. Regarding the focus of variability, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) recommends that the learner focus on

*variability within a group* at Level A, *variability within a group* and *variability between groups* at Level B, and *variability in model fitting* at Level C.

In summary, this dissertation study seeks to discuss teacher content knowledge of variability in data distributions. The purpose of this section is to help in the interpretation of the findings. The steps in doing so included:

- describing the part of statistical investigation that applies to this research
- defining some key terms used in the research
- outlining the three GAISE Levels (Franklin et al., 2007) for part three of statistical investigation: Analyzing the data
- placing key terms in the suggested levels of the 2005 GAISE Curriculum Framework (Franklin et al., 2007)

The findings of this dissertation study will be discussed in the landscape of these resources along with some additional statistics education research. For example, this research includes that of Makar and Confrey (2004) who studied teachers' informal and formal talk about variation and distribution. It is hoped that the statistics education research, the 2005 GAISE Curriculum Framework (Franklin et al., 2007), and expert definitions will guide in characterizing the formal and informal knowledge that is exhibited in this dissertation study. It is also anticipated that these three sources will guide in indicating where the knowledge lies in the continuum of formal and informal statistics knowledge. Specifically with regards to Mae (pseudonym for the teacher studied in this research), it is not just expert knowledge but also her informal knowledge that are studied. Based upon the statistical experiences Mae had in this dissertation study, it is anticipated that she will be at either Level A or Level B of the 2005 GAISE Curriculum Framework.

## **Chapter 3**

### **Research Design**

As previously stated in Chapter 2, the sensemaking model—noticing, interpreting, and implementing—could be helpful to study teacher content knowledge of variability (Weick, 1995; Drake, 2006). This model seemed to give a solution to the problem of studying teacher knowledge that I needed. It gave a more bidirectional way of seeing what teachers know. This is in contrast to previous studies that looked in one direction. These studies looked to that what a teacher knows prior to teaching to indicate what he or she would know in teaching. My view of studying teacher knowledge would connect what the teacher exhibits knowing beforehand to his or her teaching. Conversely, it would connect his or her teaching to what he or she exhibited knowing prior to teaching. My intent is to create a more complete perspective on teacher content knowledge—particularly of variability in data distributions.

In this vein, the sensemaking practices of noticing, interpreting, and implementing seemed helpful to operationalize the construct of teacher knowledge (Weick, 1995; Drake, 2006). The focus on what one notices, interprets, and implements gives insight into the sense one is making of it. In this way, sensemaking could be considered a proxy for one's knowledge. In using the sensemaking model in the context of studying teacher content knowledge bidirectionally, I focused on what the teacher noticed and interpreted prior to teaching in her—lesson planning, professional development, and performance tasks—and on what she implemented in teaching.

In summary, this bidirectional view of what the teacher exhibited knowing—prior to and during teaching—is intended to give a more complete view of her content knowledge. In this way a teacher's knowledge is not exclusive to viewing what she knows in only one

direction, either prior to or during teaching. Instead a bidirectional view of a teacher's knowledge can see the possible connections between each view of teaching. Similar to cartography, connecting both views of sensemaking—prior to making the map with the sensemaking when making the map—can give a more complete snapshot of the cartographer's knowledge of the terrain. What the cartographers notice and interpret prior to making the map can connect to what they know when making the map. Likewise, what they know when making the map can connect to what they noticed and interpreted about the terrain prior to making the map.

For this dissertation study, the first view is of the teacher's sensemaking (noticing and interpreting) prior to teaching in her professional development, lesson planning, and performance tasks, and the other view is of the teacher's sensemaking during her lesson implementation. Based on its usefulness in accomplishing this, Weick's (1995) and Drake's (2006) sensemaking model of—noticing, interpreting, and implementing—was used to frame my research questions.

### **Questions**

What does a middle school teacher know about the statistical concept of variability of data distributions that she is charged to teach?

Using Weick's (1995) sensemaking framework (notice, interpret, implement):

- a) What does this teacher seem to notice about variability in data distributions?
- b) What does this teacher seem to interpret about variability in data distributions?
  - i) in written tasks with interview
  - ii) in work done in professional development

- iii) in lesson planning
- c) What knowledge of variability in data distributions does this teacher implement when teaching this content to her students?
  - i) in inscriptions, gestures, and locutions during class
  - ii) in interview/reflection after teaching

### **Setting**

The school and the New York City setting were chosen partly based on sampling convenience because I am employed in a New York City school. However, there were two other important reasons I chose my school to study. First, New York City's diverse school system offers research a wealth of information regarding mathematics education. Its diverse teaching staff and student body give researchers potential for rich descriptions of important mathematics topics. Second, I wished to give back to a school system and teaching staff that provides services for many students. As a mathematics coach I welcome the opportunity to engage teachers in the discussion of an important statistics topic. I look forward to seeing how their understanding of this topic is manifested in the classroom. The benefits of this study can go beyond the school. Conducting this study at my school will not only help its current seventh-grade teachers understand an important statistics concept, it will also serve to inform future professional development. This professional development is an integral part of the teaching and coaching relationship that the City continues in order to support student learning.

This school is the same school as was in my practicum study. I am a mathematics coach at this school and have established relationships with the principal and the teachers. My relationship with the principal enables me to collect data in the school, and my professional relationship with the teachers affords me easy access to them during their working day. My

responsibilities include organizing staff professional development that will facilitate conducting the proposed professional development seminars.

### **The School**

The school is a multicultural magnet school serving approximately 1,500 students from the entire borough of Queens. It has a diverse student population with approximately 34% White, 17% Black, 33% Hispanic, and 16% Asian. Since its inception, the school has been dedicated to good middle school practices. It has sought to create an exemplary school and demonstration site. Under federal mandate, the school selects students to reflect the ethnic and cultural diversity of Queens. A variety of ability levels are also represented. Since the school is not considered a community school, the students are selected through an application system. Parents submit their child's applications. These applications are pooled together and then chosen via lottery. In order to attract nonlocal students, two open houses are conducted each year for prospective students and parents to learn about the school and its programs. This aspect of the school makes it unique among the other schools in the City that usually serve the children of its surrounding community.

There are 96 teachers at the school to serve its diverse student body, along with 38 paraprofessionals. The large number of paraprofessionals accommodates the school's significant population of special needs students. As much as possible, these students are mainstreamed into the regular education classes. There are 6 Cooperative Teaching classes and 10 Special education classes. The school also services a small percent of English Language Learners (ELL's). In relation to the rest of the City's community schools, this school services a greater percentage of special needs students and a smaller percentage of ELL

students. The school is responsible to instruct all students in alignment with the New York State standards.

The school is mandated to cover the New York State curriculum and to administer all State assessments. On the State and City 2004–05 English Language Arts Tests, 5.0% scored far below the standard, 31.6% were below but approaching the standard, and 63.4% met or exceeded the standard. On the 2004–05 State and City Mathematics Tests, 9.5% scored far below the standard, 27.8% were below but approaching the standard and 62.7% met or exceeded the standard. On the grade 8 statewide test, 75.8% met or exceeded the standard.

Instruction is aligned with the new State curriculum in all subject areas, and general education classes are for the majority heterogeneously grouped. Both remediation and enrichment are incorporated into instruction, with consideration given to multiple intelligences and learning styles. There are exceptions to heterogeneous grouping. Among them are: students who take Regents classes in Living Environment and Integrated Algebra; at risk students, including Instructional Support Students and English Language Learners who receive additional support in their elective periods.

The school embraces a Balanced Mathematics program that strives to create lifelong learners and to build interest in mathematics. There is a strong focus on fostering critical thinking and mathematical communication. This aligns with the City’s math program in intent. However, the curriculum used at the school differs somewhat from the rest of the City.

The school implements the following programs:

- Everyday Mathematics with Math Steps in Grade 5
- Connected Mathematics (CMP) with Impact: grades 6 – 8
- Prentice Hall’s Integrated Algebra in Accelerated Grade 8 classes

The rest of the city uses Impact Math for grades six to eight without using CMP. The school is one of the few schools in the city using this curriculum. As a result, the city offers little to no support for CMP users. Thus, I, as the math coach, provide the support and modeling necessary to implement CMP, along with Everyday Mathematics, and Integrated Algebra. Special attention is given to teaching the students with learning disabilities.

### **Professional Development**

The New York City Department of Education schedules two full calendar days of professional development. These are conducted at the end of August and one at the end of June at the teachers' respective schools. In addition, the City offers various professional development workshops and seminars throughout the City for both administrators and teachers. As do many other City schools, the school takes advantage of the various professional development workshops offered. In addition, it offers its teachers membership in a school-based Critical Friends Group. This group meets monthly after school to discuss improving teaching and student learning. Its membership is voluntary and teachers from all grades and disciplines attend. Critical Friends was initially funded by the Annenberg Institute of School Reform at Brown University. Since the City does not fund it, the school's involvement is somewhat unique compared to other City schools. In addition to general professional development opportunities, subject-specific professional development, such as mathematics is offered by both the City and this school.

Professional development for mathematics teachers in New York City has many aspects. There are ongoing monthly professional development meetings for the mathematics coaches. These sessions are usually separated by elementary and middle school levels. In addition, independent professional development workshops are offered to teachers on specific

topics such as teaching algebra, implementing the Impact curriculum, or using manipulatives. Further, there are professional development sessions offered via grants. For the past three years, Title II B grant funds have been used to sponsor professional development for mathematics coaches and lead teachers. All of these professional development opportunities take place outside of the school.

Professional development for mathematics at this school is conducted according to the City's schedule for professional development. One day at the end of August and one day at the end of June are designated for professional development. These professional development sessions are conducted in the school building, and are run by the math coach. There is usually a general meeting of all the fifth to eighth grade teachers and separate meetings for each grade. A central focus of these meetings is planning curriculum that aligns with the New York State 2005 Mathematics Standards. In addition to these two formal days of professional development, weekly meetings of the seventh and eighth grade regular and special education teachers afford opportunities for the math teachers to share materials and teaching strategies. During monthly math department meetings, grades five to eight have the opportunity to do the same. Teachers, who attend a professional development seminar outside of the school, will turn-key the information to the grade appropriate colleagues. In addition, informal professional development takes place during teachers' regularly scheduled professional periods. Finally, on occasion, math teachers and/or the coach attend national and regional conferences, such as the National Council of Teachers of Mathematics' regional conference held in Atlantic City, New Jersey in 2007.

In this dissertation study, the three professional development sessions ran after school for two hours with one follow-up session for teachers to plan their lesson(s). The aim of their lesson was the variability of data in distributions. The following dates were selected.

3:15–5:15 pm	March 20	Investigation 1 Making Sense of Variability Investigation 2 Making Sense of Measures of Center
3:15–5:15 pm	March 27	Investigation 3 Comparing Distributions: Equal Number Values Investigation 4 Comparing Distributions: Unequal Numbers of Data Values
3:15–5:15 pm	April 8	Lesson Planning—Variability in Data Distributions

Finishing the CMP curriculum on April 8 gave the teachers time to plan and implement their lessons on variability in data distributions before the end of the school year. The teachers designed their lessons on variability in data distributions. The length and duration of these lessons depended upon the time available in their curriculum-pacing calendar. There were approximately three to five days available for these lessons.

Based on the teachers' personal schedules, the actual schedule of professional development sessions changed. See the data collection section for further explanation of this change. The following sessions took place on the noted dates:

Session 1	Professional Development	May 21, 2008	after school
Session 2	Professional Development	May 29, 2008	after school
Session 3	Professional Development	June 5, 2008	after school
Session 4	Professional Development	June 16, 2008	after school
Lesson plan	Mae and Connie	June 19, 2008	during school
Lessons	Mae	June 23, 2008	during school
	Connie	June 25, 2008	during school
Tasks	Mae and Connie	June 25, 2008	during school
Tasks	Connie	June 25, 2008	after school

## **The Curriculum**

The professional development seminars for this dissertation study used the Connected Mathematics Project (1995) curriculum (CMP). Investigations 1 through 4 of the *Data Distributions* text were addressed (Lappan et al., 2006). The decision to use a curriculum in the professional development setting was based upon the fact that teachers' pedagogical content knowledge of variability in data is most likely nascent. In this regard, a curriculum that is geared toward middle school students sets up problems that focus on variability and data distributions in a developmentally appropriate way. The teacher can then be somewhat relieved of the need to make pedagogical content decisions. This does not mean that the teachers are not required to make these decisions while using the curriculum. However, part of this work is done when the tasks are predesigned and sequenced for the teacher.

The specific curriculum was chosen because it is a standards-based and National Science Foundation funded program. CMP is a comprehensive problem-centered mathematics curriculum. The U.S. Department of Education's expert panel has designated Connected Mathematics as an Exemplary Mathematics Curriculum (1999).

The overarching goals of CMP are that all students should be able to reason and communicate proficiently in mathematics. This includes knowledge and skill in the use of the vocabulary, forms of representation, materials, aspects, techniques, and intellectual methods of the disciplines of mathematics including the ability to define and solve problems with reason, insight, inventiveness, and technical proficiency. CMP makes a commitment to skill, but skill is much more than proficiency and computation and manipulation of symbols. Skill in CMP means that a student can use the mathematical aspects, resources, procedures, knowledge, and way of thinking that have been developed

over time to make sense of new situations that the student encounters. CMP development was guided by five key instruction principles: Mathematical Investigations (“big ideas” in mathematics [Lappan et al., 2002]), Reasoning (reason effectively using information represented in multiple ways), Teaching for Understanding (emphasizes inquiry and guided discovery), Connections (with other subjects and the real-world), and Technology (calculators and computers) (Lappan et al., 2002).

The curriculum was also chosen because it focuses upon understanding variability of data distributions. As Makar and Confrey (2005) conjectured, the concepts of variability and distribution seem to be coupled when teachers describe distributions using nonstandard language. Therefore, the curriculum’s focus positively influences the likelihood of the participant teachers’ discussing the variability of data. In addition, the curriculum uses precollected data because data collection is not fruitful when there is a small group as was in this research. This is important because the amount of data would be too small to compare to the ones that Canada (2004) used in his dissertation.

The CMP curriculum is familiar to me. I was involved with the CMP curriculum both as a teacher and as a math coach. Since I began student teaching in 1991, I was exposed to the CMP curriculum. The school was a pilot site for the program and my cooperating teacher was selected to test it in her sixth-grade classroom. From the first day I observed her class, I was impressed with the way the curriculum engaged the students and had them thinking about mathematics. I remember reflecting on how I could not identify the stronger students from the others because all of them were participating in the group activities with equal fervor. The following year I began teaching the seventh grade at the school, and I continued to use the

CMP curriculum. I taught to the best of my ability in accordance with their inquiry-based instruction. It became the most comfortable way for me to teach mathematics.

My involvement with the CMP extended outside of the classroom in a number of ways. I was selected to participate in its piloting; I attended training sessions in Michigan, and I spent three years working on its second iteration. The first two years of teaching I piloted the program in my seventh-grade classroom. This included keeping daily logs of what I taught, how it went, etc., and weekly meetings with Dr. Frances Curcio from Queens College to discuss our progress, issues, or concerns. After piloting the program, I attended a week-long training session in Michigan. It was a nationwide conference where teachers and administrators met to discuss in greater detail the curriculum. Our activities included doing the problems and investigations, and analyzing assessments and student work. Finally, from 2001 until 2004, as a graduate assistant I worked on the second iteration of the curriculum. My main focus was to read behind the authors in their writing and to create answer keys. I participated in author and board of advisors meetings. My involvement at CMP influenced its use in my study.

In 2004, I worked closely with the authors of the CMP book that was used in this dissertation study, *Data Distributions* (Lappan et al., 2006). My participation stemmed from becoming interested in statistics. I communicated weekly with the authors and assisted them in writing problems. On one occasion, I visited a school that was piloting the book. I videotaped lessons and interviewed the teachers. My interest in statistics and my knowledge of the curriculum made it an easy match for my research topic. I came into this study with an enthusiasm for statistics and the curriculum that brought it to the middle school level.

## **Participants**

There were four possible teachers to select for my dissertation study. All four of seventh-grade mathematics teachers are desirable because there are differences in their years of teaching. The range is from two to thirty years. However, one teacher was pregnant with twins and would not be teaching at the time of this research. This teacher is the teacher that I studied in my practicum. As a result, I would not be able to directly compare the results of her learning this year to her learning last year.

All teachers were to choose one heterogeneously grouped class that they taught to be involved in the study. During their normal scheduled school day, all participants met once a week to plan their lessons. I was a participant-observer in this study because of my current role in the school as mathematics coach. As a coach I attended the teachers' weekly planning meetings and had access to their classrooms when they were teaching. However, I was regularly assigned to work only with the teacher who had two years of experience. This work included observing a class and conducting a postlesson discussion twice a week.

Ultimately two teachers chose to be involved in the study. These teachers, Connie and Mae, pseudonyms, have long established relationships with me. It was beneficial to have two teachers involved in the research for a couple of reasons. First, they were able to bounce ideas off of each other during our professional development discussions. This included intersubjectivity, or asking the other to verify what they understood or reasoned. Second, they were able to support each other in the learning process. Particularly, going through the study together could help alleviate uncomfortable feelings. This includes feelings that might emerge when being observed while learning something new.

Inasmuch as it was beneficial to have two teachers involved in the study, time did not allow for the write up of both. Therefore, I decided to analyze and write up my findings on one of them. I believed that one thick description could be a basis in which future studies might be inspired. I also knew that the data on the other teacher could be analyzed at another time. The teacher I chose to write up was Mae. Mae was quite verbal and outspoken regarding what she knew or saw as the variability in data distributions. This was a major factor in choosing to analyze her data.

Mae and I have been colleagues since she began teaching in 1999. Since that time she taught the seventh grade on the second floor with no interruption in service. At the time she entered the school, I was the mathematics lab facilitator. During her first year of teaching, I was not assigned officially to her classroom, but I did interact with her on some occasions. Since then, our interactions have been at the seventh-grade meetings, and when I disseminated materials to her classroom. During those times, Mae has asked me questions regarding the curriculum, the mathematics, and/or her assessments. Mae prefers to initiate contact when seeking help from me regarding her teaching.

Mae and I developed a friendly relationship that remains within the confines of the school day. We did not socialize outside of school other than at school functions. Our relationship could be considered professional and friendly. If asked, she is always willing to give of her time to lend a hand or to do a favor either professional or personal.

In her practice Mae is a planner who prefers to plan with colleagues. She verbalized the need to plan during the seventh-grade weekly meeting. During these meetings she gets down to business and is often seen leading the discussions. She is generous with her time, materials, and talents not only at this meeting, but also throughout the school day. She is

known to spend her lunchtime creating an assessment with colleagues, tutoring students in mathematics, or practicing an instrument along with them.

Mae is involved in other social activities of the school. She assists the student organization coordinator with running the dances. It is not unusual to see her classroom filled with students helping prepare party decorations. Over the years, Mae has developed close relationships with students that have extended outside of school.

Mae is a disciplinarian in the classroom. Her teaching is a hybrid of teacher-directed and student-centered methods. She has used the CMP curriculum since she began teaching at the school. She was not involved in the piloting of the program. However, she did go to Michigan for a week's training. She is a proponent of this mathematics program and has not strayed from its use as outlined in the seventh-grade curriculum guide.

Along with me, Mae is a member of the school's Critical Friends Group (CFG). She has been a member of the CFG group since 1999. She is an enthusiastic participant who freely brings up dilemmas in her teaching for discussion. It was during a CFG meeting that she expressed the desire to be the initiator in seeking help from me in her practice. This year she asked fellow members to observe one of her classes and give suggestions for teaching them. Over time Mae has taken on some leadership responsibilities in the group, and this year she has gone for formal leadership training. Mae's involvement with CFG demonstrates her desire to reflect upon her practice on a regular basis. This made her a good candidate for my study, because I anticipated that she would bring the same level of reflection to it.

Mae has an undergraduate and master's degree in Mathematics Education from City University of New York at Queens College, Flushing, New York. She has taken one statistics

and probability course during her undergraduate schooling. However, the coursework only covered statistics.

### **Researcher's Roles**

I came into this research with a relationship with Mae. I was the coach for the school's mathematics department. However, I interacted with her at the grade-level meetings. In this section I discuss my role as a mathematics coach with the seventh-grade teachers and the role I had with them during the research.

The seventh-grade teachers met weekly during a regularly scheduled 45-minute period in the school. Attendance was voluntary, but the administration and I encouraged it. Most of the teachers attended the meetings, and shared ideas and/or concerns about teaching the curriculum, student learning, and the State assessment. In addition, the teachers actively planned their weekly calendar. This planning included, but was not limited to, selecting the CMP investigations, homework, and assessments for the upcoming week(s) of instruction. All of the teachers aligned their teaching with the school-created curriculum guide and strove to remain closely in pace with each other.

My involvement with the seventh-grade weekly meetings changed over time. Initially, I would ask probing and challenging questions about the mathematics that was being taught. Some discussions I initiated included unpacking the mathematical concept, making connections to other topics, and exploring methods to help students understand the topic. However, finding a common planning time was a challenge for the teachers. They wanted planning to become the focal point of these meetings. When the meeting refocused on planning, I was given the floor for approximately 5 to 10 minutes at the end of the meetings. As a result, the topics I brought up were administrative in nature, e.g., the ordering of supplies

and/or making copies. Nonetheless, I participated freely in their planning, and only when absolutely necessary I interjected a deeper discussion of the mathematics. When needed, the teachers also continued to seek guidance and direction from me. They would ask questions about the mathematics or the curriculum.

How I interacted with the teachers in the study contrasted with how I interacted with them at the seventh-grade meetings. Whereas in these meetings the teachers discussed their priorities and I interjected when necessary, during the study I directed the course of the work that was done. There was a greater dependency on me in the study. The teachers relied upon me to tell them what they needed to complete in the curriculum and when to move on in the discussion. Because they did not have familiarity with the statistics topic of variability, they looked to me to answer their questions on the content.

In the seventh-grade meetings, I would freely give all of the information the teachers requested in mathematics or otherwise. This was a part of my role. However, during the study, I was more of a judicious teller who selected what information was disclosed to them. Specifically, I initially withheld informing them about the two types of variability. Their productive struggling with the concept revealed their thinking about it. This provided valuable information for the study, and was helpful in answering my research question on what they were noticing about the topic and how they were interpreting it. However, I was uncomfortable with this role because it was unlike what I would have done and what they were used to me doing.

As a researcher, I was uncomfortable about not giving information on the two types of variability. I did not want their level of frustration to interfere with their comfort in participating in the study. One of the teachers was particularly vocal about wanting to find out

from me this information on variability. She expressed her frustration with not knowing and not being told. In an effort to mitigate her frustration and to not lose her participation in the study, I introduced the term *variability* in the glossary of the student edition and in the mathematical concepts section of the teacher edition. As a math coach, at any time I would not have withheld any information from my teachers about a mathematics topic.

Also as a researcher, I could not allow the teachers to plan a lesson on a topic in which they have serious misconceptions. For me this would interfere with my ethics on a teacher's responsibility toward her students to teach correct information. This is not to say that teachers can't make mistakes while teaching. I am addressing my knowing that a teacher has a misunderstanding of a topic and still allowing her to teach a lesson with this misconception. At the end of the professional development sessions, I attended to this problem by giving both teachers the teacher edition to plan their lessons on variability. This addressed their need to clarify their understanding of variability, feel comfortable teaching it, and at the same time provide the study with information on what they were noticing and interpreting about it. Again, as a coach, I would not knowingly permit a teacher to teach a lesson with a major misconception about the topic they were about to teach. It is my job to prepare them to teach correct mathematical concepts.

In both roles as math coach and researcher I was an active listener. What the teachers were saying or asking was important in both relationships. However, I responded to them in different ways. When I acted solely as their math coach I would give them whatever information they required interjecting that information at whatever opportunity I could. As a researcher, I listened more to their thinking and chose not to interject often, as I did not want to perturb their train of thinking. Often, my responses included but were not limited to: "Um

hum,” “Okay,” “Ah ha,” “All right.” These were indicators of my intention of not interfering with their thinking. As a researcher I wanted to find out as much as I could about their thinking, and I asked probing questions, such as “How so?” “Why?” “Because?” “When you say...what do you mean?” Finally, I sought to understand what they were saying, for example: “I have no idea what you said,” or “So you are saying...?” In spite of the fact that I was an active listener in both roles, there was greater emphasis on listening without interjecting in the researcher role.

In this research, I sought to mitigate any power relationship that might ensue based on my role as mathematics coach at the school. In my role as mathematics coach I do not have the authority to be evaluative. My role is one of support and guidance. There are no consequences in place for not following my suggestions. However, I do run the mathematics department meetings and I do hold sway in decisions that are made for the teachers. As a result, this can affect teachers’ perceptions of me. During my study, I sought to mitigate some of the assumed power differentiation as discussed below.

Meeting dates, times, and locations were flexible to accommodate the teachers’ needs. The teachers’ schedules, availability, and comfort with the meeting time were adjusted at every opportunity. Original meeting dates were changed to accommodate their availability. Meetings started when they were ready to fully participate. If they had business with which to attend, we waited until they were finished to begin. Initially meetings were held in one teacher’s classroom, but when the school was closed, they were moved to the other teacher’s home. The atmosphere of the professional development sessions was friendly and comfortable.

Part of making the professional development experience comfortable was controlling the amount of work completed. It became clear after the first professional development session that the teachers worked hard, and that not all the desired curriculum was doable in the remaining sessions. In an effort not to overtax the participants, I made the following two adjustments to the professional development curriculum. First, I scaled back on the investigations selecting only those that were essential to the topic I wished to study. Sometimes this scaling back included eliminating problems, and sometimes it involved my reading through and explaining the problems instead of having the participants do them. Second, in spite of the fact that calculators were available in these sessions, I also sought to give them the calculated answers when it did not interfere with my understanding their thinking. These two changes streamlined the work that was covered and helped maintain a less pressured work environment during these sessions.

### **Data Collection**

Data collection began during professional development, which was audiotaped to capture teachers' exact verbalizations of variability. In addition, teachers' written work was collected for analysis of their written inscriptions of variability. This process was extended with postprofessional development tasks and interviews where semistructured interviews were used.

Canada (2004) found that one of his participants used heavy gestures in discussing graphs to convey ideas that were reasonable. As a result I decided to videotape the professional development interview tasks to capture all possible descriptions of variation expressed by the participants. The interview tasks were chosen based on their use in previous studies. (See Appendix A.) This enabled me to reference my findings with that of the original

researchers. Data collection continued in the classrooms where the two participants taught the lessons they developed in professional development on variability in data distributions. As with the professional development interview tasks, these lessons were videotaped to capture verbalizations, gestures, and black board inscriptions used by the teacher during her lesson.

Table 3.1 shows the type of data collection, responsible person, and method of collection.

<b>Type of data collection</b>	<b>Responsible person</b>	<b>Method</b>
Professional development work	Researcher	Audiotaped Teacher written work
Postprofessional development task & interview	Researcher	Videotaped Field notes
Classroom observations	Researcher	Videotaped Field notes
Postlesson interviews	Researcher	Videotaped

*Table 3.1 Type of Data Collection, Responsible Person, and Method of Collection*

Table 3.2 shows the research questions in tabular format with data collection methods.

<b>Questions</b>	<b>Type of data collection and responsible person</b>	<b>Method of collection</b>
Overarching question:  What does a middle school teacher know about the statistical concept of variability in data distributions that she is charged to teach?		
a) What does this teacher seem to notice about variability in data distributions?	For a) and b):  -written task with interview	For a) and b):  -videotape interview
b) What does this teacher seem to interpret about variability in data distributions?	-written and verbal work done in professional development	-paper and pencil task  -teacher inscriptions

*Table 3.2 Questions in Tabular Format With Data Collection Methods*

Table 3.2 (cont'd)

	-interviews after professional development  -via observations in professional development  Researcher	-audiotape of professional development  -videotape; field notes  -audiotape for professional development
c) What knowledge of variability in data distributions does this teacher implement when teaching this content to her students?	-observations in the classroom  -interviews after teaching  Researcher	-videotape; field notes  -videotape; field notes

*Table 3.2 Questions in Tabular Format With Data Collection Methods*

Data collection began at the first professional development session on May 21, 2008. The following is a description of the data collection process for the study. As stated previously, my role as the mathematics coach enabled me to have access to the school, the teachers, the curriculum, and to the time needed to spend with them in the school day. On a regular basis, I am scheduled to spend time with teachers discussing their practice. I have never been assigned to the two teachers who participated in the study. However, I was given permission to do so. This allowed me to discuss and videotape their lessons, and conduct postlesson interviews and tasks. For Mae, these all took place during the school day. For Connie, the lesson planning and the lessons were conducted during the school day, and the postlesson interviews and tasks took place after school. For Connie, meeting after school was based on her teaching the lesson in the last period of the second to last day of school. Since

there was no time to meet on the final day of school, we met after school the day before, commensurate with her completing the lesson.

Collecting data was more challenging than expected for a number of reasons. It was affected by the availability of the participants and the time the study took place in the school year. In contrast to professional development sessions that are scheduled with little or no flexibility, these planned sessions changed to accommodate the participants. The intent was to make these sessions as comfortable to the participants as possible and to adhere to the integrity of the study. In line with this, no session was conducted when there were time constraints, pressure on the participants, or when one of them was unavailable. Thus, when the participant's personal commitments conflicted with planned professional development, the dates were changed accordingly.

Conducting the study in the latter part of the school year affected data collection. The ideal outline for data collection included four, two-hour sessions of professional development that would take place over four weeks. However, based on the busy time in the school year that the study began, the planned data collection was reduced to two, two-hour sessions to do the investigations, and one, one-hour session for the lesson planning. This condensing of the sessions pushed more of the curriculum into each session. Initially one investigation was planned for each session, and then it was changed to two investigations. See the "Professional Development" section for the initially proposed schedule.

As evidenced in the first professional development session, compacting more than one investigation in each session was problematic, and adjustments were made again. The professional development schedule changed to reflect the time it took for the teachers to

complete the assigned CMP investigations. The new professional development schedule became:

Session 1	Professional Development	May 21, 2008
Session 2	Professional Development	May 29, 2008
Session 3	Professional Development	June 5, 2008
Session 4	Professional Development	June 16, 2008
Lesson plan	Mae and Connie	June 19, 2009
Lessons	Mae	June 23, 2008
	Connie	June 25, 2008
Tasks	Mae and Connie	June 25, 2008

In the first professional development session, completing Investigations 1 and 2 was planned, and after working diligently, the teachers finished up to and including Problem 1.3. This left one session to do three investigations. Since they worked hard during the first session, it did not seem reasonable to expect this amount of work from them. It might prove counterproductive to the study and stressful for them. Therefore, the remaining sessions were reconfigured to accomplish the following CMP investigations and problems:

Session 2	Investigation 1: Problem 1.4
	Investigation 2: Problems 2.1 to 2.4, and
	Investigation 3: Problem 3.1
Session 3	Investigation 3: Problems 3.3A, and 3.4.
Session 4	Investigation 4: Problem 4.2 C

Considering what was reasonable to accomplish in a session, the problems were streamlined. In some instances, they were eliminated or explained by the researcher and answered verbally by the participants. In other instances, when it did not interfere with their understanding of variability, calculations were provided. This was accomplished in the following manner:

<u>Session 2</u>	<u>Problem</u>	<u>How Accomplished</u>
Investigation 2	Problem 2.2 and 2.3A	walked through verbally
	Problem 2.3 B	with researcher
	Problem 2.3 C–D	individually
Investigation 3	Problem 3.1	eliminated

### Session 3

Investigation 3	Problem 3.2 A	done verbally
	Problem 3.2 B, C, D E	done in writing
	Parts C 1 and 2	calculations were given
	Problem 3.3 A	individually in writing
	Problem 3.3 B	eliminated
	Problem 3.4 A–C, D	individually in writing

### Session 4

Investigation 4	Problem 4.1	eliminated
	Problem 4.2 C	individually in writing

The problems that were eliminated either had concepts that were repeated in another problem, such as Problem 3.1 that described and compared reactions, which was repeated in Problem 3.2. This was the same for Problem 3.3B which covered the same concept as Problem 3.3A—comparing distributions. Also, Problem 4.1 that compared distributions with unequal numbers of data values using bar graphs is similar to Problem 4.2 that compared the same type of distributions using line plots. Cutting back on the problems that might have repeated the concepts gave more time for the problems that covered them as well or better.

The problems that were done verbally covered basic knowledge that needed to be understood, but the responses of the participants were not necessary to capture their understanding of variability. Two examples of problems that were verbally answered are Problem 2.2, which discussed the mean as a balance point in a distribution, and Problem 2.3A, which discussed the usefulness of the mode to describe data distributions. The problems that were completed individually by the participants in writing were deemed important in determining their own thinking about the aspect of variability addressed in the question. Their thinking as revealed in their locutions and inscriptions was necessary data for the study. For example, Problem 4.2C required comparing distributions of unequal numbers of data values, which is an essential situation to notice and discuss the variability of data distributions.

The selected problems aligned with Canada's (2004) Evolving Framework of his elementary preservice teachers' thinking about variation. These problems fell into the second aspect of his framework, Displaying Variation. Most tasks focused upon his dimension of Evaluating and Comparing Graphs. Within this dimension, the problems gave attention to different themes that he focused on in his EPSTs' thinking that included average, range, extremes/outliers, spread and/or shape. Table 3.3 depicts the problems covered, goals accomplished, how goals were completed, and alignment with Canada's (2004) Evolving Framework for the CMP investigations.

<b>Investigation number</b>	<b>Problem title</b>	<b>Goal</b>	<b>How completed</b>	<b>Alignment with Canada's (2004) Framework</b>
Investigation 2 Making Sense of Measures of Center	Problem 2.2 The Mean as a Balance Point in a Distribution	Understand the balance model to make sense of the mean	Walked through	Focus on average
	Problem 2.3 A Repeated Values in a Distribution	Use graphical display of data to understand and decide when to use the mode to describe a distribution	Verbally	Focus on average
	Problem 2.3 B	Explore graphical display of data taking into account repeated values, and making decisions about	With researcher as facilitator	Evaluating and Comparing Graphs

*Table 3.3 Professional Development: Problems Covered, Goals Accomplished, Methods Used, and Connection to Canada's (2004) Evolving Framework*

Table 3.3 (cont'd)

		using measures of center or clusters when answering questions about the data		
	Problem 2.3 C–D	Same as 2.3 B	Individually	Evaluating and Comparing Graphs – Focus on average, range, shape, spread
	Problem 2.4 A 1, 2a Measures of Center and Shapes of Distributions	Understand when and how changes in data values in a distribution affect the median or mean	Verbally in pairs	Focus on average
	Problem 2.4 B	Same as Problem 2.4A	Individually Verbally	Focus on average
	Problem 2.4 C1	Relate the shape of the distribution to the location of its mean and median	Verbally in pairs	Focus on average, shape
Investigation 3 Comparing Distributions: Equal Number of Data Values	Problem 3.1 Measuring and Describing Reaction Times	Use properties of distribution to describe the variability in a given data set	Eliminated – somewhat redundant to Problem 3.2 B	Evaluating and Comparing Graphs
	Problem 3.2 A Comparing Reaction Times	Recognize the importance of having the same scales on graphs that are used to compare data	Verbally	Producing Graphs – Technical details

Table 3.3 (cont'd)

Investigation 3 Comparing Distributions: Equal Number of Data Values	Problem 3.2 B 1,2	Develop and use strategies for comparing equal-size data sets	Verbally	Evaluating and Comparing Graphs – Focus on spread, range
	Problem 3.2 B 3,4	Same as Problem 3.2 B 1,2	Individually in writing	Evaluating and Comparing Graphs – Focus on spread, range
	Problem 3.2 C 1, 2	Same as Problem 3.2B 1,2, and decide if a difference among data values and/or summary measures matters	Answers given by facilitator	Evaluating and Comparing Graphs – Focus on average
	Problem 3.2 C 3, 4	Decide if a difference among data values and/or summary measures matters, and decide when to use the mean or median to describe a data distribution	Individually In writing	Evaluating and Comparing Graphs – Focus on average, range, extreme/outlier, and spread
	Problem 3.2 D–E	Develop and use strategies for comparing equal-size data sets	Individually In writing	Evaluating and Comparing Graphs

Table 3.3 (cont'd)

	Problem 3.3 A Comparing More than a Few Students	Develop and use strategies for comparing equal-size data sets	Individually in writing	Evaluating and Comparing Graphs Focus on average, range, extreme values
	Problem 3.3 B	Develop and use strategies for comparing equal-size data sets	Eliminated – redundant to Problem 3.3 A	Evaluating and Comparing Graphs
	Problem 3.4 A Comparing Fastest and Slowest Trials	Recognize the importance of having the same scales on graphs that are used to compare data	Verbally as a group	Producing Graphs – Technical details
	Problem 3.4 B–D	Develop and use strategies for comparing equal-size data sets	Individually in writing	Evaluating and Comparing Graphs – Focus on average, range, extremes/ outliers, shape, and spread
Investigation 4 Comparing Distributions: Unequal Numbers of Data Values	Problem 4.1 Representing Survey Data	Develop and use strategies for comparing unequal-size data sets	Eliminated – based on time constraints and somewhat repetitive to Problem 4.2	Producing Graphs  Evaluating and Comparing Graphs
	Problem 4.2 A, B		Eliminated- covered in discussion in Problem 4.2C	Evaluating and Comparing Graphs

Table 3.3 (cont'd)

	Problem 4.2 C	Develop and use strategies for comparing unequal-size data sets to solve problems	Individually in writing	Evaluating and Comparing Graphs – Focus on average, range, shape, spread
--	---------------	-----------------------------------------------------------------------------------	-------------------------	--------------------------------------------------------------------------

*Table 3.3 Professional Development: Problems Covered, Goals Accomplished, Methods Used, and Connection to Canada's (2004) Evolving Framework*

Because the problems in the CMP text, *Data Distributions* (Lappan et al., 2006) aligned with Canada's (2004) Evolving Framework of ESPTs' thinking about variation, it was a useful guide in my data analysis. Its use in my data analysis is discussed further in the following chapter.

## **Chapter 4**

### **Data Analysis**

#### **Theoretical Framework for Data Analysis**

In choosing to understand content knowledge for teaching in my dissertation research, I needed to find a way to study it. Since I was seeking to understand the content knowledge of variability in data distributions, I saw it as a landscape of knowledge within which a teacher needs to make sense. I began to see that teachers' sensemaking about variability could be a proxy for the knowledge they had about it.

As a result, I searched for a sensemaking model to use in my dissertation study. I found that Drake (2006) adapted a model of sensemaking from organizational research (Weick, 1995) to study teachers' sensemaking of educational reform. I discuss here how Weick's (1995) work on sensemaking and Drake's (2006) use of it influenced my dissertation study.

Weick (1995) likened sensemaking to cartography where the terrain is unknown and there is an indefinite number of ways to map it out (p. 9). Similarly, I see the knowledge of variability as a landscape of which teachers need to make sense, and that there are many ways to make sense of it. The map the cartographer makes depends on what he or she looks at in the terrain and how he or she looks at it. Similarly, I see that the knowledge a teacher has of variability can be seen through what she notices and how she interprets it. Finally, I see the created map as one possible representation of the sense that the cartographer has made of the terrain. In the same vein, I see the classroom lesson as one possible representation of what the teacher knows about variability. This is how I related Weick's (1995) nature of sensemaking to my work. I also looked at Drake's (2006)

work to see how she interpreted Weick's (1995) sensemaking to study teacher implementation of education reform. Drake's (2006) work and how it relates to mine is discussed in the following.

Drake (2006) sought to understand how teachers' narratives could help explain the sense they made of educational reform. In doing so, Drake used a sensemaking model from Weick's (1995) organizational research. Drake (2006) stated that, "This [Weick's 1995] model suggests that the sensemaking process is comprised of three key actions on the part of teachers responding to reform noticing, interpreting, and implementing" (p. 593). Drake (2006) used what teachers noticed in their turning point stories and how they interpreted what they noticed as a lens to discuss how they implemented the reform curriculum in their classroom. Drake (2006) found that teachers who had turning point stories noticed certain aspects of the curriculum, interpreted these aspects in certain ways, and implemented the reform curriculum in their classroom in a certain way. Turning point stories involved teachers that initially experienced failures in mathematics, but because of a turning-point experience, now see themselves more positively as learners and teachers of mathematics (p. 579).

Drake (2006) reported two types of noticing and interpreting practices of teachers who had turning point experiences. There were teachers who noticed and interpreted the curriculum based on the tools or manipulatives that was similar to their turning point experience with mathematics. These teachers implemented the curriculum generally in a traditional way (p. 597). The other teachers noticed and interpreted the curriculum based upon content that also aligned with their turning point experience. Their levels of practice in the classroom were inclined to be reform-oriented (p. 598). In summary, Drake (2006)

found that the teachers' turning point experiences tended to clearly frame their practices of noticing, interpreting, and implementing reform and the reform curriculum (p. 598).

In contrast to using teacher stories as a way to understand teachers' sensemaking practices in implementing the mathematics education reform, I studied a teacher's sensemaking practices of a statistics topic: variability in data distributions. Similar to Drake (2006) I studied the sensemaking practices of noticing, interpreting, and implementing that the teacher used in coming to know about variability. Specifically, I looked toward these three practices to cull from them the teacher's content knowledge. I assumed that what she noticed about variability had meaning to her, and how she interpreted what she noticed gave insight into her knowledge of it. Finally, how she implemented a lesson on variability shed further light on the knowledge she had of it. My intention was that altogether these three practices would give me a more complete snapshot of the teacher's content knowledge of variability. In this vein, I looked to what the teacher noticed about variability in her lesson planning, professional development problems, performance tasks, and her postlesson discussion. Further, I analyzed how the teacher interpreted what she noticed about variability in these venues. Lastly, I studied one classroom lesson to understand the knowledge she used when implementing a lesson on variability. Again my intent was through studying the teacher's sensemaking practices in these various sites of teaching, I would come to see the content knowledge she has of variability in data distributions.

In summary, I used the process of sensemaking (Weick, 1995; Drake, 2006) to reveal the teachers' knowledge of variability. Seeing the knowledge of variability as a terrain of knowledge the teacher needed to make sense of enabled me to do this. Specifically, I looked into the sites of lesson planning, professional development,

performance tasks, and postlesson discussion to see what was noticed and how it was interpreted. Further, I looked to one lesson of her teaching on variability to study how the teacher implemented her knowledge in the classroom. From her three sensemaking practices I culled the teacher's content knowledge based upon what she noticed about it how she interpreted it and, finally, how she implemented it in a classroom lesson. These sensemaking processes gave me a way to operationalize the large concept of teacher knowledge in order to study it.

In this way, these sensemaking practices (Weick, 1995; Drake, 2006) also aligned with my bidirectional view of teaching. This bidirectional view involved focusing on teacher sensemaking prior to teaching and during teaching. Together these views are intended to give a more complete snapshot of what the teacher knows. Similar to cartographers who are making sense of an unknown terrain, connecting both views of their sensemaking—prior to making the map and when making the map—can give a more complete snapshot of the cartographer's knowledge of the terrain. What the cartographers notice and interpret prior to making the map can connect to what they know about the terrain when making the map. Likewise, what they seem to know when making the map can connect to what they seem to know prior to making the map.

For this dissertation study, the first view is of the teacher's sensemaking (noticing and interpreting) prior to teaching in her professional development, lesson planning, and performance tasks, and the other view is of the teacher's sensemaking during her lesson implementation.

Weick's (1995) and Drake's (2006) work gave me a broad analytic lens with which to direct my data analysis. However, it did not fully explain the finer-grained process needed to

analyze my data. To this end, I looked to research in the field of mathematics education. In my analysis, I came to rely upon the work of Miriam Gamoran Sherin. Sherin (2007) studied professional vision, which consists of two main sub-processes (a) selective attention and (b) knowledge-based reasoning. Selective attention concerns how the teacher decides where to pay his or her attention. Knowledge-based reasoning refers to the ways in which a teacher reasons about what is noticed based upon his or her knowledge. For example, a teacher might reason about a particular event based on her knowledge of the subject matter, knowledge of the curriculum, or knowledge of students' prior commitments. They interact in a dynamic manner. That is, the kinds of interactions that a teacher notices will likely influence how the teacher reasons about those events. In addition, a teacher's knowledge and expectations can be expected to drive what stands out to the teacher in any given situation. Sherin's (2007) meaning involved that which the individual was choosing to focus upon. This focus included among others the various aspects of the context, concept, and/or situation that might be noticed.

Sherin's (2007) research is significant to my study for multiple reasons. First, it falls in the domain of my work in mathematics education research. Second, it creates a dynamic picture of how knowledge-based reasoning influences what one notices and how what one notices influences the kinds of reasoning he or she does. In addition to Sherin's (2007) work that illuminated the dynamic process between these, Sherin's and van Es's (2009) research introduced me to a way to analyze the data. Specifically, they used a plan to identify what a teacher notices and his or her knowledge-based reasoning of it. For example, in their analysis of the video clubs, they initially divided the transcripts into "idea units" (which they took from Jacobs & Morita, 2002)—segments in which a particular idea was discussed. This method of

analysis influenced how I studied my data. This is discussed further in the “Data Analysis” section that follows.

### **Data Analysis**

The sensemaking model—what teachers notice, interpret, and implement—was used as a broad lens of directing my analysis (Weick, 1995; Drake, 2006). Data collected from all contexts—professional development, lesson planning, interview tasks, postlesson discussion, and classroom lesson—were analyzed separately and then cross-sectionally with the intent to develop rich descriptions of what Mae knows about the variability of data distributions. Finding what Mae notices via her utterances, gestures, or inscriptions about the variability of data distributions in these contexts is the first focus of the analysis. Before I formally checked for notices, I did an informal analysis while I transcribed.

Data analysis began at the time of transcribing. It was an iterative process. I did not transcribe for more than an hour or two when I began to connect the teacher comments to the various studies on variability discussed in my “Literature Review” section. Canada’s (2004) *Evolving Framework on EPSTs’ thinking about variation* was often referred to during my transcribing and comment writing. Makar’s and Confrey’s (2005) work on “variation talk” was also referenced. On occasion, I turned to Hammerman’s and Rubin’s (2004) work as a guide regarding the tension between minimizing data without compromising its fullness. During and after transcribing, I made comments on Mae’s utterances. The connections I made to these researchers’ work were included in my comments, questions, and hunches that I wrote on the side of my transcriptions of Mae’s work.

Initially, I also conducted a cross-sectional analysis among the various sources of data for Mae. It began with the transcribing of each of the four professional development sessions

and the lesson planning session. On my first few readings of these sessions I placed comments, questions, and hunches in margins of these documents. Next, her comments were extracted from the transcribed professional development sessions and placed in separate documents. An analysis of these extracted comments included writing comments, questions, and hunches in its margins. All the while, I searched for possible assertions and patterns. Next, the same process was conducted for the performance tasks. A separate typed document was prepared, comments, etc. were written in the margin. At the time the performance tasks were ready to be analyzed (i.e., after the professional development data had several rounds of analysis), I began to confirm or refute the current possible assertions and patterns and to search for new ones.

Overall, my analysis encompassed analyzing Mae's notices, her interpretations, and her lesson implementation. First it began with some initial thoughts when transcribing. Then in a more systematic fashion I found what she noticed. Next I used Mae's notices to guide what she seemed to interpret about her notices. Finally, I looked toward her lesson to find out how she implemented what she noticed. In particular, I compared how she interpreted variability in several venues: her professional development, her lesson planning, her performance tasks, and her lesson implementation with postlesson discussion. Here I explain how I determined Mae's notices, how I counted them, and the steps and criteria I used to understand how she seemed to interpret her notices. Finally, I describe how I analyzed the lesson she taught to her class on variability.

### **What Is a Notice**

Noticing variability for this research encompassed Mae's comments or inscriptions. These were either made as a result of direct questions about variability or made freely by Mae.

It was expected that what she noticed would be repeated throughout her writing and speaking. The repetitions occurred within and among all the contexts—interview tasks, lesson planning, professional development, and in the postlesson discussion. See next section on counting notices for how the repeated notices were handled.

Research was used as a guideline in determining a notice. Sherin's and van Es's (2009) research helped with the finer-grained work of selecting notices. Their use of idea units (Jacobs & Morita, 2002) helped me to consider one idea or concept of variability as a notice. The statistics education field helped to select these concepts. Specifically, Canada's (2004) Evolving Framework that characterized EPSTs' thinking on the variation, as described in the literature review, was used. Within his concepts, that included dimensions and themes, I selected Mae's notices. Since the focus of my research is variability in data distributions and not variation in repeated samplings and probability outcomes, as was explored in his study, only parts of his framework seemed to apply. For example, his dimension of Evaluating and Comparing Graphs that included the themes of average, range, extremes (which Mae called outliers), spread, and/or shape—appeared to be more visible in Mae's work. However, for the purposes of being thorough and accurate in selecting Mae's notices, all concepts in his framework were referred to when determining a notice.

Other research helped in determining what Mae noticed. The “variation talk” in Makar's and Confrey's (2005) research helped direct my attention to the informal language Mae drew upon to describe variability. This research helped me discern Mae's nonstandard language about the variability. For variation, the nonstandard language might include *spread*, *clustered*, *clumped*, *grouped*, *bunched*, *gathered*, *spread out*, *evenly distributed*, *scattered*, or *dispersed*. For distribution, it might include low-middle-high clumps called *triads*, or *modal*

*clumps* (Konold et al., 2002). The modal clumps were the middle portion of the distribution. The nonstandard language for distribution also might include distribution chunks. One example of the distribution chunk is discussing the extreme part of the data distribution, for instance, the handful of students who improved the most. Once I highlighted the word or phrase as a notice, it was aligned with Canada's (2004) Evolving Framework and counted.

Questions were also included with the words or phrases that emerged from the research. They too guided what counted as a notice. As previously stated, the words and phrases were extracted from the research of Makar and Confrey (2005) and Canada (2004). They were the words and phrases that their teachers and preservice teachers used when discussing variation. The questions along with the words and phrases on variability emerged through an iterative process of going back and forth from Mae's utterances to the findings of research. The questions came from both asking what is the question other researchers' teachers are answering as well as what are the questions that Mae is answering about variability. The questions, words, and phrases were used to identify the notices in Mae's statements. The following Table 4.1 depicts the questions, words, and phrases that were referenced when selecting Mae's notices. These questions, words, and phrases were used later to find out how Mae interpreted what she noticed. This is discussed in the section on what Mae interpreted.

<b>Concept</b>	<b>Questions to select notices for interpretation</b>	<b>Words or phrases (some that might connect to variability)</b>
Graphs	<ul style="list-style-type: none"> <li>▪ How does she view the effects that the type of graph or its scaling has upon seeing the variability of data?</li> <li>▪ How do the graphs she constructs or have students construct affect the ability to see the data's variability?</li> </ul>	Scaling Type of graph

*Table 4.1 Questions, Words, and Phrases That Emerged When Selecting Mae's Notices*

Table 4.1 (cont'd)

Average	<ul style="list-style-type: none"> <li>▪ How do data spread out or cluster around the measure of center?</li> <li>▪ What does her use of the mean tell you about their ability to see variability?</li> <li>▪ How does she see the relationship between the location of mean and the median in describing the data's variability?</li> <li>▪ How does she express tolerance for variability when the average is not an exact number?</li> <li>▪ How does she interpret typical as a range of values (thereby displaying attention to the variability in the data)?</li> <li>▪ How does she compare graphs using the average?</li> </ul>	Mean Median Mode Typical Expected Value Expected Value to be somewhere around _____
Range and Extremes (which were outliers to Mae)  Metric meaning	<ul style="list-style-type: none"> <li>▪ How does she see ranges?</li> <li>▪ How does she see range as important when describing, analyzing or comparing the data distribution?</li> <li>▪ How does she create reasonable ranges for a set of data?</li> <li>▪ How does she predict ranges of values?</li> <li>▪ How does she discuss the influence of outliers on the mean, and/or variability of data in general?</li> <li>▪ How is her discussion about minimum and maximum values pertinent to the variability of data?</li> <li>▪ What is her notion of outliers? How does it influence her analyzing or comparing data distributions?</li> <li>▪ How does she discuss obvious outliers or extreme values?</li> <li>▪ How are outliers viewed as individual points or as a contiguous chunk?</li> </ul>	Range Minimum/maximum values Big difference in the values of the variables  Outliers Extremes Extreme values Values separated from a cluster Lower/upper values Values on the ends

Table 4.1 (cont'd)

<p>Spread</p> <p>Distribution as spatial object</p>	<ul style="list-style-type: none"> <li>• How does she discuss the data's clustering, clumping, grouping, or concentration?</li> <li>▪ How does she describe these in relation to the center or other reference point?</li> <li>▪ How does she describe the relative grouping of the data?</li> <li>▪ When and how does she use phrases such as more evenly distributed, more predictable pattern, most of the data falls here?</li> <li>▪ How did she quantify where she noticed the data clustered?</li> <li>▪ How does she use percentage (proportional reasoning) to describe the spread of the data?</li> <li>▪ How does she talk about the concentration or the bulk of the data?</li> <li>▪ How does she discuss the influence of outliers on the mean or on the variability of data in general?</li> <li>▪ How is the range influential in her thinking about spread?</li> <li>▪ How were differences in spread noticed when comparing or analyzing graphs?</li> <li>▪ How naïve/sophisticated/deterministic were her notions of spread?</li> <li>▪ How does she discuss gaps in the graph? Do they influence the size of the cluster? If so, explain.</li> <li>▪ How does she see the relationship between the location of mean and the median in describing the spread of the data?</li> </ul>	<p>Spread</p> <p>Spread out</p> <p>Scattered</p> <p>Dispersed off</p> <p>Cluster</p> <p>Clumps</p> <p>Grouped</p> <p>Concentration</p> <p>Bulk of data</p> <p>Bunched up</p> <p>Close or spread out from the mean</p> <p>Cluster to a center</p> <p>Concentrated at various intervals within the range</p> <p>More evenly distributed over the _____</p> <p>Steady predictable pattern</p> <p>More or most of the variable here and less there</p> <p>Middle 50% range</p> <p>Upper 75%</p> <p>Where data most likely or least likely will be</p> <p>How close or tight the graph is</p> <p>Main group</p> <p>Majority</p> <p>Modal clump</p>
-----------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Table 4.1 (cont'd)

Shape	<ul style="list-style-type: none"> <li>▪ How does she use shape (visual features) to depict the data's variability?</li> <li>▪ How does she describe the shape of the data?</li> <li>▪ How does she attend to the visual features (or shape) of the distribution in comparing or analyzing graphs?</li> <li>▪ How is her perception of shape influenced or not influenced by outliers?</li> </ul>	Taller, higher, central peaks/bars Highest amounts or bars Skewed Symmetrical Pyramid or inverted V Pyramid or inverted V shape Bell curve Uniform Waves
-------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

*Table 4.1 Questions, Words, and Phrases That Emerged When Selecting Mae's Notices*

### Counting Notices

As previously discussed, Sherin's and van Es's (2009) use of idea units (Jacobs & Morita, 2002) helped me code for Mae's notices. In alignment with Canada's (2004) Evolving Framework, I first separated my data into units that aligned with his concepts of variation. This included for example, his themes of Evaluating and Comparing Graphs: average, extremes (which Mae called outliers), range, and/or shape.

The following Table 4.2 was used to keep track of Mae's notices. Each notice was sorted by the appropriate concept. Its place in Canada's (2004) Evolving Framework of aspects, dimensions, and themes of characterizing thinking about variation was also noted. The total number of notices in each concept was tallied and then a percentage of each concept for its specific data source was calculated.

<b>Data source</b>	<b>Concept noticed</b>	<b>Canada's Framework: dimensions aspects themes</b>	<b>Number of notices</b>	<b>Percent of notices</b>
Professional development session				
Performance tasks/interviews				
Lesson planning				
Postinterview				

*Table 4.2 Chart for Keeping Track of Mae's Notices*

Each time Mae shifted to another concept, it was coded as a different unit. For example, a natural shift occurred within each of the questions embedded in the tasks. Usually, there were multiple concepts brought up in one turn of speaking. Therefore, each concept was coded separately as a notice. Also, when a concept was brought up again in another part of the conversation, it was coded as a separate notice. Exceptions included when those concepts stated at different times were a continuation of the earlier discussion of this concept. These notices were not coded separately as long as they originated from the same data source, such as a professional development session.

This coding process was divided up into discrete segments. There was one for each professional development session, one for each of the performance tasks and interview, one for the lesson planning session, and one for the postlesson interview. See Table 4.3 for how it was done for Mae's performance tasks, lesson planning, postlesson interview, and professional development problems.

<b>Stages in data analysis</b>	<b>Method used for tasks</b>	<b>Method used for lesson planning</b>	<b>Method used for postlesson interview</b>	<b>Method used for professional development</b>
Separating data	By task then by each task's questions  By Canada's (2004) framework and other concepts	By topics of discussion  By Canada's (2004) framework and other concepts	By topics of discussion  By Canada's (2004) framework and other concepts	By each session then by each problem in the session  By Canada's (2004) framework and other concepts
Counting notices	Did not count statements made to clarify or amplify earlier comments in the same task.  Counted statements separately when Canada's (2004) category (concept of variability) changed, even if it was spoken in a statement meant to clarify.	Did not count statements made to clarify or amplify earlier comments in the same topic.  Counted statements separately when Canada's (2004) category (concept of variability) changed, even if it was spoken in a statement meant to clarify.	Did not count statements made to clarify or amplify earlier comments in the same topic.	Did not count statements made to clarify or amplify earlier comments in the same problem.

*Table 4.3 Criteria for Counting Mae's Notices*

Table 4.3 (cont'd)

	Comments that were repeated throughout the different tasks were counted as separate notices, even if they seemed to be similar to a comment made in other tasks.	Comments that were repeated throughout the different topics were counted as separate notices, even if they seemed to be similar to comments made in discussions of other topics.	Comments that were repeated throughout the different topics were counted as separate notices, even if they seemed to be similar to comments made in discussions of other topics.	Comments that were repeated throughout the different problems were counted as separate notices, even if they seemed to be similar to comments made in discussions of other problems.
--	------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

*Table 4.3 Criteria for Counting Mae's Notices*

The table explains the method used to count notices in each of the contexts of Mae's work in this research. What follows is a brief discussion of its use for her performance tasks and her lesson planning.

For Mae's performance tasks the data were separated into segments based on each of the six tasks. Next, the data were sectioned by the questions in each task. Within these sections of the data, Canada's (2004) concepts of variation, either his dimensions or themes, were noted and labeled. In the next stage of analysis, Mae's comments noted on these concepts were read to determine whether they were a continuation of an earlier statement. If this was the case, then it was not counted as another notice. What was considered a continuation of an earlier statement included a comment that was meant to clarify or to amplify the original utterance on the concept. Statements that were considered separate notices were those that, for example, focused upon a new concept in Canada's (2004) Evolving Framework. Comments that were repeated throughout the different tasks were counted as separate notices, even if they seemed to be similar to a statement made in another task.

Task number 5 offers an example of not counting a comment more than once for a concept. In her response, Mae mentioned what she perceived as an outlier but gave it no numerical value. Soon thereafter in the discussion, I asked what the outlier was and she identified the outlying values:

Researcher: Which graph shows more variability in student's heights?

Mae: Graph A has most of the data shaped in a bell distribution while Graph B has, almost has, 3 clusters and one outlier...

Researcher: And what was your outlier for B...?

Mae: 162, and almost 155 even though it is with the group it's still just that one person who is just not there

Researcher: So you said that it was an outlier? I don't know what you meant by 155.

Mae: I don't think that is an outlier, but it is definitely not part of any of the larger clusters that are put together.

(Performance Task 5, Lines 242–269)

These comments were counted as one notice of outliers because all statements made after the first one clarified the original statement.

In the first stages of analysis for lesson planning, the data was also separated into segments based upon the topics of discussion. In contrast to the performance tasks, where there was one separate interview for each task, the lesson planning discussion was not bound or ordered by the structure of the task. Instead, the topics emerged spontaneously and appeared recursively in the lesson planning discussion. It is believed this was based on the nature of lesson planning that was a thinking-out-loud or a figuring-out-and-creating process of the lesson. The list of topics discussed during the lesson planning were brought up and then revisited. This revisiting of the topics was done either by Mae, who was still thinking it through, or by me when I was seeking to clarify or to amplify her thinking. For example, the

following topics were repeated more than once in the lesson planning discussion, which made them less bounded than the tasks. They are not listed in chronological order:

- task selection (Should I do this? Why? What does it explain regarding variability?)
- homework (What aligns with the goal of the lesson?)
- mathematics/variability in lesson (What mathematics or variability does this go for?)
- materials for task (What do I need to teach this lesson?)
- aim of lesson (What do you expect students to be able to do?)
- timing of lesson (What task(s) reasonably fit into this lesson?)
- task content (What to include/not include in the task? Why?)

Within these topics, the concepts that were found in Canada's (2004) Evolving Framework were noted and labeled. These concepts were counted as notices in the same way as her performance tasks.

Next, there was a cumulative tally of these notices showing the percent of notices in each concept of Canada's (2004) Evolving Framework and their general location in the different contexts. This was done on Table 4.1. This table helped me see the frequency and the density of Mae's notices (how many times and where) for the various concepts of variability of data distributions. I used this information to guide the next stage in my analysis, which was to find out what Mae interpreted about her notices. The concepts that are noticed more frequently, and the data sources with the higher density of the notices, warranted my initial attention. In addition, as discussed earlier, in alignment with the sensemaking model

(Weick, 1995; Drake, 2006), I attended to what was manifested in Mae's lesson to ensure the bidirectional flow of my analysis.

### **How Mae Interpreted**

The next stage of analysis focused on how Mae interpreted the variability of data distributions. In this stage, how Mae interpreted the variability based on her notices was addressed. As previously discussed, transcriptions were analyzed and coded to find the concepts that highlighted Mae's thinking about variability in data distributions. These concepts were aligned with research, for example, the conceptual themes of Canada's (2004) Evolving Framework.

After I recorded the percentages of Mae's notices, I needed to select the concepts with which to analyze her interpretations. Canada's (2004) dimension of Evaluating and Comparing Graphs and its themes of average, range and extremes (which Mae called outliers), shape, and spread, came up with the highest percentages of Mae's notices. The questions on Table 4.2 were answered for each of the themes of evaluating and analyzing graphs. When I finished analyzing Mae's interpretation of these themes in this way, I found that much of what she said about the spread of the distributions involved the average, range, extremes (which she called outliers), or shape. Further, I found that most of her comments on spread overlapped with these other themes. Based on this, I did not consider them mutually exclusive. Based on this and not wanting to be redundant in analyzing all of the data, I decided to focus my analysis of Mae's interpretation of variability on how she discussed spread. In analyzing Mae's utterances on spread, the other themes of measures of center, range, extremes (which she called outliers), and shape naturally came out and were analyzed. Again, this was based on her utterances of

spread encompassing the other themes. Therefore, within the context of discussing spread, these themes seemed to warrant their own discussion.

### **What Mae Implemented**

As opposed to what Mae planned or intended, her knowledge of variability that she actually implemented in the lesson was analyzed. Inscriptions on the blackboard and statements made to students were considered for analysis.

In the following part of my data analysis, I transcribed the tapes from Mae's lesson. Using the same procedure as I did for professional development and the learning tasks, I wrote comments, questions, and hunches on the margins of the document. Next, I viewed the videotape of Mae's lesson looking simultaneously for confirmations of my possible assertions, for contradictions to my assertions, and possibly for any new ones.

In summary, I used both these prospective and retrospective processes to analyze my data. I searched to determine what Mae knows about the variability in data distributions through her teaching (action). This was done to confirm what I found she focused upon and how she interpreted it in her performance tasks, her lesson planning, and her professional development sessions. In addition, I looked to affirm it in the opposite direction—that what Mae focused upon in her performance tasks, her lesson planning, and her professional development sessions—was manifested in her lesson. This is the bidirectional perspective I took in studying teacher knowledge. In other words, what Mae knows will be what she noticed in all three data sources, and what she knows will be revealed in how she interpreted her notices. Finally, what she knows will be revealed in her actions (lesson). Conversely, her actions (lesson) will be a window to what Mae focused upon and what she interpreted in the other data sources. Thereby, confirming one another. In this way, I hope to not just triangulate

my results but to take a bidirectional approach to study teacher content knowledge. In doing so, I wished to give a more complete foundation to my findings of what Mae knows about the variability in data distributions.

In the following results chapters a further review of what the Mae noticed and interpreted is discussed with regard to the findings of other researchers. Along with Canada's (2004) research, Hammerman's and Rubin's (2004) and Makar's and Confrey's (2005) work are used. In addition, a curriculum framework for pre-K–12 statistics education—the 2005 GAISE Report (Franklin et al., 2007)—is referenced to place Mae's knowledge of variability in the broader field of statistics education.

## **Chapter 5**

### **A Closer Look at What Mae Noticed**

As discussed in previous chapters, I used a sensemaking model to operationalize Mae's knowledge of variability in data distributions (Weick 1995; Drake, 2006). I focused on what Mae noticed, interpreted, and implemented to discuss her content knowledge of this topic. My intention was to give a more complete view of her content knowledge by analyzing it in a bidirectional way—connecting what she knew prior to teaching to what she knew when teaching. The sites studied prior to teaching included what was noticed and interpreted in professional development, lesson planning, and performance tasks. What Mae noticed and interpreted in these sites was connected to what she implemented in her teaching. This was done to portray a dynamic view of what she exhibited knowing prior to her teaching.

As previously discussed, I used Canada's (2004) Evolving Framework of EPSTs' thinking on variation to guide the coding of the notices. This was done to extend the use of the framework to include in-service middle school teachers. The framework provided a list of concepts I used to code Mae's notices. Based on the nature of the problems and tasks used in the research, parts of his framework seemed to be more applicable than others. The entire framework was used, however, in the initial coding of notices. The framework is depicted in Table 5.1. The bolded items represent what Mae seemed to notice. This chapter discusses the results of what Mae noticed in her professional development, lesson planning, performance tasks, and postlesson discussion.

Evolving Framework
<p><b>[1] Expecting Variation</b></p> <p>A] Describing What is Expected</p> <ul style="list-style-type: none"> <li>i) Concerning Expected Value</li> <li>ii) Concerning Repeated Values</li> <li>iii) Concerning Range or Extremes</li> </ul> <p>B] Describing Why (Reasons for Expectations)</p> <ul style="list-style-type: none"> <li>i) Involves Possibility or Likelihood</li> <li>ii) Involves Experiential Reasoning</li> <li><b>iii) Involves Proportional Reasoning</b></li> <li>iv) Involves Distributional Reasoning</li> </ul> <p><b>[2] Displaying Variation</b></p> <p>A] Producing Graphs</p> <ul style="list-style-type: none"> <li>i) Technical Details</li> <li>ii) Characteristics of the Distribution</li> </ul> <p>B] Evaluating and Comparing Graphs</p> <ul style="list-style-type: none"> <li>i) Focus on Average</li> <li>ii) Focus on Range or Extremes</li> <li><b>iii) Focus on Shape</b></li> <li>iv) Focus on Spread</li> </ul> <p>C] Making Conclusions about Graphs</p> <ul style="list-style-type: none"> <li>i) Emphasizing Decisions in Context</li> <li><b>ii) Emphasizing Consistency or Reliability</b></li> <li>iii) Emphasizing Level of Detail &amp; Usefulness</li> </ul> <p><b>[3] Interpreting Variation</b></p> <p>A] Causes and Effects of Variation</p> <ul style="list-style-type: none"> <li>i) Definitions &amp; Descriptions</li> <li>ii) Examples</li> </ul> <p>B] Influencing Expectations and Variation</p> <ul style="list-style-type: none"> <li>i) Naturally Occurring Causes</li> <li><b>ii) Physically Induced Causes</b></li> </ul> <p>C] Effects of Variation</p> <ul style="list-style-type: none"> <li>i) Effects on Perception</li> <li>ii) Effects on Decisions</li> </ul> <p>D] Influencing Expectations and Variation</p> <ul style="list-style-type: none"> <li>i) Quantities in Sampling</li> <li>ii) Number of Samples</li> </ul>

*Table 5.1 Applicable Parts of Canada's (2004) Evolving Framework*

The variation talk of Makar's and Confrey's (2005) in-service teachers was also used to select Mae's notices. Specifically, phrases discussed by their teachers, such as *evenly distributed* and *clustered together* were sought out and highlighted as notices.

These notices were collected and tallied according to their respective location in Mae's work: lesson planning, performance task with interview, postlesson discussion, and professional development problems. The tallies were then used to calculate the percent of times this concept appeared in each particular location. The total for each concept was based on the total number of notices for that specific location, such as lesson planning. See Chapter 4 for a further discussion on how these notices were counted. The results of the identifying and sorting of these notices are depicted in Table 5.2.

<b>Alignment with Canada's (2004) Evolving Framework</b>	<b>Lesson planning</b>	<b>Tasks with interview</b>	<b>Postlesson interview</b>	<b>Professional development 1–4</b> <b>Inv 1 Prob 1.1–1.3</b> <b>Inv 2 Prob 1.4–2.2</b> <b>Inv 2-3 Prob 2.2–3.4</b> <b>Inv 4 Prob 4.2</b>
Causes for variation:  Measure-ment error/ Natural causes	<b>0%</b>  Although the task of comparing data distributions of head sizes could involve measurement error, it did not come up in planning the lesson.	<b>6%</b>  16% (1/6) of the tasks explicitly asks about the reasons for different results when collecting measure-ment data.  <b>Task 1</b>	<b>12%</b>  Task entailed students measuring head sizes.	<b>15% 0% 4% 0%</b>  Investigation. 1 discussed trends, patterns and differences in: 1.1 candy colors in M&M's 1.2 immigrants to US 1.3 head measurements.  Investigation 3 discussed measurement data.  <b>PD Summary 9.5%</b>

*Table 5.2 The Results of Mae's Notices*

Table 5.2 (cont'd)

<p>Comparing/ Analyzing graphs and Evaluating graphically displayed data:</p> <p>Focus on average, range and extremes, shape, spread some with proportional reasoning (see below*)</p>	<p><b>45%</b></p> <p>Task involved describing a data distribution and finding the typical value of it.</p>	<p><b>67%</b></p> <p>83% (5/6) of the tasks explicitly asked to compare graphs.</p> <p><b>Tasks 2, 3, 4, 5, 6</b></p>	<p><b>54%</b></p> <p>Task involved finding the typical head size for a student. The data collected included values perceived as outliers.</p>	<p><b>35% 66% 76% 74%</b></p> <p>Investigation 1 involved analyzing different kinds of data:</p> <p>1.1 categorical data 1.2 data as counts 1.3 measurement data</p> <p>This is in contrast to the tasks, lesson planning, and Inv. 2–4, that used predominantly measurement data.</p> <p>Using measurement data facilitated describing data distributions with average, range, extremes/outliers, spread, and shape.</p> <p><b>PD Summary 60%</b></p>
<p>Expect/ predict (includes influences):</p> <p>-Variation in data distribution: proportional distributional, and/or probabilistic reasoning</p> <p>-Sample size</p>	<p><b>20%</b></p> <p>In choosing the content of her task, Mae mentioned having the students predict from the collected data to the whole class or school.</p>	<p><b>0%</b></p> <p>No tasks asked to predict</p>	<p><b>15%</b></p> <p>Mae discussed predicting as something she would have done if she had more time, and as something she wanted to include in the homework.</p>	<p><b>0% 11% 2% 5%</b></p> <p>No problems asked to predict.</p> <p><b>PD Summary 5.5%</b></p>

Table 5.2 (cont'd)

<p>Displaying variation:</p> <p>-Technical detail/level of usefulness</p> <p>-Produce/Choose appropriate graph</p>	<p><b>13%</b></p> <p>The aim of the planned lesson included producing an appropriate graph.</p> <p>While discussing the content of the task, Mae mentioned the details and usefulness of the graphs.</p>	<p><b>11%</b></p> <p>50% (3/6) of the tasks asked explicitly for details or usefulness of graph.</p> <p><b>Tasks 2, 3, 5</b></p>	<p><b>3%</b></p> <p>There was one discussion on scaling the line plot for the head measurement data.</p>	<p><b>15% 11% 4% 5%</b></p> <p>Investigation 1: All problems involved creating graphs. This facilitated a discussion on the technical details and sometimes usefulness of a graph.</p> <p><b>PD Summary 7.5%</b></p>
<p>Proportional reasoning*:</p> <p>-Unequal size data sets</p> <p>-Equal size data sets</p>	<p><b>3%</b></p> <p>Mae discussed proportional reasoning when selecting the task (that is comparing classes based upon clusters).</p>	<p><b>11%</b></p> <p>Mae answered 33% (2/6) of the questions using proportional reasoning</p> <p>-one with equal size data sets</p> <p>-one with unequal size data sets.</p> <p><b>Task 3, 6 respectively</b></p>	<p><b>3%</b></p> <p>Mae mentioned that if there was more time in the lesson she would have compared unequal data sets (boys to girls).</p>	<p><b>5% 11% 6% 11%</b></p> <p>Investigation 1: proportional reasoning was used to analyze trends in immigration</p> <p>Investigation 4: proportional reasoning was used to compare distributions of roller coaster speeds.</p> <p><b>PD Summary 4%</b></p>

Table 5.2 (cont'd)

Explicit connection to consistency:  -in lesson planning  -in categorical data	<b>3%</b>  When choosing tasks, Mae discussed consistency in analyzing categorical data.	<b>6%</b>  16% (1/6) of the tasks asked explicitly about the consistency of a data distribution.  <b>Task 1</b>	<b>0%</b>  There was no explicit connection to consistency or categorical data.	<b>10% 0% 4% 5%</b>  1.1 Used the phrase “evenly distributed” to discuss variability in categorical data (M&M colors). 1.3 Used “more evenly distributed” to discuss clustering in measurement data (head measurement). Inv. 3 asked about consistency in computer reaction times.  <b>PD Summary 4%</b>
Connection to variability:  in lesson planning and the postlesson interview – define/ describe	<b>19%</b>  Created a lesson focused on students’ understanding of variability of data distributions.	<b>11%</b>  Approximately 33% (2/6) of the tasks questioned explicitly about the variability of a data distribution  <b>Tasks 2, 5</b>	<b>12%</b>  Mae was insecure regarding the meaning of variability.	<b>15% 0% 4% 0%</b>  1.1 Defined variability 1.2 Asked to describe variability by writing comparative statements 1.3 Asked specifically about variability  <b>PD Summary 9.5%</b>

Table 5.2 The Results of Mae’s Notices

The bulk of Mae’s notices fell into Canada’s (2004) dimension of Comparing and Analyzing Graphs. This dimension included focusing on the following themes to discuss variability: average, range, extreme values (which Mae called outliers), and shape. This is not surprising. All of Mae’s work in this research involved comparing and analyzing data distributions. She chose a problem for her lesson that used measurement data to describe a

data distribution. Most of Mae's performance tasks also required comparing or analyzing distributions. Further, throughout her professional development problems she analyzed and compared distributions of categorical and numerical data. Finally, the problem she chose for her lesson included analyzing the distribution of her students' head measurements. Specifically, she wanted her students to find one hat size to fit all of their head measurements. Her discussions in these various sites of teaching included the themes of average, range, extremes (which Mae called outliers), spread, and shape.

In Mae's lesson planning, she had plenty of opportunity to discuss these themes in analyzing data distributions. Her lesson planning involved selecting the aim of her lesson, the task, the components of the task, and the homework. Within these, Mae discussed the various themes in her sensemaking of variability: average or measures of center, range, extremes (which Mae called outliers), spread, and shape. Mae noticed these themes in her sensemaking of variability 45% out of her total notices. This 45% is out of the notices counted throughout the concepts on variability noticed in her lesson planning. (See above listing of all concepts from Canada's (2004) Evolving Framework used for this research.) In a not-so-close second, predicting from her data distribution followed this category with 20% of the notices. This was based on a brief discussion Mae had on using her class's data to predict the head size for the rest of the school.

Based upon the opportunities her lesson planning afforded her to discuss variability, the 45% in Canada's dimension of Evaluating and Analyzing graphs and its accompanying themes of average, range, or extremes (which Mae called outliers), spread, and shape was not surprising. Specifically, Mae's discussions involved deciding what measurement data would support her students in understanding variability. Also, in her lesson planning deliberations

she determined how reference lines would enable her students to find the typical value. These topics of her lesson planning brought into view what she noticed in terms of variability.

Mae's performance tasks also gave her much opportunity not only to analyze graphs, but also to compare them. (See Appendix A for actual performance tasks.) Five out of six or 83% of the tasks required Mae to do this. As a result, she noticed average, range, values she perceived as outliers, spread, and shape 67% of the time when solving these tasks. This was a greater percentage of notices in these themes than in her lesson planning (45%), postlesson interview (54%), or professional development sessions (60%). This was probably based on the explicit and direct way these tasks required Mae to discuss these themes when comparing and analyzing distributions.

Since most of Mae's performance tasks (two through six) required her to compare or analyze distributions, it is not surprising that she used these themes of average, range, values she perceived as outliers, spread, and shape to do so. She used these themes 67% of the time she solved these performance tasks. When she used these themes, it was in response to the task specifically asking her to. No other conceptual focus from Canada's (2004) Evolving Framework came close percentage wise to this category. Thus, as might be expected, what Mae noticed in her performance tasks was prompted by the problems in these tasks.

This also holds true for her professional development problems from the *Data Distributions* text (Lappan et al., 2006). These problems were written to help the learner make sense of the variability in data distributions. Based on this, the percentages of Mae's notices in professional development reflected this. Overall, when comparing and analyzing data distributions, she spoke about average, range, extremes (which she called outliers), spread, and shape 60% of the time.

Again, this is not surprising. These problems included comparing and analyzing data distributions of categorical and numerical data. Based on this, the amount of times Mae described distributions with average, range, extremes (which she called outliers), spread, and shape depended upon the type of data that was used. In the first professional development session, categorical data was used. Therefore, in contrast to the other three professional development sessions, it accrued the lowest percent of notices of these themes of variability. It had 35% compared to the 66%, 76%, and 74% in sessions two to four respectively. However, for the professional development sessions overall 60% of the notices fell into these themes of variability. Percentage-wise there were no other categories that came close to this one when solving her professional development problems. Perhaps this was based on the intent of the problems as previously discussed.

Finally, in Mae's postlesson interview, she focused upon average, range, values she perceived as outliers, and spread in 54% of her total notices. The reason for this percentage might be based on the nature of her lesson in which her students created a data distribution of their head measurements. She wanted them to determine whether it was possible to find a typical hat size for the class. Her discussions with her students included describing the distribution's measures of center, values she perceived as outliers, and range. Also, in a not-so-close second, 15% of Mae's notices involved predicting. Specifically, she hoped that in a follow up lesson on variability her students would predict from their typical head size to the whole school's typical head size.

In summary of the counting of Mae's notices, the preponderance of her notices were a part Canada's (2004) Evolving Framework of his EPSTs' reasoning about variability; specifically, Mae's notices were a part of his themes from Evaluating and Comparing Graphs.

These themes included average (measures of center), range or extremes (which Mae called outliers), and shape of the data distributions. They represented most of Mae's notices in her lesson planning, performance tasks, postlesson discussion, and professional development problems. As a result, these themes were chosen initially as the focus for Mae's sensemaking of variability and are considered her *sensemaking notices*.

In addition to Canada's (2004) themes, other notices of Mae's were considered her sensemaking notices. These other notices did not receive as high a percentage as the themes did. For a number of reasons, *partitioning* and *defining and describing variability* were also considered part of her sensemaking notices.

In spite of the fact that partitioning and defining and describing variability did not receive high percentages of Mae's notices, they were considered potentially fruitful in Mae's sensemaking of variability. This might also be said of Canada's (2004) themes of: extremes (which Mae called outliers) and range, because they are included in describing variability (Lappan et al., 2006, p. 5). However, this might not be the case for his other themes—measures of center and shape—because in and of themselves they are not conventional means of making sense of variability. In contrast to this, partitioning and defining and describing variability could prove to be productive in making sense of variability. The following paragraphs discuss why this might be so. References are made to the field of statistics and statistics education as part of this discussion.

Partitioning was chosen as one of Mae's sensemaking notices in part because it is a strategy used in the *CMP Data Distributions* text that focuses on describing variability (Lappan et al., 2006, p. 5). The authors of this text state that in comparing parts of the distribution, it is useful to partition or divide the distribution (Lappan et al., 2006, p. 104).

One way the authors suggest doing this is by identifying benchmarks and drawing reference lines on the distribution. When using these reference lines both the number and percent of data above and below them can be analyzed (p. 104–105). In this way, the variability of the data or how the data is dispersed throughout the distribution can be discussed.

Another reason for considering partitioning as one of Mae's sensemaking notices is that partitioning is also listed as a way of analyzing distributions in the statistical investigation process (Lappan et al., 2006). Recall the concept map discussed in Chapter 2 that named analyzing distributions as the third part of the statistical investigation process. In that map partitioning the data was listed as a way to analyze distributions, such as by part-whole or relative frequency reasoning. This is not to say that its placement on the concept map in itself warrants partitioning as being potentially fruitful in Mae's sensemaking. Characterizing measures of center and shape are also on the map. Yet, they are not in and of themselves conventional means of making sense of variability. Therefore, they might not prove equally fruitful in Mae's sensemaking of variability. Nevertheless, how partitioning has been used in statistics education research in making sense of variability might be further evidence of its potential fruitfulness in Mae's sensemaking of variability.

In addition to partitioning being chosen as a sensemaking notice, because it is integral in the statistical investigation process, partitioning is also a strategy used by teachers when making sense of variability. Hammerman and Rubin (2004) reported on their teachers' use of a computerized partitioning tool called binning. Through their use of this tool, the teachers were able to deal with variability in analyzing data (p. 17). In

summary, partitioning as a way of analyzing data distributions and as a strategy used by teachers in statistics education, made it seem potentially fruitful to include as part of Mae's sensemaking notices.

In addition to partitioning, defining and describing variability also was chosen as one of Mae's sensemaking notices. Even though it did not receive a high percentage of her notices, defining and describing variability was seen as integral to Mae's making sense of variability. How Mae defines and describes variability can give insight into the sense she is making of it. Additionally, defining and describing variability was chosen because it is a part of Canada's (2004) themes that emerged from his EPSTs' reasoning about variation. Therefore, defining and describing variability seemed potentially fruitful as well to consider in Mae's sensemaking of variability.

In summary, the final list of Mae's sensemaking notices includes:

1. partitioning
2. measures of center
3. range and values perceived as outliers
4. shape
5. defining and describing variability

Mae's sensemaking notices represent a mixture of concepts, definitions and descriptions, and a strategy used in the fields of statistics and statistics education. For reasons previously discussed, all of these were considered a part of her sensemaking notices. Yet, each of these sensemaking notices might not prove to be equally fruitful in Mae's sensemaking of variability. The next chapter discusses Mae's use of her sensemaking notices in her other sensemaking practice—interpreting. This chapter

answers the second research question: What does Mae seem to interpret about variability in data distributions?

## Chapter 6

### A Closer Look at Mae's Interpretations of Variability

As discussed in Chapter 5, from an analysis of Mae's first sensemaking practice, *noticing*, five sensemaking notices emerged. To varying degrees these sensemaking notices were perceived as possibly being fruitful in her second sensemaking practice, *interpreting*. This chapter focuses on Mae's interpretation of variability of data distributions during her lesson planning, professional development, and solving her performance tasks. Using her five sensemaking notices as discussed in Chapter 5, this chapter is set up to explore how these five sensemaking notices emerge across Mae's utterances when interpreting variability.

In this findings chapter Mae's interpretations of variability are situated in reference to a statistics education framework and research. In particular, the 2005 Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report (Franklin et al., 2007; see Appendix B), and the research addressed in the literature review is referenced: Canada (2004), Hammerman and Rubin (2004), and Makar and Confrey (2005). Reports, such as the 2005 GAISE Report (Franklin et al., 2007) propose a curriculum framework for statistics education for pre-K–12. Specifically, the framework proposes “‘must-have’ competencies for graduates to ‘thrive in the modern world’” (p. 4). It intends to complement other standards documents such as the NCTM *Principles and Standards* (2000). In fact, its foundation rests on them. The 2005 GAISE Curriculum Framework (Franklin et al., 2007) seeks to help teachers; specifically to help those teachers who might not see how the overall statistics curriculum provides a developmental sequence of learning experiences. To that end, it offers them a conceptual structure for statistics education that creates a coherent picture of the overall curriculum.

The 2005 GAISE Curriculum Framework (Franklin et al., 2007) provides levels to support the developmental stages of these conceptual structures. However, these levels are not meant to signify grade levels. If a learner has limited or no experience with a concept they are expected to be at Level A. Here are examples of analyzing data at each level.

- Level A: What type of music is most popular among students in our class?

A bar graph is used to display the number of students who choose each music category.

- Level B: How do the favorite types of music compare among different classes? For each class, a bar graph is used to display the percent of students who choose each music category. The same scales are used for both graphs so that they can easily be compared.

- Level C: What type of music is most popular among students in our school?

A bar graph is used to display the percent of students who choose each music category. Because a random sample is used, an estimate of the margin of error is given.

In this findings chapter, these levels are used as a reference to discuss Mae's interpretation of variability. Specifically, Mae's interpretations are discussed along the continuum of these levels. This is not intended to assess her, but rather to help understand where her interpretations lie in the landscape of statistics education.

When her interpretation falls at Level A, it does not indicate knowledge of lower value. Instead it signifies her knowing the foundational concepts on which Levels B and C are built. In addition, Mae was not expected to discuss Level C concepts. Concepts in Level C involve formal measurement of variability, chance variability, and variability in model fitting. Mae

was not charged to teach these topics. Further, throughout this discussion of findings, it is assumed the knowledge Mae exhibited is not indicative of her complete or incomplete knowledge. Rather, it is considered to be the knowledge that she specifically revealed at various occasions in this research.

In each of the commentaries, along with the 2005 GAISE Curriculum Framework (Franklin et al., 2007), Mae's interpretation of variability is discussed in light of recent research. In some instances, Mae's interpretation aligned with other researchers' findings and other times it diverged. As stated previously, the research addressed in the literature review is used: Canada (2004), Hammerman and Rubin (2004), and Makar and Confrey (2005).

Further, the different data sources used in this research were noted. They included her professional development, performance tasks, and lesson planning problems and experiences. The sources or locations for each of her utterances are indicated. Also, in that reporting, the specific problems she worked on are described. This is intended to help give a sense of where and when Mae's sensemaking of variability emerged. However, based upon the intent of this research to study Mae's sensemaking of variability—operationalized as what she noticed, interpreted, and implemented—these data sources were not reported chronologically. Instead the exemplar utterances relevant to each of the five sections were selected irrespective of the source.

In short, the five sections are based on the following: the first, partitioning, is based on a common strategy described in statistics education research. The succeeding three are based on Canada's (2004) Evolving Framework (including themes of Evaluating and Comparing Graphs): measures of center, range and extremes (which Mae called outliers), and shape. The fifth section involves defining and describing variability, and is based on Mae's constructing

definitions of variability. In the commentaries, Mae's interpretation is discussed against the 2005 GAISE Curriculum Framework (Franklin et al., 2007) and the research. The use of these is intended to give the reader an indication of where Mae's interpretation lies in the findings and expectations of the statistics education field.

### **Mae Used Partitioning as a Way to Interpret Variability**

Partitioning is discussed in the research literature as a strategy for making sense of variability. For statistical analysis, statisticians partition data in many statistical representations. For example, histograms and box plots partition the data into groups. In this dissertation study during professional development, Mae used partitioning to compare data distributions of roller coaster speeds.

When partitioning, Mae used proportional reasoning. She partitioned the data in various ways that included moving from using multiple partitions or reference lines to single partitions or benchmarks. Her process of using partitioning related to the findings of statistics education research and the requirements of the 2005 GAISE Curriculum Framework (Franklin et al., 2007). One example is when she solved professional development Problem 4.2C. This problem asked Mae to determine which roller coaster was faster, wood or steel.

#### **Multiple Reference Lines**

When solving this problem, Mae began with partitioning the distributions of roller coaster speeds using multiple reference lines. She separated the distributions into four sections: 0–29, 30–59, 60–89, 90–120 miles per hour because 120 is divisible by four. Initially Mae compared the percent of roller coasters that were in these discrete sections, "...so I got that 11% of the steel was below 30 and only 4% were wood; and from 30–59,

63% [of the steel] were at that speed and then 84% of the wood were at that speed; and then that's like I found the percentage for each interval. So for 60–89 is 25[%] for the steel and 12[%] for the wood, and then 90–119 was 2[%] for the steel and 0[%] for the wood and 0[%] for both of them at 120” (Professional Development 4, lines 450–4). Comparing these subsections of the distribution did not seem to help Mae in determining the roller coaster with the faster speed. She then moved on to using single partitions or benchmarks to compare their speeds.

### **Single Partitions or Benchmarks**

Single partitions or benchmarks appeared to enable Mae to make more sense of the variability when comparing the distributions. She used them in two ways. One of her strategies for using the single benchmarks was to compare large chunks of the distributions to each other. She stated, “Yeah, it's [the % using multiple reference lines] a lot of numbers; and I couldn't figure out what statement to write [that concludes which roller coaster is faster], but I know that 88% of them [wood roller coasters] would go from 0–69 and only 73% of steel, which is a large part of 100%” (Professional Development 4, lines 470–2). Here Mae seemed to take into account most of the data in comparing the distributions. It appeared that to her the 88% was a large enough proportion of the roller coasters from which to draw a conclusion about the coaster with the faster speed.

Mae also used the mean as the benchmark to compare the speeds of roller coasters. At the end of solving Problem 4.2, she partitioned the distribution at the mean. She spoke about using the percentage of data above the means to determine the faster roller coaster speed.

Mae: If I would have looked at the mean, because that is going to be the average of all of the roller coasters and see how that was, in that interval, how many were above that, just in that interval or below.

Researcher: You mean that Connie did put the mean as your position point?

Mae: For that interval, well 'cause if the mean is 53, and you want to look like how many are above 53 in that interval and get that percentage; and if that percentage is greater for the steel than for the wood, then I would say steel is faster because more of them are faster than the mean for the wood.  
(Professional Development 4, lines 648–666)

In short, Mae used partitioning to make sense of the variability of the roller coaster speeds. First she compared discrete subsections of the distributions, then she compared larger chunks of the distributions, and finally she compared the proportion of data above the mean in these distributions. Mae's flexible movements in making sense of the variability using partitioning were meaningful for a number of reasons.

First, Mae used proportional reasoning to compare the roller coaster distributions. This is important because it is a powerful technique used widely throughout statistics. Examples include rescaling variability in standard deviation units when calculating  $z$ -scores, as well as using percentages instead of counts to deal with differences in sample size (Hammerman and Rubin, 2004, p. 21).

Second, Mae not only used proportional reasoning, but she used it appropriately when comparing distributions of unequal size. (There were 50 wood roller coasters and 100 steel roller coasters represented in their respective distributions.) This is significant because it was not true of teachers that were reported in other research. Hammerman and Rubin (2004) found their teachers persisted in using counts when comparing unequal-size data distributions.

Third, Mae did not use the mean alone in comparing the roller coaster distributions. For her, it was not the sole determining factor in comparing the speeds of the roller coasters. Instead, she looked to the variability of the data around the mean to determine which roller coaster was faster. The authors of the *Data Distributions* text used in this study state that using only a measure of center to analyze or compare distributions can be misleading (Lappan et al., 2006). Therefore, Mae's beginning to take into account how the data varied in relation to the measure of center was important in comparing the distributions.

Finally, Mae's sensemaking of variability was similar to the way in which Hammerman and Rubin (2004) found that their teachers made sense of it when using technology. Mae's progression from discrete subsections to a larger chunk replicated how Hammerman's and Rubin's (2004) teachers dealt with the complexity of variability. These researchers found that their teachers tried to make sense of variability by narrowing the perspective of the data. The teachers did this by using smaller sections of the distribution to reduce the variability to attend to, on one hand, and expanding the scope of the data to include a minimum number of data values on the other. The minimum level was the lowest amount of data that the comparer was comfortable with that he or she felt represented the distribution. It appeared that Mae's shift from comparing smaller subsections of the distributions to a larger chunk of it exhibited this tendency. She tried to make sense of the variability of the roller coaster speeds by shifting from comparing the smaller subsections of the distributions to a greater section of it in order to make a comparison.

Here Mae compared large chunks of the distributions to each other. Yet, she did not mention where most of the data was clustered. This clustering of most of the data can be

considered a *modal clump* that indicates the variability along with the location of measures of center (Konold et al. (2002) as cited in Makar's and Confrey's (2005) research). Mae's use of main cluster or modal clumps (Konold et al., 2002) in making sense of variability is discussed more in the next section on her use of measures of center in interpreting variability.

Regarding the 2005 GAISE Curriculum Framework (Franklin et al., 2007), Mae's use of the mean to compare roller coaster distributions began to align with Level B. This level of the framework suggests introducing learners to the idea of comparing data values to a central value. In a primitive way Mae began to address this by coupling the mean (the distribution's location) with the percentage of data above it, which is good. Yet, she could still progress in her sensemaking of variability by measuring the distance the data values are from the mean.

In summary, Mae's use of the strategy of *partitioning* gave insight into her interpretation of variability. Her use of percents of data indicated that she used proportional reasoning as a part of discussing the variability of a distribution. In addition, her use of reference lines to make sense of variability indicated that she interpreted larger sections of the distribution as being more useful for comparing than smaller discrete sections. Finally, Mae's seeing the percent of data above and below the mean indicates her interpretation of the mean as an important benchmark when discussing at a somewhat gross level the variability of the data. This snapshot of Mae's sensemaking and the interpretations she seemed to make of variability in this process are displayed in Table 6.1. As Mae continued to interpret variability, measures of center were further used as an integral part of her sensemaking. This is discussed in the next section.

What Mae did during sensemaking with partitioning	What Mae seemed to interpret about variability
a) Used percent of data appropriately when analyzing unequal size data distributions b) Used multiple reference lines to make sense of variability c) Saw the percent of data above and below the mean	Proportional reasoning is a part of discussing variability  Larger sections (encompassing more data) of distributions are more useful for comparing than smaller discrete sections  The mean is an important benchmark to discuss the variability of data (percent above and below it)

*Table 6.1 A Snapshot of Mae's Sensemaking and Interpretation of Variability*

### **Mae Used Measures of Center as a Way to Interpret Variability**

Measures of center were found to be important for Mae in three ways when interpreting variability. First, they were an essential part of planning a lesson for her students on variability. When she chose a task for her lesson, measures of center were the focal points or benchmarks in determining the percent of data above and below them. Second, measures of center played an integral role in her decision to use reference lines as a part of this lesson in order to identify the main cluster or modal clump (Konold et al., 2002). And, lastly, Mae connected the relative locations of the mean and the median to the distribution's variability in an unconventional way. This last part of her sensemaking of variability was unconventional in that a statistician would have connected the distribution's shape to the relative locations of these central measures. The following describes the ways that Mae used measures of center in making sense of variability.

### **Mae Chose a Classroom Problem for Her Lesson Based Upon Measures of Center**

Measures of center were important to Mae when choosing a problem for her upcoming lesson on variability. She picked a problem that used measurement data over a problem that

used categorical data. Initially, she grappled between using a problem with measurement data (students' head sizes) versus a problem that used categorical data (colors of M&M's). She wondered what each one would afford her students in terms of learning the concept of variability. Mae did not think she could use the problem with categorical data for her lesson on variability. She stated:

Well this one (categorical problem) is hard for me to understand how the variability works. Whereas opposed to when they measure their heads, you could use the definition of what it means to be clustered together and farther apart and outliers. This one (measurement problem) would be easier. So yeah, I think this (measurement problem) would probably be better; only because it is interesting and it is easier to see above and below the mean and the mode.

(Lesson Planning, lines 29–30)

Here Mae chose to use the measurement problem over the categorical problem for her classroom lesson on variability. Her decision was based on what variability meant to her—being clustered together, farther apart, and with outliers. However, she also based it on what the type of problem afforded regarding measures of center. For her, seeing above and below the mean and the mode was a prominent benefit to using measurement data in her classroom problem.

Mae's work is important. Level B of the 2005 GAISE Curriculum Framework states that learners need to be introduced to the idea of comparing data values to a central value, such as the mean or the median, and quantifying how different the data are from these central values (Franklin et al., 2007). Mae's work here is approaching Level B requirements. She did know the mean is important to discuss the variability of the distribution. Yet there is still room for her to determine how different the data are from it, for example, using the Mean Absolute Deviation (MAD), variance, or standard deviation. Nonetheless, Mae's work indicates that she

knows the importance of measures of center as reference points in making sense of the variability of the distribution.

Mae's interpretation of variability in relation to measures of center and the percentage of data around them are important considerations when determining the faster speeds of roller coasters. Yet, according to Makar and Confrey (2005) a more robust understanding of the context and an examination of the whole distribution are desired in using descriptions to make comparisons. These descriptions might include the clustering of the data, its gaps, and its outliers in light of the context. Mae moved toward doing this when she chose the components of the problem for her classroom lesson on variability. In determining the components of the problem, Mae used measures of center to make sense of variability via the clustering of the distribution. This is discussed in the following.

### **Mae Determined That the Purpose of Reference Lines Is to Find the Main Cluster or Modal Clump**

After Mae selected the type of problem for her classroom lesson, she set out to determine the components of the problem. She was less challenged constructing some parts of her classroom problem over others. This was based on using some questions directly from Problem 1.3 Variability in Numerical Measurement that required collecting the head sizes of students. The particular questions she chose were from Part C:

1. What are the minimum and maximum values of the distribution?
2. What is the range of the distribution?
3. Do you think the range of the measurements is great enough that recommending a single cap size would be difficult? Explain.

Are there any unusually high or low data values, or outliers? If so, what are they?

5. Do some or most of the data cluster in one or more locations? If so, where does this occur?
6. Are there gaps in the data? If so, where do they occur?
7. What would you describe as the typical head size for these data? Explain.
8. Use these ideas to describe the variability in the data.

Mae decided to use these questions as part of her classroom problem. Yet, she labored whether to have her students partition the distribution with reference lines to answer the questions. At first she could not see their purpose in answering question 7—determining the typical head size of the students.

Mae: But I want to know if I should include this piece over here in Investigation 4 [Are Steel Roller Coasters Faster than Wood Roller Coasters?]. Where is the thing that we did the other day? It was only one part. I know that we only did the reference lines. Would that be applicable?

Researcher: What would you want to get out of it by doing that? Just to get an idea about variability?

Mae: Well, it would be more than nice...if it could help you answer question 7: What would you describe as a typical head size for the data? Would that be helpful? No, that would be the mode, and you would not have to concern yourself about the ranges or the reference points. I don't think it would correlate.

(Lesson Planning—Task Content, lines 183–90)

Here, in the beginning of her lesson planning, the mode, the most frequent data value, was the typical value for Mae. However, later on in her planning she reconsidered this when she was still trying to figure out how reference lines would be helpful.

Researcher: So you are saying the benchmarks (another term for reference lines) would be used if you were comparing multiples like five data sets. Is that what you are saying?

Mae: No, I don't think so. That's more along for the mean and the median and where you are more likely to find that information not so much to understand the definition of what it means to be variability. So I don't know.

So we looked at this book (*Data Distributions*); and then, well, we were just looking at the speeds and it was more for just the clusters. So I guess that would be easier to explain like if you looked at the range of (pause) I guess it could be helpful more along the lines of to understand what it means to be clustered.

Here Mae realized that understanding what it means to be clustered was a purpose for partitioning the distribution using reference lines. She went on to clarify what they meant and when she would use them.

Mae: But that is going to be a lot of information in one period. That would be too much.

Researcher: So what would you like to do then?

Mae: We still have a couple more days of school. I might even add on the benchmarks, the reference lines. But that would be another day because that would be a lot to do in one day. Only so that they could say, because based on this (distribution of students' head sizes): What would we expect for the whole school using the reference lines? Which group of data is more clustered? And that would help them to distinguish: Okay most of them are in that range or this interval. So yeah.

(Lesson Planning—Task Content, lines 334–49)

Here in this last utterance on reference lines, it seems that Mae saw the value in using them in her classroom lesson. She would have her students find the typical head size for the whole school based upon them. She would expect this to be a process for them. Through partitioning with reference lines, she expected her students to identify where more of the data was clustered in the distribution. Then, from there the students would determine the range where most of the data were located, that is, the modal clump (Konold et al., 2002).

Mae's utterances imply that she tolerated variability in the typical value. That is, she expected that the predicted value for the school's head size would fall into a range of values in contrast to a single value. Mae's work is similar also to Makar's and Confrey's (2005) findings. Their preservice teachers found modal clumps that are the range of data in the heart of a distribution of values (Konold et al., 2002). These modal clumps allowed their preservice

teachers to express simultaneously the average and how variable the data are (Konold et al., 2002). This view might be considered foundational to discussing the more formal measure of variation in the data that is required at Level B of the 2005 GAISE Curriculum Framework (Franklin et al., 2007).

In short, when determining the interval on the distribution where most of the data are clustered, that is, the modal clump (Konold et al., 2002), Mae began to discuss its variability. Earlier she stated that reference lines would only be helpful to find where the mean and the median most likely are. However, she thought further and began to discuss their use in describing the distribution's variability. Specifically, she saw their use in determining the range of typical values for the main cluster or modal clump (Konold et al., 2002) of the distribution. She stated earlier that variability was what it means to be clustered together or farther apart including what she called outliers. So for her, it seems that identifying the main cluster pointed toward how the data varied within the distribution.

### **Mae Used Relative Locations of Measures of Center When Discussing Variability**

Mae again used measures of center when interpreting variability. In an unconventional way she focused on the relative locations of the mean and the median when discussing the variability of the distribution. To her, where they were situated relative to each other was influenced by the pattern of where most of the data clustered and the deviations from this pattern.

Problem 3.4, Comparing Larger Distributions, required Mae to compare all students' fastest video game reaction times to their slowest video game reaction times. (Note the ranges of these data values provided in the problem were: all students' fastest times 0.58 to 1.19 seconds and all students' slowest times 0.85 to 2.48 seconds.) When solving this problem,

Mae mentioned that the main clustering of the data without any values she called outliers, could be related to the close proximity of the mean and the median.

In this problem Mae saw that the relative location of the mean and the median interacts with how most of the data are clustered in relation to them including whether there were extremes (which Mae called outliers). In discussing the students' fastest video game reaction times, she stated, "...[in] the whole class, the fastest times, the mean and the median are very close because it [the data] was mostly clustered around them and there was no serious outliers, that is, like below them..." (Professional Development 4, lines 156–9). Here Mae saw that the main clustering of the data about the center without any values that she called "serious outliers" can have an effect on where the mean and the median are situated in relation to each other. Specifically, in this distribution she saw that the close proximity of these central measures to each other was based on the fact that most of the data was clustered around them without any "serious outliers" below them to pull them farther away from each other. By referencing the main clustering and "serious outliers" of the distribution, Mae was looking at its variability. As previously discussed, the main cluster or *modal clump* describes both the general location of the central measures and the variability of data around them (Konold et al., 2002 as cited in Makar and Confrey, 2005).

A statistician might not connect the concept of variability of the distribution directly to the relative location of the mean with regard to the median. However, Mae's connecting the variability of the distribution to the relative location of the mean and the median is reasonable for someone who is just beginning to make sense of variability using informal language. A statistician uses the relative location of the mean and the median to determine skewness of the distribution.

Skewness is the measure of the degree of asymmetry of a frequency distribution. When a distribution is symmetrical with a single mode, the mode = median = mean. When a distribution stretches to the right more than to the left, it is *right-skewed* and then the relationship tends to be mean > median > mode. The opposite is true for a *left-skewed* distribution that stretches more to the left than to the right where the relationship tends to be mean < median < mode (Aczel, 1996, p.25). Skewness can be measured by a normalized difference between mean and median.

Whereas statisticians might compare the relative location of the mean and the median to interpret the distribution's skewness, nonstatisticians, such as Mae, most likely would not. Instead, in her interpreting of variability she began to connect informally the location of the mean and the median to the variability of the distribution via its modal clustering and "serious outliers." Discussing the presence of outliers is discussing any deviations from the distribution's main pattern of clustering. Discussing the presence of outliers with regards to the main clustering, especially in a certain direction from the cluster, below them, is also seeing if there is any data that might influence the location of the central measures—the mean in relation to the median. The mean, which takes into account all of the data values, is influenced by unusual or extreme data values whereas the median is not. In contrast to the median, the mean can be pulled more closely towards the location of outliers or extreme values. Thus, outliers and extreme values can influence the closeness of the mean and the median.

Describing whether a distribution stretches more to the left or right would be a reference to its skewness. Yet, Mae's way of mentioning values that she called outliers, such as those below the cluster, suggests that she is beginning to see that aspects of the distribution,

such as deviations from the main clustering can affect the relative location of the mean and the median. In this somewhat primitive way she is developing a mindset toward viewing variability that might be foundational to viewing skewness in relation to the locations of the measures of center to each other. In analyzing Mae's interpretation of the variability of the distribution, she might be looking more than at the main clustering of data to determine the relative location of the mean and the median. The distributions' main clustering and location of what she called outliers, such as below this cluster, gives us an idea of where most of the data are situated plus where there are data values that might deviate from the modal clustering.

After a brief amount of professional development that used a middle school text, it would not be expected yet for Mae to have a full working definition of variability and its connections. In fact, it is reasonable to expect that her ideas as expressed through informal language would be the beginning of interpreting variability and its connections to other aspects of the distribution. Therefore, in an attempt to interpret variability, she made informal connections that made sense to her. This has potential benefits because informal language can be a basis with which to switch to more formal terms when (and if) appropriate (Makar and Confrey, 2005).

Mae's interpretation of variability in connection to the relative location of these measures of center might be somewhat unique. In the research reported in this study no teachers were noted to have made this connection. This is not surprising because as previously stated relating the variability of the distribution to the location of the mean and the median is not a practice of statisticians. Nonetheless, as I have attempted to point out, it emerged from the informal language Mae used in making sense of variability.

In short, it appeared that measures of center were important to Mae in interpreting variability. She chose a task for her classroom lesson that involved using measurement data based on the mean and the mode. She did this for two reasons. First, when represented graphically, measurement data would keep the mean and the mode in view, and, second, that this was important when discussing how the data was proportioned around them. In addition, Mae chose to include the use of reference lines in her classroom lesson because of measures of center. Specifically, she expected that her students would use reference lines to locate the modal clump or main cluster (Konold et al., 2002, as cited in Makar and Confrey, 2005) of data in the distribution. In doing so, she would have them locate the range of typical values for the distribution.

Mae also used this modal clump (Konold et al., 2002) or cluster and values she called outliers to discuss the relationship between these measures of center. In solving her professional development problem, Mae's interpretation was that the clustering of most of the data in the distribution including its deviations—values she called outliers—interacted with the relative location of the mean and the median. Specifically, to her the relative locations of the mean and the median to each other could be based on the data being mostly clustered around them without any of the values she called outliers below (perhaps pulling the mean toward these extreme values). This awareness points toward a mindset of seeing that the variability of the distribution via its modal clustering and outliers might interact with the relationship between the mean and the median.

Overall, Mae's use of measures of center in discussing variability might be considered foundational to meeting the requirements of Level B of the 2005 GAISE Curriculum Framework (Franklin et al., 2007). Through her use of modal clumps (Konold et al., 2002) or

clustering of most of the data, Mae began to see informally data values in relation to a central value. In this way, she seemed to value using the measures of center and the data around them to discuss variability.

In summary, Mae's use of *measures of center* gave insight into her interpretations of variability. Her choice of a task using measurement data for her class lesson on variability demonstrated that to her identifying percents of data above and below measures of center could indicate variability. In addition, determining that reference lines could support locating the modal clump (Konold et al., 2002) or cluster, indicates her seeing that identifying the modal clump (Konold et al., 2002) or cluster, is a part of interpreting variability. Finally, Mae's unconventional connection between the distributions' main clustering, the values she called outliers, and the relative location of the mean and the median indicates how she saw that the pattern of variability of the data along with its deviations might affect the relative locations of the mean and the median. These summative statements of Mae's sensemaking and the interpretations she seemed to make of variability in this process are displayed in Table 6.2. In both her professional development problem and her performance task, Mae mentioned another feature of distributions—data values she perceived as outliers. Along with them, Mae used the range to interpret variability. How she used them is discussed in the next section.

What Mae did during sensemaking with measures of center	What Mae seemed to interpret about variability
a) Chose a problem with measurement data in order to see variability	Identifying percent of data above and below measures of center indicates variability
b) Saw reference lines as helpful to identify the main cluster or modal clump (Konold et al., 2002)	Identifying modal clump (Konold et al., 2002) (measure of center and dispersion of most of the data around it) is part of seeing variability
c) Saw that the relative locations of mean/median could be based on the data's clustering and values perceived as outliers	Variability of data (and its deviations) might affect the location of measures of center

Table 6.2 A Snapshot of Mae's Sensemaking and Interpretation of Variability

### Mae Used Her Perceived Outliers and Range as a Way to Interpret Variability

Range is a key concept in understanding variability of data distributions. It is a measurement of the spread or dispersion of the distribution. Range can be expressed as a number representing the difference between this maximum (highest) value and the minimum (lowest) value of the distribution. The *Data Distributions* text used in this study defines it as “A number found by subtracting the minimum value from the maximum value. If you know the range of the data is 12 grams of sugar per serving, you know that the difference between the minimum and maximum values is 12 grams” (Lappan et al. 2006).

On the other hand, outliers are extreme observations (Aczel, 1996). Outliers include those values that are far away from the rest of where the data set is clustered. They seem to stand out in some way at either the high or low end of the distribution. In the *Data Distributions* text, Lappan et al. (2006) state that outliers are: “Unusually high or low data values in a distribution” (p. 92). While other aspects used to describe distributions—shape,

center, and variability—focus on the overall patterns in the data, outliers describe deviations from the pattern (Reading and Reid, 2006).

Outliers need special attention. They might represent an error or might be caused by special circumstances. In the latter instance, the information outliers provide might be important. For example, the head sizes collected in Mae's class might have represented a student with a much larger or smaller head than his or her classmates. As such, outliers need to be evaluated. Aczel (1996) states, "Because of the possible information content in outliers, they should be carefully scrutinized before one decides to discard them" (p. 503). Therefore, it is important to determine, particularly with measurement data, whether outliers can be traced to an error in measuring. If so, they should be disregarded.

Recall from earlier discussion on variability in the literature review, that outliers are identified by how far they are from the interquartile range, which is the middle 50% of the distribution. Specifically, outliers are determined as being three times the interquartile range  $3(IQR)$  more or less than the third or first quartile respectively. This means that the value of  $3(IQR)$  is added to or subtracted from the third and first quartiles respectively to determine the outer fences of what are considered outliers for the distribution. In addition to outliers, there are also suspected outliers that are determined when  $1.5(IQR)$  is added to or subtracted from the third and first quartiles respectively. The values determined by this method will be the inner fence of what is considered suspected outliers.

In this section, the following examples of Mae's work with values she perceived as outliers shows that based upon the statistical methods described previously to determine them, some of the values she called outliers were either outliers, suspected outliers, or close to suspected outliers. For consistency in this dissertation study, the values that Mae refers

to as outliers are labeled *perceived outliers* when they were outliers, suspected outliers, or not outliers. This is done to keep the focus on the purpose of this section, that is, to study how Mae used the concept of outliers when making sense of variability.

In other words, in spite of Mae's inexperience in naming outliers, how she dealt with and what she said about her perceived outliers would help indicate what she seemed to know about them. Also, based on having no experience with calculating outliers in this dissertation study, it is not expected that Mae would be exacting when labeling outliers. Nevertheless, how she handled her perceived outliers can give insight into what she knows about them when interpreting variability. This section discusses how Mae defined her perceived outliers and how she dealt with them in different contexts.

### **Defining Variability Includes the Concept of Outlier**

The concept of outlier was an integral part of Mae's interpreting variability. During her lesson planning session, Mae chose a problem to use in her introductory lesson on variability in data distributions. For Mae, choosing a task for this lesson, where the definition of variability would be used, was important; and the concept of outlier was a part of her definition. She stated, "[In this task] you can use the definition of what it means to be clustered together and farther apart and outliers" (Lesson Planning, lines 34–35). Here Mae's definition of variability includes the concept of outlier. Her interpretation is that the concept of outlier is considered when discussing the variability of a distribution. Based on this, Mae chose a task that asked questions addressing it. She stated, "Because they are even using those questions in there: the outliers, the clusters" (Lesson Planning, line 74). Here her choice of a task, including questions directly asking the students about outliers and clusters, stresses their importance to her in interpreting variability. Specifically, it

indicates that she wanted her students to answer them in making sense of variability.

Mae's inclusion of the concept of outlier in discussing variability was evidenced at other times. When solving professional development Problem 1.4, Comparing Student Head Measurements, Mae briefly mentioned what the values she perceived as outliers meant to her.

### Mae Defines Her Perceived Outliers

In solving Problem 1.4, Variability in Numerical Measurements, Mae spoke about her perceived outliers. This problem asked her to determine a typical head size for distributions of a boy, a girl, and a class, respectively. The problem provided Mae with three mock distributions of these head sizes. The purpose of determining the typical head size was to select a hat size for the respective individual or class. When discussing the distribution of the male student, Jalin, Mae intimated what the values she perceived as outliers might mean to her. The distribution of Jalin's head measurements is displayed in Figure 6.1.

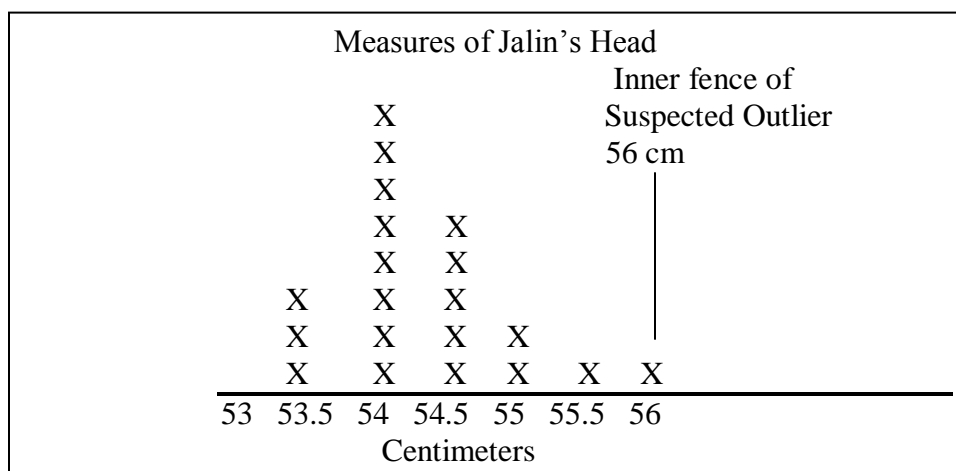


Figure 6.1 Distribution of Jalin's Head Measurements with Suspected Outlier

In her professional development notebook, Mae did not note any values she perceived as outliers for the distribution of Jalin's head measurements. However, when

responding to the other participating teacher's expressed view that 56 cm was an outlier for Jalin's head measurements, Mae stated, "Well, you mean 56 because it is not around the cluster? But even still the class got quite a few..." (Professional Development 1, lines 1420–1). Here Mae gave insight into what she might think an outlier is, that is, a data value that is not around the cluster. Mae's brief comment does not give a full view of how she saw outliers. Yet, it does give us some insight into her thinking about them.

Here Mae gave a qualitative and informal description of how she defined outlier that shed some light on her way of interpreting variability. As one might reason from the previously described method used by statisticians to quantify outliers, Mae's description of outliers is not precise enough. Nonetheless, it seems to indicate that Mae's thoughts when she stated, "because it is not around the cluster" might be headed in the right direction of seeing how outliers are determined: specifically, that they are determined in relation to the main clustering of a distribution, which could be similar to the interquartile range (IQR) that encompasses 50% of the distribution.

Based upon having no experience with calculating outliers during professional development, it would not be expected that Mae would have determined through calculation whether 56 cm was an outlier. For this distribution, when using the previously noted method for determining the inner fence of the distribution, that is  $1.5(\text{IQR}) + \text{the third quartile}$ , the value of 56 cm can be called a suspected outlier. Figure 6.1 depicts the distribution of Jalin's head measurements marked with the inner fence of 56 cm. Later, Mae further discussed what her perceived outliers meant to her in a classroom lesson when she was introducing her students to variability in data distributions. Chapter 7 discusses how she defined her perceived outliers in her lesson in more detail.

Mae very informally began to show the importance and meaning of outliers to her. Based on her informal treatment of outliers, Mae met the expectations of Level A learners of the 2005 GAISE Curriculum Framework (Franklin et al., 2007) that required only that type of exposure to them. In addition, the very brief and informal meaning Mae gave to her perceived outlier can give a hint of her making sense of the conceptual underpinning of outliers—that they are a certain distance from the interquartile range (IQR) where most of the data are located. Besides defining what her perceived outliers meant, Mae also exhibited how she would handle them. In particular, Mae demonstrated when she would include her perceived outliers in the analysis of the distribution and when she would not. Context seemed to play a part in how she handled her perceived outliers. This is discussed in the following section.

### **Context Involved in Mae’s Treatment of Her Perceived Outliers**

Through the following examples it is conjectured that context might have influenced Mae’s treatment of her perceived outliers when making sense of variability. The first example demonstrates a context when Mae included values she perceived as outliers in her sensemaking of variability. The last example provides a context when she did not.

**Including Her Perceived Outliers.** Mae showed how she treated her perceived outliers when discussing the mock students’ head measurements in Problem 1.4. In this problem she perceived that there were outliers and she included them as part of her analysis. When referencing the typical head size for the class of students, Mae stated, “I wrote 54[cm] but there were more outliers” (Professional Development 1, lines 1529–30). Here Mae included her perceived outliers in the analysis of typical head sizes for the

distribution. Shortly thereafter, she elaborated on what these perceived outliers were to her.

Later on in solving the same professional development problem Mae gave more of a description of these values that she perceived as outliers. She stated, “It could be that those extra people can’t wear size 56 cm and you cannot exclude them just because most people have between 54 and 56” (Professional Development 1, lines 1539–40). Here, perhaps unintentionally, Mae misstated 56 cm as the typical size for what she meant to be 54 cm. (This is assumed because she had written in her professional development notebook 54 cm to indicate where most of the data are located for the class. She also confirmed 54 cm in another statement made in professional development. The following discussion is based upon assuming Mae meant 54 cm). Here Mae addressed how she saw these perceived outliers and as a result how she would treat them. She viewed them as people who do not have the typical head size of 54 cm. Mae also decided not to exclude these people just because they were not part of the main cluster of head sizes.

In this problem, Mae’s perceived outliers were suspected outliers. For the mock distribution of the class’s head measurements, an outlier would have to be a value of 60 cm or greater, and a suspected outlier would have to be a value of 57.75 to 59.9 cm. As a result, two values could be identified statistically as suspected outliers, 58 cm and 59 cm, and a third value, 57.5 cm, was very close to being a suspected outlier. Since Mae did not specify the values she perceived as outliers, it is hard to tell which exact ones she referred to in discussing “those extra people who can’t wear size 56 cm (assumed to mean 54 cm).”

It seems as though Mae accepted her perceived outliers as valid measurements of students’ head sizes. It is conjectured that Mae might have done this because she expected

more variability in the students' head measurements. She might have thought there would be some students whose head sizes were not going to fit the typical head size of the class. The speculation that Mae expected more variability for the class's distribution was possibly confirmed during her postlesson discussion. During that discussion Mae expressed expecting a range of values for the typical students based on the 30 different data values representing the 30 different student head sizes. This is discussed further in Chapter 7.

As a result of the variability Mae expected in the class's head measurements, she seemed not to question the presence of any of her perceived outliers in the distribution. She expressed no doubts about them being valid members of this distribution of students. Once Mae decided that her perceived outliers were students who just could not wear the typical hat size of 54 cm, she included them in her analysis. This is despite the fact that she considered them outliers. It is conjectured that based on the context Mae realized that she could not exclude these students to determine a hat size for the whole class.

Also, in the context of student head measurements, Mae did not seem to attribute her perceived outliers to measurement error. This might also be based on the variability she expected, that is, 30 different head measurements for the class. In addition, this expectation might have obscured her thinking that any of the head sizes might result from errors. Finally, based on the fact that the distribution of head sizes was a mock one, she did not see the actual head measuring to pick up any errors.

Similar to Mae, a statistician might have expected certain variability in the head measurements of 30 different students. A statistician also might have agreed with Mae when she included "those extra people" who could not wear size 54 cm. Yet, the statistician might have done so for different reasons. He or she might not have considered

those extra people outliers—one, because they (two of the three) were suspected outliers, but also because of his or her knowledge and experience with the variability of data in this context.

Mae was not expected to be where a statistician would be in terms of the knowledge of variability (inclusive of its deviations—outliers) in this context. Based upon her experiences in this dissertation study, Mae is where she was expected to be. Pursuant to the 2005 GAISE Curriculum Framework (Franklin et al., 2007), she had informal experiences with the concept outliers as a Level A learner. She informally used the concept of outlier to mean deviations from the pattern in the variability, and then she decided whether to include the values she perceived as outliers in the analysis. Her work with making sense of variability inclusive of its deviations might be considered a foundation upon which to build.

In summary of this part of Mae's treatment of her perceived outliers, it is speculated that the inclusion of her perceived outliers for the class might be related to the context. In this professional development problem the context was analyzing a distribution of students' head measurements in order to find a hat size for the whole class. It might be that Mae expected greater variability in the data on 30 different students with 30 different head sizes. In addition, in her inexperience with outliers she might have expected there to be some students that would be considered outliers. At the same time, because the context required Mae to determine one hat size to fit the class she could not exclude her perceived outliers.

Mae also did not address possible measurement error. Perhaps this is based on the variability that she expected in the class's head sizes obscuring the possibility of an error; or, because it was a mock distribution and she did not see the measuring take place to pick

up errors. One way to possibly test this conjecture is to see how Mae handled her perceived outliers in another context. The next section discusses this other context.

**Excluding Her Perceived Outlier.** When solving professional development Problem 3.2, Comparing Reaction Times, Mae was asked to comment on students' reaction times to a computerized video game. In analyzing Henry's times, Mae made an entry in her professional development notebook. She noted the consistency of Henry's times and an outlier. Henry's reactions times were 1.15, 1.25, 1.34, 1.47 and 2.48 seconds. Regarding Henry's times, Mae wrote, "Henry who was pretty consistent, except for 2.48 [seconds], an outlier" (Professional Development Notebook, p. 11). Here Mae described the pattern of Henry's video game reaction times as pretty consistent and the outlier as an exception to these times. Her comments intimate that she is seeing his time of 2.48 seconds as a deviation from his pretty consistent pattern of speed in reacting to the game.

When answering a question about ranges and consistency, Mae gave some insight into how she might have seen the outlier in Henry's times. In Mae's answer to question B3 of the same Problem, 3.2, that asked about the usefulness of the range in comparing consistency of times, Mae stated, "No, it [the range] doesn't because they could have one bad reaction time causing them to have a large outlier and the other values are small" (Professional Development Notebook, p. 11). Here the meaning Mae seems to give outlier in this context is a "bad reaction time." Specifically, in this instance it is a reaction time that is large compared to the other reaction times that are small.

It could be argued that the context of computer reaction times would have made it easier for Mae to exclude an outlier. It was not like excluding students when determining a hat size for the whole class. Yet, for Mae it might have been reasonable to expect a student

to have some bad reaction times in a computer game. And based on this expectation, it seemed reasonable to disregard any times when they were an exception and fell outside the overall pattern of time, such as in Henry's times.

It also could be argued that Mae excluded 2.48 seconds not because of the context, but rather because it is an outlier. In this problem, 2.48 seconds is an outlier of Henry's reaction times. The outer fence to determine outliers of the distribution of his times was 2.13 seconds. This is in contrast to the data values in the context of students' head measurements where Mae included them when they were suspected outliers. There might be some truth to the fact that she excluded 2.48 cm because it was an outlier. However, when looking at the common approach used by Hammerman's and Rubin's (2004) teachers to discount extreme values when the context did not prove them helpful in comparing the distributions, it could be possible that Mae used the same reasoning.

When comparing the context of the number of weekly hours that students from two different locations spent on homework, Hammerman's and Rubin's (2004) teachers discounted the extreme values when they did not represent something about the typical student. The teachers stated, "The top [amount of weekly hours doing homework] is a lifestyle no matter where they [the students] live" (p. 29). Hammerman and Rubin (2004) conjectured that this approach was an extension of the teachers' strategy of "disregarding outliers" (p. 30). (Note that the values that Hammerman's and Rubin's (2004) teachers' were disregarding were extreme values and not outliers. Nonetheless, based on the informal way Mae identified her perceived outliers in this dissertation study, their findings seemed applicable.)

For Hammerman's and Rubin's (2004) teachers the purpose of the analysis seemed to be a factor in deciding to disregard outliers. In their study, when the purpose was based on comparing the typical amount of weekly hours that students spend doing homework between two locations, the extreme amount of homework did not seem to be relevant to the analysis. Likewise, when applying this reasoning to Mae's treatment of the outlier in Henry's times, the outlier represented a deviation from the overall pattern of his times. Therefore, in this context, 2.48 centimeters did not seem relevant to discussing the consistency of his times and, thus, it could be excluded.

As mentioned earlier, the 2005 GAISE Curriculum Framework does not expect the learner to experience formal treatment of outliers until Level C (Franklin et al., 2007). Yet, they recommend informal experiences with them at Levels A and B. At Level A and B, the framework suggests that an understanding of error versus natural variability will help students interpret whether an outlier is a legitimate data value that is unusual or whether an outlier is due to a recording error (p. 33). As discussed in this section on outliers, Mae had no formal experience with identifying outliers in this dissertation study. As a result, she did not calculate whether the data values she perceived as outliers were, in fact, outliers. It was found that sometimes her perceived outliers were either outliers, suspected outliers, or close to suspected outliers. Mae was not expected to be at Level C of the 2005 GAISE Curriculum Framework (Franklin et al., 2007) in this dissertation study. Therefore, informally addressing outliers placed her where she was expected to be, that is, at Levels A or B.

In short, in spite of Mae's lack of experience in identifying outliers, Mae's treatment of her perceived outliers helped to indicate what she might know about them

when making sense of variability. This entire section on outliers focused on how Mae informally defined, referred to, and treated her perceived outliers. First, it was shown that the concept of outlier along with clusters were important to Mae to include in her task that introduced variability to her class. Second, it was demonstrated that she informally identified her perceived outliers as data values that were not around these clusters. Third, through examples it was conjectured that context influenced how she treated her perceived outliers, that is, whether she excluded or included them in her analysis. In conjunction with her perceived outliers, Mae also mentioned range when making sense of variability. This is discussed in the following.

### **Mae's Perceived Outliers and Range**

As previously mentioned in the discussion on Mae's excluding her perceived outliers, she also referred to them when discussing range. Range is an integral part of variability. It is the measure of the distance from the minimum data value to the maximum data value. Expressed as a number it is the measure of the spread of the distribution. When solving professional development Problem 3.2, Comparing Reaction Times, Mae accurately calculated the ranges for all of the students' reaction times. She also discussed consistency in connection with the range, and when she did, she expressed that the range was not useful when the distribution contained a value she perceived as an outlier.

Mae's response to question B3 of Problem 3.2 indicates what she noticed about the usefulness of the range when comparing the consistency of times. The question asked, "Does comparing ranges of reaction times help you decide if one student is more consistent than another student?" (Lappan et al., 2006, p. 58). To this Mae responded, "No, it [the range] doesn't because they could have one bad reaction time causing them to have a large

outlier and the other values are small” (Professional Development Notebook, p. 11). Here in her response, Mae is showing us how she sees that the range is not helpful when comparing consistency because it might be influenced by a value she called an outlier, that does not represent the other values of the distribution. (As mentioned previously, 2.48 seconds is considered an outlier.) Mae seems to know that 2.48 seconds falls out of the typical pattern of the reaction times and, therefore, in itself it does not tell us much about the pattern or variability of most of the distribution.

According to the 2005 GAISE Curriculum Framework (Franklin et al., 2007), Mae’s work was foundational. Her work with range aligned with Level A. At Level A, learners are introduced to range as a measure of spread in numerical data. Mae’s comparing of the ranges in these distributions of computer reaction times placed her on this level. She was able to calculate the range and, thereby, she began to quantify how much variability there is in a distribution of numerical data. However, since the range is only one quantity that measures the degree of variability, more analysis can be involved in discussing variability. To some extent Mae seemed to know this when she looked more closely at how the data was spread to discuss its consistency.

In this way, Mae’s work is significant. She knew more than just how to compute the range. Specifically, she critiqued the usefulness of the range when she perceived an outlier in the distribution. She also knew to look at the pattern in the data to determine its consistency. Specifically, Mae knew to look back into the data for a consistent grouping of values when determining the consistency of these times. Thus, Mae knew not only that the range was not useful when she perceived an outlier, but also how to determine consistency without completely relying on range to indicate variability.

This example seems to demonstrate how Mae went beyond a rote use of a measure to a more critical use of it. In the context of video game reaction times, Mae considered the presence of outliers in her statistical reasoning. Makar and Confrey (2005) discuss the importance of this:

Variation encompasses more than just a measure, although measuring variation is an important component in data analysis. In considering variation, one must consider not just what it is (its definition or formula), or how to use it as a tool (related procedures), but also why it is useful within a context (purpose). (p. 28)

In line with this statement, knowing why range would not be useful when an outlier is present is also important. In short, Mae determined the usefulness of the range in discussing consistency. She took a critical view of using range. She did not use it deterministically in that she looked to the purpose of determining consistency in video game reaction times to evaluate its use. When an outlier was present to her, she did not use the range because it proved not helpful in and of itself in discussing variability.

In summary, Mae's use of her *perceived outliers* in making sense of variability gave insight into her interpretation of variability. Her definition of variability demonstrated her interpretation that the overall pattern in a distribution's variability has deviations. Her definition of her perceived outliers also indicated her interpretation of the conceptual underpinning of calculating outliers. In addition, her seeing that context influences the inclusion and exclusion of her perceived outliers points toward her seeing that variability can be based on natural variation. These summative statements of Mae's sensemaking and the interpretations she seemed to make of variability in this process are displayed in Table 6.3.

<b>What Mae did during sensemaking with her perceived outliers</b>	<b>What Mae seemed to interpret about variability</b>
a) Defined her perceived outliers  b) Saw that context influences including and excluding her perceived outliers	Overall pattern in variability has deviations  Conceptual underpinning of calculating outliers  Variability can be based on natural variation

*Table 6.3 A Snapshot of Mae's Sensemaking and Interpretation of Variability*

In summary, Mae's use of *range* in making sense of variability gave insights into her interpretation of variability. Basically her work showed that she knew how to calculate range in order to measure variability. In addition, Mae interpreted the range as not being a useful measure when an outlier is present. These summative statements of Mae's sensemaking and the interpretations of variability she seemed to make in this process are displayed in Table 6.4. Another way Mae interpreted variability was in discussing the shape of a distribution. This is discussed in the next section.

<b>What Mae did during sensemaking with range</b>	<b>What Mae seemed to interpret about variability</b>
a) Calculated range  b) Determined the usefulness of range when to her an outlier is present	How to calculate range to measure variability  Range is not a useful measure to describe variability when an outlier is present

*Table 6.4 A Snapshot of Mae's Sensemaking and Interpretation of Variability*

### Mae Used Shape as a Way to Interpret Variability

Shape is a feature of the distribution that answers the question, “What do I see?”

(Moore, 2000, p. 12). A distribution can be seen as symmetrical or skewed. Moore (2000) states:

A distribution is **symmetric** if the right and left sides of the histogram are approximately mirror images of each other. A distribution is **skewed to the right** if the right side of the histogram (containing the half of the observations with larger values) extends much farther out than the left side. It is **skewed to the left** if the left side of the histogram extends much farther out than the right side (p. 12, emphasis in original text).

Moore (2000) further states that in mathematics symmetry means the two sides of a histogram are exact mirror images of each other. Histograms can be called approximately symmetric as an overall description because data are almost never symmetric (p. 12).

On a gross level, a distribution’s shape is closely related to its measures of center.

When a distribution is symmetrical with a single mode, the mode = median = mean. When a distribution stretches to the right more than to the left, it is *right-skewed* and the relationship between measures of center tends to be mean > median > mode. The opposite is true for a *left-skewed* distribution that stretches more to the left than to the right where the relationship between the measures of center tends to be mean < median < mode (Aczel, 1996, p.25).

Mae attended to shape in describing variability. Mae used *shape of the whole distribution* and *shape of the main cluster* of the distribution in making sense of variability. When solving her professional development problems and her performance task, she used informal or nonstandard language to do so.

## **Shape of Whole Distribution**

Mae addressed the shape of the whole distribution when solving her professional development problem and her performance task. She used informal language to discuss asymmetrical shapes.

In discussing variability during the third professional development session, Mae described the shape of the whole distribution. She did this when the problem explicitly asked her to. She discussed how the shape was related to measures of center. Professional development Problem 2.4 required Mae to reflect on the relationship between the whole distribution's shape and the relative locations of the mean and the median. As implied in its title, this problem aimed at having Mae make connections between the measures of center and shape. Specifically, it required her to connect the locations of the mean and the median to the shape of the distribution. It accomplished this by having Mae sort data distributions according to the relative locations of the mean and median. Then, it asked her to describe how their locations appear to be influenced by the shape of the distribution (Lappan et al., 2006, p. 43).

In solving this problem, Mae described asymmetrically-shaped distributions. In the following instance when solving professional development Problem 2.4C, Mae referred to how the data was spread along the scale of the distribution. She did this by referencing how the low and high values were dispersed in the distribution. She stated “and in [distribution] two, there is a lot of gaps and a lot of single values that are low where most of the values are at the higher end and it skews the information” (Professional Development 3, lines 1094–5). Here Mae is attending to the asymmetry of the distribution. When asked what exactly skew means she replied “The gaps and the unevenness about the amount that are on each value.” In her explanation, Mae used informal language and did not define skewness in the statistical sense of

the word. Yet, her description of the distribution leaned toward her seeing asymmetry and possible skewness.

It is not surprising that Mae did not use the term *skewness* as a statistician would. As previously stated, in this dissertation study Mae had a brief amount of professional development that used a middle school text that did not mention skewness. Therefore, at that time it would not be expected that she would have a full working definition of the skewness of a distribution. However, Mae's sensemaking about how the distribution was spread out lends itself to viewing the distribution as asymmetrical or possibly extending to the left as somewhat skewed. The paragraph below explains how this could be so.

A distribution is skewed to the left if the left side extends much farther out than the right side (Moore, 2000, p. 12). Through Mae's use of informal language, it might be reasonable to see that she described a distribution that is asymmetrical and extended out or more variable on the left. By describing lots of values that are low and most of the data on the higher end, it is possible that Mae's description points toward her beginning to see asymmetry and possible skewness in a distribution. In another professional development problem, Mae also discussed the shape of clusters in interpreting variability.

### **Shape of Clusters**

Professional development Problem 1.3, Variability in Numerical Measurements, was another opportunity for Mae to use shape in discussing variability. In this problem she described the shape of the main cluster of the distribution.

Early on in the first professional development session Mae discussed the location of the clusters. She stated her understanding of cluster was the location of most of the data. In Problem 1.3, Mae described the location of most of the students' head measurements. In

identifying the cluster of head measurements for the female student Sarah, Mae identified 56 centimeters. To her this was because “it seems everything is just around 56[cm]” (Professional Development 1, line 1451). This is in contrast to the clustering she named for the distribution of the male student, Jalin’s, head measurements. Mae chose a range of values for Jalin as opposed to the one value she chose for Sarah. She described the clustering of Jalin’s head sizes as a range of values—between 53.5 cm and 54.5 cm. She stated this was because “it is more evenly distributed as to how much is most for him” (Professional Development 1, lines 1449–50). Here Mae’s description of Jalin’s cluster alluded to its shape. She described both the parameters of the cluster, from 53.5 to 54.5 cm, along with how it is spread, that is, evenly. From her description, you could get the sense of where most of the data was located on the scale, its spread, and how it was generally shaped.

In this problem, Mae described the shape of the main cluster where most of the data was located. She gave both the location on the scale of this main cluster and a description of how it was evenly distributed. In these two examples, Mae gave a description of the clustering and to her how the particular pattern of its shape, or lack thereof, affected the typical value. Because most of Jalin’s head measurements were more evenly distributed, Mae provided a range of values to describe his head measurement. This was in contrast to her description for the female student, Sarah. To Mae, most of Sarah’s head measurements pretty much fell around 56 cm with no remarkable pattern of distribution. This led Mae to give one measurement to describe the cluster.

In the distribution of Sarah’s head measurements, 56 cm was the mode. The mode obviously stood out among the other data values and represented 60% of them. In contrast, she gave a range of measurements for Jalin’s head because most of his data was evenly distributed

in between the two data values. In the distribution of Jalin's head measurements, 80% of the values fell between 53.5 and 54.5 cm with 40% at the mode of 54 cm. Based on her solution to this problem, it seems that the shape of the cluster influenced the values that Mae chose as typical.

Mae's discussion of shape of the main cluster of a distribution in interpreting variability is significant for a number of reasons. It demonstrated her knowing variability through a distribution-view of the data. This knowledge was important. It was evidence that Mae was not just seeing measure of center (typical value) as a calculation from individual points, but rather as a characteristic of the distribution (Bakker et al., 2004 as cited in Makar and Confrey, 2005). In particular, Mae discussed how the characteristic of shape interacted with her distributional conception of typical value. When the main cluster was evenly distributed, Mae chose a range of values to describe the typical value. In this way, she exhibited a tolerance for variability in the measures of center based on the shape of the main cluster.

In short, in Mae's use of shape when interpreting variability, she described the shape of the whole distribution as well as the shape of its main cluster. Mae's description of the shape of the whole distribution was interpreted as asymmetrical. When Mae discussed an asymmetrical distribution, she described how the data were dispersed at the low end of the distribution in contrast to how they were dispersed at the high end of the distribution. She did not use the term *skewed* in the statistical sense of the word, and she was not expected to, based on the brevity and content of the professional development offered her in this dissertation research. Yet, her description of the asymmetrical distribution pointed toward her seeing the variability in a distribution that was possibly somewhat left skewed.

Further to describing the shape of the whole distribution, Mae also described the shape of the main cluster. This is significant because it influenced how she described the typical value. When the data fell pretty much around a number with no distinctive shape or pattern of variability, Mae gave a single value to describe this main cluster. Conversely, when the shape of the main cluster had a specific pattern, such as more evenly distributed, Mae gave a range of values for typical; thus, exhibiting her tolerance for variability in the typical values.

Overall, Mae's discussions of shape in interpreting variability were important. In general her work corroborated the research findings of Canada (2004) and Makar and Confrey (2005). All of these researchers found that their teachers used nonstandard or informal language to discuss shape. Along with this alignment with research, Mae's work also met the requirements of the 2005 GAISE Curriculum Framework (Franklin et al., 2007) and other aspects of statistics education research.

Level A of the framework states that looking at the distribution's variability—its clusters and gaps—helps the learner identify its shape. Mae did this when she discussed the gaps and most of the data (main cluster or modal clump, Konold et al., 2002) in the distributions. When the 2005 GAISE Curriculum Framework relates the distribution's clusters and gaps to naming its shape, it is identifying a connection between variability and shape (Franklin et al., 2007). This connection was found in research as well (Makar and Confrey, 2005).

Makar and Confrey (2005) found that in teachers' informal "variation talk" the concept of variation and distribution are closely related. Their preservice teachers' informal variation talk included phrases to express *spread*—*clustered*, *clumped*, *grouped*, *bunched*, *gathered*, *spread out*, *evenly distributed*, *scattered*, and *dispersed*—that highlighted attention to the more spatial aspects of the distribution. To Makar and Confrey (2005) these terms took on a

meaning that implied attention to variation more as a characteristic of shape than as a measure. They also stated that this is in contrast to concepts of variation in conventional statistical language that are articulated by terms like *range* or *standard deviation*, both of which are measures (pp. 47–48).

In summary, Mae’s use of *shape* in making sense of variability gave insight into her interpretation of variability. Her use of shape to describe the asymmetry of the distribution demonstrates her interpreting that the variability in data can affect asymmetry (possible skewness). In addition, Mae’s description of the shape of the cluster points toward her seeing how the cluster’s shape might indicate how the typical value varies. These summative statements of Mae’s sensemaking and the interpretations of variability she seemed to make in this process are displayed in Table 6.5. In this study, Mae also stated what variability meant to her. She then used her definition to guide her in describing variability. This is discussed in the next section.

What Mae did during sensemaking with shape	What Mae seemed to interpret about variability
a) Discussed asymmetry b) Discussed shape of main cluster	Shape describes asymmetry (possible skewness)  Variability in data affects asymmetry (possible skewness)  Shape of main cluster might indicate how the typical value varies

*Table 6.5 A Snapshot of Mae’s Sensemaking and Interpretation of Variability*

### **Mae Constructed Definitions and Descriptions of Variability**

Throughout her professional development and performance tasks Mae discussed the meaning of variability. She defined what variability was to her and described it in her problem solving. There were instances when Mae *defined variability*, instances when she both *defined*

*and described it*, and other instances when she *described* it. Here *defining* variability refers to the utterances Mae used to explain what variability is to her; *describing* refers to how Mae depicted or described the variability in certain distributions. Lastly, *defining and describing variability* refers to the instances when she did both.

### **Defines Variability**

During her lesson planning Mae expressed what variability was to her. She used her definition as a lens to choose the type of problem for her classroom lesson. She stated:

Well this one (categorical problem) is hard for me to understand how the variability works. Whereas, opposed to when they measure their heads, you could *use the definition of what it means to be clustered together and farther apart and outliers*. This one (measurement problem) would be easier. (Lesson Planning, lines 29–30; emphasis added)

Here Mae defined variability in terms of what it means to be clustered together and farther apart with outliers. Her definition included the pattern of how the data varied and deviations from this pattern (Reading and Reid, 2006). In a very basic way, her definition could be pointing toward variability as a measure of the spread of the data values. This might be a beginning from which to grow in her defining variability. Besides what it means to be clustered together and farther apart, there is room in her definition to include where the data values lie in relation to a central measure.

For example, a statistician might see that variability is how varied the data values are in relation to their mean (Aczel, 1996, p. 17). Based on her experiences in this dissertation study, Mae was not expected to have this knowledge of variability. Yet, she continued to make sense of variability. How she further defined and described it might be considered a foundation upon which to build. This is discussed in the next section.

### **Defines/Describes Variability**

When solving her professional development problem, Mae defined and described variability. She used her definition of variability to guide her in describing the variability of the distribution in the problem. On two instances she stated variability involved data clustering around certain values and those not around them.

The first time Mae discussed the meaning of variability was in inquiring about question eight of Problem 1.3, Variability in Numerical Measurements. This question asked Mae to describe the variability of the distributions of student's head sizes. After she read the question she inquired, "What are they asking me to do? I want to just describe how variability is used in the class we are looking at, or just in general say how it shows variability because there is a cluster around a certain numeric and then there is a few outliers and not everyone had that" (Professional Development 1, lines 1325–7). Here Mae added to what she defined in her lesson planning. In planning her lesson, she stated that variability involves being clustered together and spread out and outliers. Here she included in her definition that the clustering is around a certain number. She also repeated that there were values she perceived as outliers that were head sizes that not all students had. Because she included "around a certain numeric" when describing and defining distributions, her sensemaking of variability might be considered a basis from which an understanding of measuring variability, such as in relation to the mean, could be developed.

In the second instance, Mae continued to describe variability. The same problem asked her to describe the variability of the distribution of student head sizes. She used her definition of variability and stated, "Because we talked about what it was before, that is, how widely spread or how closely clustered, I talked about how they were clustered around certain

numbers. I said that Sara's variability was clustered around 56, so things around there; and Jalin's was around 54; and then the class was also around 54, but you had those couple of people that were not close to that number" (Professional Development 1, lines 1572–74; 1578–80). Here Mae describes variability in the distribution of students' head measurements. She added to her description a specific data value around which the data are clustered.

These two examples of Mae's utterances show how she included in her defining and describing variability the clustering around a certain number. In this way, how she defined and described variability might be considered a foundation upon which to build. This might include defining variability as a measure of how the data values are spread from a certain value, for example, the mean.

It is important to note that Mae's work here is similar to the modal clumps in Konold et al.'s (2002) research as described in Makar and Confrey (2005). Similar to Mae, the teachers in Makar's and Confrey's (2005) study focused on the middle portion of the distribution when comparing data distributions: modal clump (Konold et al., 2002). In essence, Mae's description of the numeric around which most of the data are clustered and located seems to encompass the center of the distribution. By discussing the center of the distribution via its major cluster, Mae's utterances showed how she described the variability of the data. Besides clusters, Mae also included a range of data values inclusive of gaps in describing variability. This is discussed in the next section in which Mae describes variability.

### **Describes Variability**

Mae also described the variability of data distributions when solving her Performance Task 2. This task required Mae to choose, among three different graphs of car stopping

distances, the one that showed more variability than the others. These graphs are depicted in Figure 6.2.

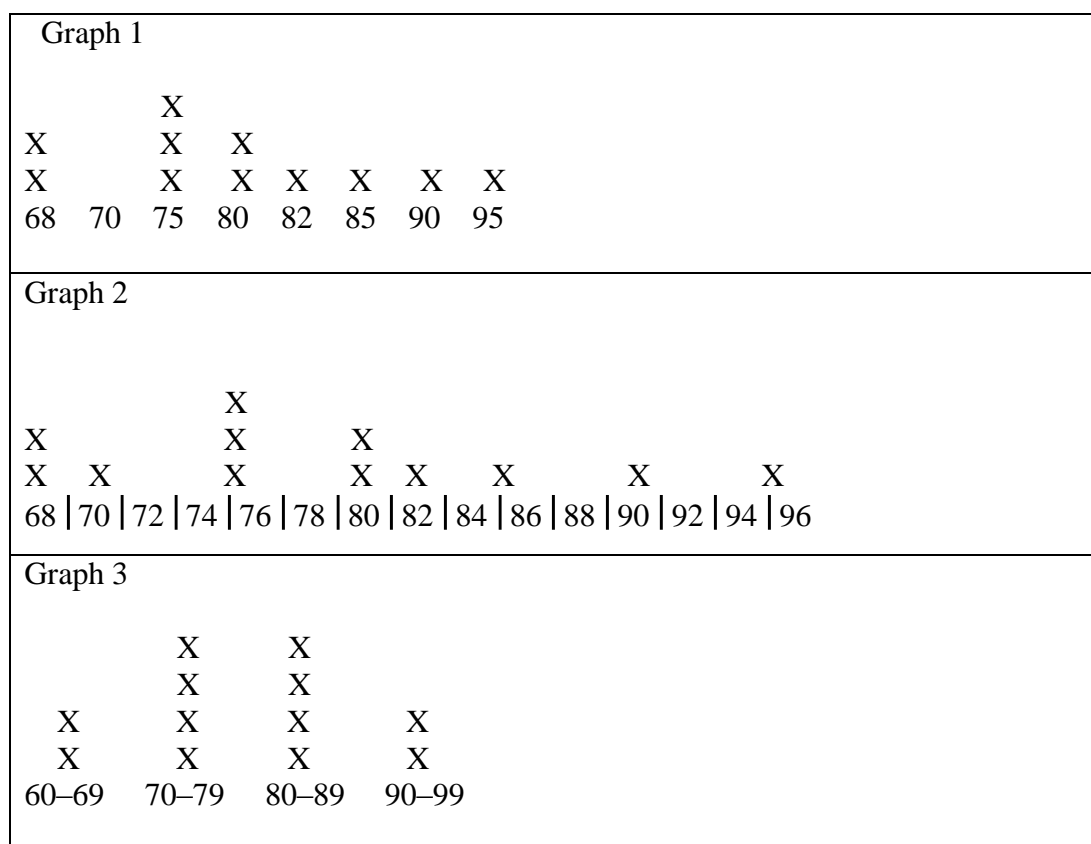


Figure 6.2 Graphs From Performance Task 2

In choosing Graph 2 as the graph that shows more variability, Mae articulated its variability by describing its clusters including its gaps and her perceived outliers. She stated, “because it has those outliers of 90 and 95, and then it has most of the data clustered between 75 and 85; even though there is gaps, most of it is in that range” (Performance Task 2, lines 109–111). Here in her description Mae tells how the data are spread out on the scale. She gives more specifics, such as the location of the range of where the data are mostly clustered in the distribution. Also included in her description are those data values that she calls outliers, and are not parts of the cluster. Finally, gaps were in her description even though to her they did not affect the main clustering of data that she had described. Thus, Mae’s utterances on

variability involved the location of where most of the data are clustered (despite gaps) on the scale and the location of values that are not included in this main grouping.

Regarding the concept map depicted in Chapter 2, Mae's defining and describing variability might be considered a foundation upon which to build. This map illustrated that characterizing variability includes: clusters, gaps and measures, such as standard deviation, Mean Absolute Deviation, and outliers. Mae's description of variability, including the range of values where the data mostly clustered inclusive of its gaps, to some extent aligns with how the concept map depicts characterizing variability. Mae did not mention the measures of standard deviation, the Mean Absolute Deviation, or range. In answering the question, she did not mention the range possibly because the range was detectable on both Graph 1 and Graph 2. Mae also did not mention standard deviation or the Mean Absolute Deviation most likely because she did not have experiences with these measures in this dissertation study.

In short, in Mae's constructing definitions and descriptions of variability, she was consistent. She defined variability as how far apart and how close the data values are inclusive of data that was not close to the cluster. She described variability based upon the distribution's range of data values where most of the data clustered, including its gaps, and the values she perceived as outliers. Her definition and description did not include the formal measure of variability such as the range and standard deviation. Nonetheless, through her use of informal language her definitions and descriptions of variability might be viewed as a foundation upon which to build the formal knowledge of measuring variability (Makar and Confrey, 2005).

In summary, Mae's *defining and describing variability* gave insight into her interpretation of it. Through her informal definition of variability she seemed to have made basic sense of what variability involves, that is, how far apart and close together the data

values are (as indicating what is meant when calculating the range). She also had a sense of describing variability to include the cluster of data around a certain number. This might be considered a basic foundation upon which to build, for example, on how the data values are spread from the mean when measuring variability. Also through her definitions and descriptions of variability, Mae interpreted clusters, gaps, and the values she perceived as outliers as parts of discussing the variability of a distribution. These summative statements of Mae's sensemaking and the interpretations of variability she seemed to make in this process are displayed in Table 6.6.

<b>What Mae did during sensemaking when defining and describing variability</b>	<b>What Mae seemed to interpret about variability</b>
a) Defined variability	Defining variability involves how far apart and close together the data are
b) Described variability	Describing variability involves a modal cluster around a certain number
c) Defined and described variability	Clusters, gaps, and outliers are a part of describing variability

*Table 6.6 A Snapshot of Mae's Sensemaking and Interpretation of Variability*

This chapter summarized Mae's sensemaking practice of *interpreting*. Through the lens of her sensemaking notices, how she seemed to interpret variability was discussed. Chapter 7 focuses on Mae's sensemaking of variability in the third part of her sensemaking practices—*lesson implementation*.

## Chapter 7

### **Mae's Sensemaking of Variability Through Her Lesson Implementation**

This research employed a bidirectional approach to studying teacher content knowledge of variability in data distributions. In accordance with this bidirectional approach to studying teaching, Chapter 6 focused on the one view of knowledge that Mae exhibited prior to teaching, that is in her professional development, lesson planning, and performance task work. In turn, this chapter focuses on the other view of knowledge that is exhibited by Mae when teaching her lesson. It was hoped that together these views would give a more complete snapshot of Mae's content knowledge of variability in data distributions. It was anticipated that what she exhibited knowing prior to teaching would help guide the analysis of the knowledge she would exhibit while teaching. It was also expected that Mae would use informal language as she had done previously. This is based on her experiences in this dissertation study and the fact that Mae was teaching this concept to her class of middle school students.

As such, this chapter reports findings to the third research question: What knowledge of variability in data distributions does this teacher implement when teaching this content to her students? Here another view of Mae's knowledge of variability is presented. It is a snapshot of how she represented variability to her students during a classroom lesson that she constructed for this research. Her lesson was part of the planning that was referenced in Chapter 6. The *sensemaking notices* that emerged from her sensemaking prior to teaching—*partitioning, measures of center, range and values she perceived as outliers, shape, and defining and describing variability*—became the lenses for the analysis of sensemaking during her lesson implementation. Mae's introductory lesson on variability does not capture all of

Mae's knowledge. Through its implementation, however, it gives another perspective on how Mae makes sense of variability.

This being said, the analysis of the classroom data began with a curiosity about how Mae would translate the knowledge she exhibited in lesson planning, professional development, and her performance tasks onto a lesson for her students. It was found that Mae drew upon some but not all of the knowledge she had developed prior to teaching. The *sensemaking notices* that surfaced during her lesson included: *measures of center*, *values she perceived as outliers*, and *range*. Her use of these was particularly interesting because unlike the problems she had worked on prior to teaching, she and her students worked with a previously unknown set of data, a data set that was generated during the lesson implementation from the students in her classroom.

Research on teaching statistics and reports, such as the 2005 GAISE Curriculum Framework (Franklin et al., 2007) encourage the use of data that emerges from the classroom. Collecting data relevant to the students with whom you are working can help the learner engage more in exploring the data. At the same time, it can create a greater challenge to the teacher who guides the learner through the analysis. In the instance of Mae's lesson, she walked into the classroom with her recent experiences with variability using a middle school text, a lesson with guiding questions, and a yet to be revealed distribution of data values.

Based on previous findings, it might be reasonable to expect that some different information might be revealed in this analysis about Mae's knowledge. In contrast to learning in professional development with a colleague, or her individual response to a task, implementing a lesson involves actively engaging with students' ideas. In addition to the actual student data Mae and her students collected during class (measurements of their head

sizes), she would also answer their questions, challenge their thinking, and introduce vocabulary to them. This is especially true for an introductory lesson such as the one Mae chose to try out with her students. It was anticipated that interacting with students could shed additional light on her knowledge of variability. Additional insight was found also regarding Mae's knowledge in the postlesson discussion. Specifically, these data were analyzed for anything regarding variability that did not surface during her lesson and/or that confirmed or refuted what happened in her lesson. What follows is a brief overview of the lesson, and then a catalog of the lesson. Finally, several segments of the lesson and two from the postlesson interview illustrating my findings are presented with a brief commentary on each.

In summary, the aim of Mae's lesson was to help students understand variability in data distributions. At the beginning of her lesson, she wrote the following two questions on the board: (a) What is variability? (b) How do we decide what hat size we should get for our homeroom? Mae's leading question was: Can we get one same hat size for everyone in the class, or for each of the two chosen students? To this end, she had her students measure each others' and two individual students' head sizes. The collected data was represented in respective distributions and analyzed separately. Before Mae began the analysis of the data with her students, she discussed how variability might help them interpret their data. At the end of her lesson she reflected with her students how variability helped them determine (or not) a typical hat size.

To get a better view of the components of Mae's lesson, here is a brief catalog of its highlights. About one half of Mae's 40-minute lesson was spent on collecting data and recording it. Her lesson began with the students measuring each others' heads and writing their head measurement in their notebooks. Afterwards for approximately seven minutes, Mae

conducted an introduction that included discussing the terms *variability*, *mode*, *cluster*, and *range*. During this discussion Mae spoke more about variability. Specifically, she emphasized how the range and the modal clump or cluster (Konold et al., 2002) could help in determining a typical hat size for the class. Next on the blackboard Mae constructed with her students a line plot of the class's head sizes. See Appendix C for the distribution of her students' head sizes. The term *cluster* was mentioned prior to the actual construction of the line plot, and afterwards the terms *cluster* and *outlier* were defined and identified.

After creating the class's line plot, Mae had her students report out the measurements of the individual male, Kevin's, and female, Stephanie's, respective heads. She did this because some students were absent when the data were collected and they all needed it for their homework. Finally, she allowed her students to answer the questions on the lesson's worksheet about the class's distribution. While the students were working on the questions, she discussed with the whole class a student's response to the problem regarding the maximum value. During this time she also had small group discussions on the minimum value and on possible measurement error. She summarized her lesson by having the students share out their answers to the worksheet. At this time she clarified to a student why 56 is not an outlier to her, and used her perceived outliers and clusters to describe the distribution. At the end of her class, a postlesson discussion took place in her classroom.

The sensemaking notices generated from addressing the previous research question guide these findings. The analysis included whether Mae used the sensemaking notices the same way or in contrast to what she did prior to teaching her lesson. Also, extensions of what she previously made visible in her lesson planning, performance and professional development problems were sought. For the most part through her use of measures of center, range, and the

values she perceived as outliers, more of her knowledge of variability was exhibited during this first teaching lesson on variability. Specifically, in the introduction of the lesson, her use of measures of centers is similar to her use in professional development. In addition, during the middle to end of her lesson new knowledge was revealed regarding her use of the range. Finally, her use of her perceived outliers was extended as she instructed and interacted with students in the middle and latter parts of her lesson. The postlesson interview also revealed more information on her knowledge of outliers as she perceived them.

### **Measures of Center: Mae's Sensemaking During Lesson Implementation**

This section compares how Mae used measures of center in her teaching to what she did previously. Of all that Mae had done prior to this lesson regarding measures of center, it was believed that she would use the mean and the median to analyze the distribution or use modal clumps (Konold et al., 2002) to locate the typical value. It seemed less likely that she would connect the relationship between mean and median in discussing variability as she had previously done. Based on the fact that it is an introductory lesson, it was felt that she might not find it important to discuss this with her students.

In the very first discussion of Mae's lesson on variability, measures of center surfaced. After Mae's students measured each other's head sizes she discussed variability, mode, cluster, and range. Mae directed conversations with her students around the definition of variability—how far apart and how close together the data values are. Specifically, she wanted her students to use the definition of variability to help determine a typical hat size for her class. In this discourse she moved her students' thinking from the mode to more of the modal clump of the data distribution (Konold et al, 2002). This is significant in that the mode is not an appropriate measure for numerical data and at best it is an unreliable one.

Mae: What is variability? What do you think that is going to mean for our data?

Mae: What information is going to tell us that that is the best choice?

Student: Average size

Mae: Well, not so much the average

Student: The mode

Mae: The mode, okay

Student: The range

Mae: The range that's a good word. What else?

Student: The one that has the most counted votes of that number

Mae: Okay that's the mode, but I like the range. Because, let us say that we have a hat size, that is, the most we have is 55 as a hat size; should we get a hat size that is 57?

Student: No

Mae: Should we get a hat size that is bigger? What else should we look at besides the range, which is how far apart the data are? What else should we look at? Looking at the definition, what else should we look at?

Student: The mode of letters.

Mae: Is there one mode? Not always. What if I told you most of the hat sizes were here [motions with her hands a certain section on the scale] and only one or two were over here?

Student: The closer data

Mae: That's called a cluster—the amount of data that are altogether. That is a cluster of values. So you will most likely get a hat size that fits this [she again motions with her hands the same section on the scale] data but not one that is larger.

Here Mae did not seem to accept the student's suggestion that the average size would be the best choice of head size for the class. Doing so implies that she did not acknowledge

that the mean in and of itself is enough to indicate variability. On the other hand, she accepts the mode initially. However, she then proceeded to move her students' focus more toward the modal clustering or clump of the data. In this context, it seemed she did this in order to describe variability and to determine the hat size for the class.

Mae's movement from mode to modal clump (Konold et al., 2002) in the beginning of her lesson is similar to what she did during lesson planning. In her lesson planning, she focused upon the questions to ask her students during her lesson. She wondered if she would ask her students to partition the data with reference lines. Specifically, she was trying to determine what using reference lines would afford the students in terms of describing the typical value. Originally when she wanted to find the typical value she mentioned the mode. Yet, later in her planning she determined that including reference lines in her lesson's problem would help her students identify the modal clump of the distribution (Konold et al., 2002). In that lesson planning discussion, she stated that reference lines would locate the range where most of the data were clustered. This is similar to what she did here in her lesson. In her lesson she discussed the variability of a distribution of student head sizes. As noted in this introductory segment of her lesson, she refocused her students' attention from the mode toward the modal clump of the distribution (Konold et al., 2002). That is, from a singular measure of center to a modal clustering where most of the data are located. She did this when motioning with her hands the section of the scale where the data clustered.

It is not surprising that Mae was consistent in her use of modal clumps (Konold et al., 2002) to discuss variability for two reasons. First, she focused upon her definition of variability as how far apart and how close together the data values are. Based on this, Mae sought to emphasize the distribution's clustering representing the closeness of the data.

Second, in contrast to the mode, clusters, especially the one where most of the data are located, modal clump (Konold et al., 2002), gives a great deal more information about variability. With knowledge of the modal clump (Konold et al., 2002) you can see how the data are dispersed: in other words, how far apart and how close together most of the values are in the distribution. Thus, with her definition guiding her it is not surprising that Mae would emphasize modal clumps (Konold et al., 2002) in her lesson on variability.

Mae's leading her students' thinking from mode to the modal clump (Konold et al., 2002) is significant. Mode in and of itself is not appropriate to analyze numerical data. There sometimes also can be more than one mode as Mae pointed out to her students in this discussion. In contrast, locating the modal clump (Konold et al., 2002) gives a visual sense of where most of the data are located. It also can indicate where the center of the distribution is. Thus, it ties in a data distribution's center and dispersion that is essential when describing distributions (Makar and Confrey, 2005).

In addition to measures of center, Mae also used values she perceived as outliers to discuss variability with her class. The following section discusses Mae's knowledge of her perceived outliers that emerged during her teaching.

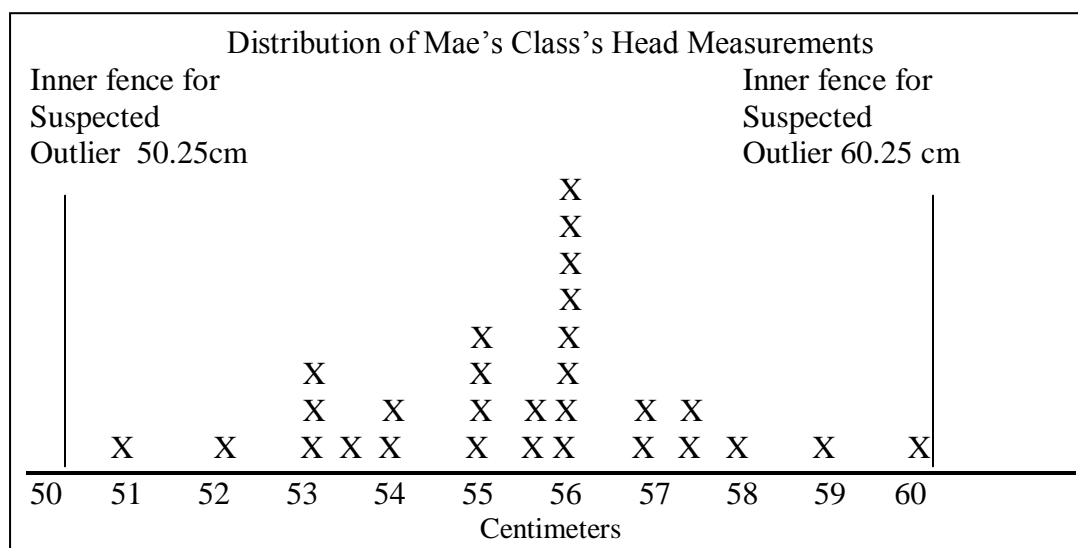
### **Perceived Outliers: Mae's Sensemaking During Lesson Implementation**

This section compares how Mae used her perceived outliers in her teaching to what she did previously with them. Entering into the analysis, there was speculation about the knowledge of outliers Mae would exhibit while teaching. Based on her experiences in this dissertation study, it was very probable that Mae would not formally treat her perceived outliers. Instead, it was anticipated that she would use informal language to discuss them. Although it was not certain how Mae would implement them, there were some expectations

based upon how she worked with them in lesson planning and professional development.

Based on the nature of collecting her class's head sizes, it was conjectured that she would treat her perceived outliers the same as she did in professional development Problem 1.4 (which had Mae analyze distributions of hypothetical students' and a hypothetical class's head measurements similar to those of her class).

Just as in Chapter 6, this section focuses on how Mae defined, referred to, and treated the values that she perceived as outliers. Also in accordance with Chapter 6, the values Mae considered outliers are called her *perceived outliers* whether they were outliers, suspected outliers, or not outliers. The following parts of this section are based on the distribution of her class's head measurements (in centimeters) that is illustrated in Figure 7.1.



*Figure 7.1 Distribution of Mae's Class's Head Measurements With Inner Fences for Suspected Outliers*

### **Defining Variability Includes the Concept of Outlier**

For the purpose of consistency and coherence with the discussion of outliers in Chapter 6, a brief comment on Mae's inclusion of her perceived outliers in her introductory lesson on variability is warranted. Mae followed through with her plan to discuss the concept of outlier

and clusters in her lesson. She both discussed with her students her definitions of them and placed questions about them on the worksheet that was used in her lesson. Since this section is on her perceived outliers, the focus is on Mae's use of them in interpreting variability and her treatment of them.

### **Mae Defines Her Perceived Outliers**

When the values she perceived as outliers came up in her lesson, Mae expanded on the meaning she had previously given them in her professional development problem. Specifically, before answering the questions on the lesson's worksheet about outliers, Mae discussed with her students what outliers were to her. In the process of introducing them to her students, she revealed more of the meaning of her perceived outliers.

Mae: Now an outlier is a number that is nowhere near the data.  
Can anyone find at least one?

Student: 58

Mae: 58? Why would you say 58 appears as an outlier?

Student: Because there is only like 4 of them that have only one square on it—  
an X.

Mae: Okay, but an outlier means a number that is nowhere near the rest of the data.  
But that is one that is not a popular size. That's good information.

Student: 60.1

Mae: Yeah, that 60.1 that is all the way on the other board [classroom blackboard]  
by itself. What else might be an outlier? Not only that, but there is another  
one.

Student: 51

Mae: It is nowhere near this huge cluster.

Here Mae described her perceived outliers as “nowhere near the data,” and more specifically, as “nowhere near the rest of the data,” and “nowhere near this huge cluster.” She also characterized her perceived outlier in the distribution of her students’ head sizes as being “by itself.” Specifically she did so in using as an example her student’s head size of 60.1 centimeters. This student’s head measurement required an extension of the line plot onto another section of the blackboard.

In her lesson, Mae’s description of her perceived outliers is clearer than what she previously discussed in solving her professional development problem. Problem 1.4 in the *Data Distributions* text (Lappan et al., 2006) involved analyzing hypothetical students’ head measurements. When describing her perceived outliers in the professional development problem, Mae labeled them as “not around the cluster.” In contrast, here in her lesson Mae described her perceived outliers more clearly by stating they are “nowhere near the rest of the data” and “nowhere near this huge cluster.” Thus, during her lesson she informally emphasized in a more descriptive way the distance her perceived outliers were from the data: specifically, that the perceived outliers were far from the rest of the data values. Table 7.1 depicts the phrases Mae used to define her perceived outliers prior to and during her teaching.

<b>Mae’s definition of her perceived outliers</b>	
<b>Prior to teaching</b>	<b>During teaching</b>
“not around the cluster”	“nowhere near the data” “nowhere near the rest of the data” “nowhere near this huge cluster”

*Table 7.1 Phrases Mae Used to Define Her Perceived Outliers Prior to and During Teaching*

Mae’s informal description of her perceived outliers is closer to the formal definition cited earlier in Chapter 2, that is, an outlying observation that appears to deviate

markedly from other members of the sample in which it occurs (Grubbs, 1969, quoted in “Outlier,” n.d.). Mae’s use of the phrase “nowhere near” informally indicates that she perceived outliers to be very far away from the major cluster of data. In this way, Mae’s informal definition also aligns with the conceptual underpinnings for the “add or subtract  $3(IQR)$  to the third and first quartile respectively” method of determining outliers. This further indicates that Mae perceives that outliers are a certain distance from most of the data, such as the interquartile range (IQR).

It was not surprising that when teaching her students Mae would discuss more about values she perceived as outliers than previously reported in her professional development. In an introductory lesson on variability, understanding terminology is key. Thus, it was assumed that Mae might be more explicit about her meaning of them when teaching than in discussing them during professional development. It seemed highly likely that she would explain more clearly to her students what her perceived outliers meant to her. She did this explaining just before the students began working on the lesson’s worksheet. Perhaps she did this to prepare her students to answer the question she placed on the sheet that asked if there were any outliers in the distribution.

When using the “ $\pm 3(IQR)$  to the third and first quartile respectively” method to identify outliers as described previously, outliers in this distribution were considered values at and above 64 cm at the high end of the distribution and at or below 46.50 cm at the low end. On the other hand, when using the “ $\pm 1.5(IQR)$  to the third and first quartile respectively” method, suspected outliers were located at the inner fences of 60.25 cm at the high end, and 50.25 cm at the low end of the distribution. Based upon these calculations, the values that Mae perceived as outliers were close to being suspected outliers. The largest outlier Mae perceived was 60.1

cm, which is .15 cm away from the inner fence for the suspected outlier of 60.25 cm. The smallest outlier Mae perceived was 51 cm, which is .75 cm away from the inner fence for the suspected outliers of 50.25 cm. Figure 7.1 depicts the inner fences for the suspected outliers in the distribution of Mae's class's head measurements.

Despite the fact that Mae did not identify outliers as a statistician would, her perceived outliers were close to what would be considered a suspected outlier. As discussed earlier, without any formal experience with identifying outliers, Mae's perceived outliers were based on identification of extreme values rather than actual calculations. In accordance with one of the middle school texts used throughout the large urban system in which Mae works (titled, *Quick Review Math Handbook*)—"outliers are data that are more than 1.5 times the interquartile range from the upper and lower quartiles" (p. 34). Although Mae seemed unaware of this definition, based upon it, Mae appeared to be close to identifying outliers for the level of students she was charged to teach. Mae's informal use of the concept of outlier also aligns with the 2005 GAISE Curriculum Framework that suggests informal treatment of outliers for the learner at Level A (Franklin et al., 2007). Basically in her work here, Mae performed appropriately based on her limited experiences with outliers as observed in this dissertation study.

### **Context Involved in Mae's Treatment of Her Perceived Outliers**

Similar to Mae's work in professional development, through the following examples it is conjectured that context might have influenced Mae's treatment of her perceived outliers when discussing variability. Just as in Chapter 6, the first example presented demonstrates a context in which Mae included her perceived outliers in her discussion of variability. However, in contrast to her work with her perceived outliers discussed in Chapter 6, there were

no examples that represented a context in which Mae excluded her perceived outliers in her analysis.

**Including Her Perceived Outliers.** After describing her perceived outliers, Mae shifted to evaluating them. It was somewhat uncertain how Mae would evaluate her perceived outliers. She might evaluate them as she did in solving her professional development problem. In professional development Problem 1.4 she analyzed hypothetical distributions of a class's head measurements. It was assumed that she might evaluate her perceived outliers in her lesson comparably to what she did in this problem because the context was the same as in her lesson: using a class's head measurements to determine one hat size to fit all.

It was during the analysis of her class's distribution of head measurements that Mae evaluated her perceived outliers. Specifically, during a summary point of her lesson she began to examine them. In the context of students' head measurements, she revealed that her perceived outliers were not just outliers to her, but rather they were data values that represented students.

Mae summarized the discussion of her students' head measurements by describing the distribution's variability and its deviations. It was then she gave her perceived outliers meaning in the context. She stated, "And then you would find two atypical people. One that is on the extreme of 51 and the one that is on the other extreme of 60.1." Here Mae categorized her perceived outliers in this distribution as "atypical." Her choice of adjective infers that her perceived outliers are individuals in the class whose head measurements do not fit the typical range of students' head sizes in the class. She also identified these values as being located on the extremes, which can imply values as different as possible from each other and/or values

deviating furthest from a central position or grouping. In either case, Mae knew her perceived outliers to be atypical students.

Mae's comment here reinforces her making sense of her perceived outliers as students located on the extremes of the distribution, and as students whose head measurements do not fall into the typical group of students. This view of her perceived outliers is supported by comments she made in her postlesson discussion. There Mae reiterated accepting a student's remark in her lesson that referred to her perceived outlier as a girl with a big head. Please note that the 60.1 cm value was a female student.

Particularly in this postlesson discussion, Mae summarized the students' use of new vocabulary. She acknowledged it took some processing for her students to learn the meaning of the term *outlier*. She stated:

They [the students] were confused because there was new terminology like the word outlier... Some of them started using it 'cause they heard me using it/explaining it. Then that one kid was like, yeah, I think that one is an outlier because it is just totally by herself with the big head.

Here Mae reiterated a student's comment about her perceived outlier. The student informally identified her perceived outlier in the class's distribution of head sizes as a fellow student "totally by herself with a big head." By restating this comment, Mae implied that to her an outlier in her class's distribution is an atypical student. The student's statement about being "totally by herself" could possibly be based on the fact that this value literally required an extension of the class's line plot, and as a result, was placed on another section of the board. This extension of the line plot to accommodate the 60.1 cm head measurement on another board gave this student the impression that it was totally by itself. In actuality there was a data value of 59 cm that is numerically close to the 60.1 cm, but this was possibly obscured by the way 60.1 cm was graphically depicted on the board.

Mae was consistent in how she evaluated her perceived outliers. In her lesson and postlesson discussion, she evaluated them the same way as in solving her professional development problem. In the context of the class's head measurements, she treated her perceived outliers as valid data values. To her they were atypical students' head sizes that were located "on the extreme." She did this even when discussing her perceived outliers in the professional development problem's hypothetical class. Despite the fact that the students Mae perceived as outliers for the class in professional development Problem 1.4 were fictitious, Mae included them in her analysis. It was as if she expected that in a class there would be students whose head measurements would not be a part of the typical head sizes of the rest of the students.

In both her professional development and her lesson, it was conjectured that context played a role in Mae's treatment of her perceived outliers. Specifically, it was speculated that when the context in her analysis of student's head sizes involved finding a hat size for the whole class, she would not exclude her perceived outliers. It was also speculated that Mae's inclusion of her perceived outliers, which she referred to as atypical students, was further influenced by the fact that she expected greater variability in the head sizes for the whole class. Based on this, it seems that for her to find one hat size to fit all, she could not disregard the atypical students who were on the extremes. Finally, it was conjectured that Mae's lack of catching any measurement errors at the time of measuring also contributed to the inclusion of her perceived outliers in this context.

To shed additional light on Mae's including her perceived outliers for the distribution, two segments from our postlesson discussion are presented. This first segment addresses Mae's expectation of variability for the class:

Researcher: So why did you think that they would be able to get a typical hat size for Stephanie and Kevin?

Mae: Just more data about one person as opposed to data about everybody... And there is only that one person they have to focus on. Whereas here we are looking at 30 different people and we are trying to find one hat size for 30 different people—as opposed to finding one hat size for one person.

Later on, Mae continued to articulate the difficulty in finding one hat size to fit the class as opposed to the individual student. She stated,

Because there would be more values [for the class] Wait, when we did Jalin and that other kid [Problem 1.4], we were able to find a more probable head size as opposed to the whole class. Because there were different people with different head sizes, and here there is only one person, so it has got to be closer to the hat size...

Here Mae discusses that finding one hat size for only one person is easier than finding one to fit 30 different people. She attributes this to more data about one person as opposed to 30 data values for 30 different people. Based on her comments, it seems as though she sees more variability in the whole class's head sizes than the one student. It is not uncommon to expect less variability in repeated measures of the same person than in measures representing 30 different people (Konold and Pollatsek, 2002).

As previously mentioned, it was also conjectured that Mae included her perceived outliers in her analysis based on not seeing any measurement errors take place in the measuring. It is not that she was unaware of the possibility of measurement errors, because she did see one when the students measured the head size of the male student, Kevin. Evidence of this is again present in our postlesson discussion, when she stated,

Is it possible that their hat [head] size would be all the way over here with just that one number with that data? Or, where they, where it might be too small? Because it is only one person who measured that you know? The height difference, because he would be too short to measure up any higher. I am serious. Yeah, they were sitting down on the stool. But then we had some people who were

shorter than others; and Kevin sitting on the stool was pretty tall; and Kenny's not that tall at all. I am serious that changes where they are measuring, because they just cannot reach.

Here it seems that because Mae saw the error occur when Kevin's head was measured, she noted it. Based on this, it is conjectured that Mae did not mention measurement errors when the class measured their heads because she did not observe any occurring. Therefore, she did not have a need to exclude her perceived outliers for the class based upon measurement errors because to her it appeared that no errors occurred.

In her lesson, it was not surprising that Mae used informal language to discuss her perceived outliers in her lesson. Her informal language use along with her informal treatment of her perceived outliers is possibly based on her experiences in this dissertation study. During the professional development, Mae used a text, *Data Distributions*, which was written for middle school students (Lappan et al., 2006). Appropriately so, this text did not formally present outliers because it was geared toward middle school students, who might not be ready for formal treatment of outliers until they make sense of these formal measures and why they would use them.

**Excluding Her Perceived Outliers.** For the purpose of consistency and coherence in how outliers were discussed in Chapter 6, a brief comment on Mae's excluding her perceived outliers is warranted. During her lesson, Mae did not exclude any values that she considered outliers. It is conjectured that because the context was the same as in her professional development problem, choosing a hat to fit all her students' heads, she seemed not to exclude any of her perceived outliers in her lesson. Again, this might be based on her viewing the perceived outliers as atypical students that could not be excluded when determining the class's hat size, barring her knowing about any measurement errors.

This is in contrast to her treatment of them during professional development when she excluded an outlier. When analyzing Henry's computer game reaction times in Problem 3.2, Mae excluded the outlier because it was a "bad reaction time" and it did not follow his typical pattern of speed in reacting to the game. It was conjectured that Mae was able to exclude her perceived outlier because she expected bad reaction times when playing a computer game, and because excluding the outlier did not affect the context or purpose of the inquiry (describing general consistency). This is opposed to eliminating an atypical student when finding a hat size for the class.

Evaluating outliers is important. Mae's early stage of seeing when to include or exclude them is critical for two reasons. First, as Aczel, (1996) stated, "because of the possible information content in outliers, they should be carefully scrutinized before one decides to discard them" (p. 503). If they are based on an error in measuring they can be discarded. Second, curriculum frameworks, such as the 2005 GAISE Report (Franklin et al., 2007), emphasize being able to informally differentiate variation from error. Mae's work with her perceived outliers hints at a beginning to develop a sense of variation and error. At this time, she knew that there would be natural variability or differing head sizes based on the different students in her class. She also knew there could be measurement errors. Further, she accepts a certain level of variability among her class's head measurements that includes atypical people because they are not near the cluster. Mae's work might be considered foundational to further developing (possibly through more experience with analyzing data distributions from different contexts) knowledge about variability, including outliers.

In summary of this view of Mae's use of her perceived outliers, she basically extended the meaning she gave them in professional development. She added more to her informal definition of her perceived outliers that included a qualitative description of their distances from the main cluster of data, such as being "nowhere near this huge cluster." This informal description was close to the conceptual underpinning of the " $\pm 3(\text{IQR})$  to the third and first quartiles respectively" method of determining outliers that identifies the main grouping of data and how far the outliers are from it.

Also in her teaching, Mae confirmed the way she evaluated her perceived outliers in her professional development problem. In Chapter 6, it was conjectured that context influenced how Mae treated her perceived outliers. It is believed that based upon her treatment of her perceived outliers in her lesson, this conjecture is still credible. In the context of finding a hat size that would fit the whole class, Mae included those values she perceived as outliers. In addition, Mae's expectation of greater variability with the class's head sizes (than with the individual students), coupled with her not detecting any student errors while measuring, influenced the inclusion of her perceived outliers in her analysis.

One final commentary on Mae's use of her perceived outliers in interpreting variability is warranted. It would be simple to say that Mae did not know what outliers were because she did not determine them as a statistician would. However, it is hoped that the presentation of Mae's informal language and treatment of her perceived outliers in this section points toward a mindset of her knowing the conceptual underpinnings of the formal treatment of outliers. That is, they are a marked deviation from a cluster of typical students. Also, Mae began to see informally the natural variability in a distribution of her class's head sizes. Mae's informal use of language and treatment of her perceived outliers might be considered a positive place for

Mae to be when beginning to make sense of outliers in discussing variability. Makar and Confrey (2005) stated, informal language could bring to life the connections between the various themes of analyzing and comparing distributions. They stated,

The nonstandard language used...by its very nature *integrates* the important statistical ideas of variation and distribution...The process of integrating rather than separating concepts in statistics has been shown to be a productive avenue for developing statistical thinking and reasoning (p. 49, emphasis in original text).

In addition, they stated that the informal language could be used as a bridge to the more formal meaning of this concept and the more formal treatment of this concept when and if it is necessary (p. 49). Thus, Mae's use of informal language regarding outliers can help to connect the relationships in the data as well as to connect to more formal treatment of outliers in the future.

In addition to discussing the values she perceived as outliers, Mae knew to use range to help her students come to know variability. This is discussed in the following section.

### **Range: Mae's Sensemaking During Lesson Implementation**

This section compares how Mae used the range in her teaching to what she did previously. It was uncertain how Mae would use range in her lesson. It was anticipated that she would address its meaning as a measurement. However, it was unclear how she would use it to analyze her class's distribution.

From the very beginning of her lesson Mae addressed range. In her initial discussion on variability she defined range as an indicator of how far apart the data values are. She also used the maximum value to set the outer limit of where possible data values would be found on the scale of the distribution. Finally, she used the context of students' head measurements to determine whether the magnitude of the range was significant.

Mae discussed range in the beginning of the lesson. She noted that the range is

important in determining the typical hat size for the student(s). When a student offered it as a helpful way to determine the typical hat size, she approved.

Student: The range

Mae: The range that's a good word. What else?

Student: The one that has the most counted votes of that number

Mae: Okay that's the mode, but I like the range. Because, let us say that we have a hat size, that is, the most we have is 55 as a hat size; should we get a hat size that is 57?

Student: No

Mae: Should we get a hat size that is bigger? What else should we look at besides the range, which is how far apart the data are?

Here Mae accepts the range as a way to help determine the class's hat size. She used a hypothetical maximum value of the distribution that is integrally tied to the range. In particular, she used this hypothetical value to set the parameter for the possible hat sizes from which to choose. This is a basic use of the range. Later on in the middle of her lesson, Mae states that the range is an integral part of variability. Then, she also extended her use of the range as a measurement when she discussed its magnitude.

When students were completing the worksheet for the lesson, Mae addressed the maximum value and range as a part of variability. While they worked, she conversed with each of the small groups in her class. In response to question C1, Mae noted a student's answer on the maximum value. She disagreed with the student's answer regarding range and maximum value and mentioned it to the whole class.

Mae: Someone said that the maximum was 56 and I disagree. Now we are talking about the range because the range is a part of our variability of our data.

Student: I think I know why the person thought that the maximum was 56, because it had the most Xs.

Mae:                So what would be our maximum?

Here Mae disagreed with a student's choice of maximum value. In doing so, she explicitly connected the range and maximum value to variability. Her statement makes clear that knowing variability includes this measure of dispersion. Mae went on shortly thereafter to elaborate her use of the range as a measurement.

As noted above, in the middle of her lesson Mae mentioned that the range is a part of variability. Further along in that part of her lesson, when analyzing the data distribution of the class's head measurements, she addressed the magnitude of the range. This came up in reaction to a student's comment about picking Mae's perceived outlier as the typical hat size to get for the class.

Mae:                Why would it make sense that we should get 56?

Student:           Because 56 has the most Xs, that means most of the people have a head size that is 56.

Student:           I think we should get 60.1 because that is the biggest person.

Mae:                Okay, but what about this 51 all the way over here? That's a huge, that's a 9-centimeter difference. What would be a typical hat size if I ran into someone in your class based on these clusters? Can you give me a range of values that would be typical?

Student:           55 to 57.5

Mae:                So here this range [points to that section of the scale with her hands], because most of the data is there. What else? Any other group that you see?

Student:           53–54

Mae:                Okay so that is another cluster

Here Mae's student chose the maximum value as the appropriate hat size for the class because it represented the biggest person. In response, Mae questioned her student's thinking

and brought up the magnitude of the range. She seemed to challenge the student who thought that buying the hat size for the largest head would be appropriate. In reaction to his comment, Mae emphasized what it means to consider a hat size that is 9 centimeters away from the minimum value. She wanted her student to realize that 9 centimeters is a huge difference in head measurements.

It was somewhat surprising that Mae brought up the magnitude of the range. It is the first time in this research Mae mentioned this aspect of the range. Previously, when solving professional development Problem 1.3 she discussed how knowing the range is not enough to discuss the consistency of data. Particularly she saw this to be true when analyzing a student's (Henry's) computer game reaction times. She realized that when an outlier was present the range did not tell you in and of itself about the pattern of consistency in his reaction times.

However, in her lesson this interaction with her student indicates that Mae knew the greatness in the magnitude of the range—9 cm in relation to head sizes. Her use of range indicates that she evaluated the magnitude of the range in context. The context of student head measurements in centimeters influenced her determining the significance of the range's magnitude. A 9 cm range seemed large to her and, therefore, choosing a value she perceived as an outlier for the class's hat size seemed very unreasonable to her.

It is important to use the range as a measurement. It is an integral part of the variability of the distribution and a measure of the spread of a distribution. Curriculum frameworks such as the 2005 GAISE Report (Franklin et al., 2007) stress this importance. Level B of this framework advocates students using measures of variability including the range.

It is also important that Mae knew when the magnitude of the range was useful in this particular context. It demonstrates a more critical use of the range. Makar and Confrey (2005) discuss the importance of this:

Variation encompasses more than just a measure, although measuring variation is an important component in data analysis. In considering variation, one must consider not just what it is (its definition or formula), or how to use it as a tool (related procedures), but also why it is useful within a context (purpose) (p. 28).

Mae used the context in her lesson to determine the usefulness of the range in discussing variability. When Mae evaluated the magnitude of the range in the context of head measurements, she used it critically. Her knowledge went beyond a rote use of it as a measure. In this way, Mae exhibited knowing when and how the range is useful for analysis of data distributions.

In summary, in the beginning of her lesson, Mae defined range. She referred to its use in measuring how far apart the data are. This use of range was picked up in the middle of her lesson when she began analyzing the data distribution of her class's head sizes. At that time, she affirmed the integral role range plays in discussing variability or the spread of the distribution. She also extended her use of the range by discussing its magnitude. For the first time in this study, Mae brought out how the magnitude of range was significant in discussing variability. Specifically, this happened in her teaching when she questioned a student's suggestion of considering her perceived outlier as the hat size to buy for the class.

In an overview of her lesson, Mae did not use all of her sensemaking notices when analyzing distributions that she used prior to teaching to discuss variability. However, regarding the ones she did use, she was consistent in using them. She used *measures of center*, *values she perceived as outliers*, and *range* in discussing variability. In particular, she continued to use modal clumps (Konold et al., 2002) to locate the typical value. Further,

context continued to play a role in evaluating whether to include or exclude her perceived outliers. Mae included them based upon the context or purpose of the analysis, such as in determining one hat size to fit her class. Mae's expectation of variability along with not catching any measurement errors might have influenced when she would include her perceived outliers. On the other hand, Mae excluded an outlier when the context or purpose of analysis supported doing so.

Further, context played a role in Mae's analysis of her class's head sizes with the range. Mae did not take the range as a measure at face value in discussing the variability. Similar to how she treated her perceived outliers, the context played a role in Mae's evaluating the range's usefulness. When Mae calculated the range, she did not use it by rote. For example, the significance of its magnitude was based upon the context of students' head measurements.

Regarding Mae's use of the other sensemaking notices reported in Chapter 6 for interpreting variability, Mae did not use *shape* or *partitioning* in her lesson. This is in contrast to her use of them in interpreting variability during professional development, lesson planning, and her performance task. Specifically, in her lesson she did not use descriptive terms, such as *skewed*, to discuss the shape of the distribution of her students' head sizes. Nor did she partition the distributions to discuss variability. Perhaps she did not use this sensemaking notice because she was not comparing unequal size distributions, and based on this, she did not need to partition sections in order to analyze percentages of clusters of data. Finally, regarding her use of the *definition of variability*, Mae used her definition throughout her lesson as a guide. As such, it was not treated separately in reporting the implementation of her knowledge of variability during her lesson.

This dissertation research used a bidirectional approach to studying teacher content knowledge that is a twofold view of teaching that encompassed studying the teacher prior to her classroom lesson, and during her classroom lesson. Chapter 6 focused upon the knowledge Mae exhibited in her work prior to teaching in the classroom. This chapter focused upon viewing the knowledge Mae exhibited when teaching in the classroom. The bidirectional approach that guided this analysis and that produced these findings is discussed further in the Chapter 8 that follows.

## **Chapter 8**

### **Discussion**

This study was designed to explore an in-service middle school teacher's (Mae's) content knowledge of variability of data distributions. This dissertation study's primary contribution is a bidirectional view of studying teacher content knowledge. This is in contrast to research that focuses on a unidirectional view of teacher knowledge, such as looking at performance tasks with interviews to indicate a teacher's knowledge for teaching (e.g., Ma, 1999). Instead, this bidirectional view of content teacher knowledge involved focusing on what teachers exhibit knowing prior to teaching and what they exhibit knowing during teaching. In particular, this bidirectional view supported a more complete snapshot of the teacher's content knowledge of variability in data distributions. This snapshot of the teacher's content knowledge was multifaceted. In one direction, what the teacher exhibited knowing prior to teaching was confirmed, extended, or not visible during her teaching. And in the other direction, new content knowledge was exhibited during her teaching that was not seen prior to teaching.

In addition to the bidirectional view, Mae's use of informal/nonstandard language added complexity to the analysis of her content knowledge. In particular through her informal/nonstandard language the teacher captured cognitive relationships, for example, between notions of center, shape, and dispersion. These relationships are not found in standard Statistics language, which tends to treat these statistical ideas as conceptually separate (Makar and Confrey, 2005). Through Mae's informal/nonstandard language she used these statistical ideas with varying degrees of fruitfulness in making sense of variability.

Some of the additional contributions of this study include:

1. Studying teacher knowledge as a dynamic phenomenon by using the construct of sensemaking (Weick, 1995) as interpreted by Drake (2006) as cognitive work that involves acts of noticing, interpreting, and implementing.
2. Empirically testing Canada's (2004) Evolving Framework that was derived from work with preservice teachers with an experienced middle school mathematics teacher.
3. Developing a framework for analyzing middle school teachers' sensemaking of variability of data distributions.
4. Assessing the usefulness of a statistics education curriculum framework, that is, the 2005 Guidelines for Assessment and Instruction for Statistics Education (GAISE) Curriculum Framework for pre-K–12 (Franklin et al., 2007) to analyze a teacher's content knowledge of variability.

In this chapter I will first review and interpret the results from this dissertation study. Then, I will discuss the contributions of this study, the limitations of the study, the implications of these results for mathematics classroom practice and teacher education, and future research suggested by these results.

### **Review and Interpretation of Results**

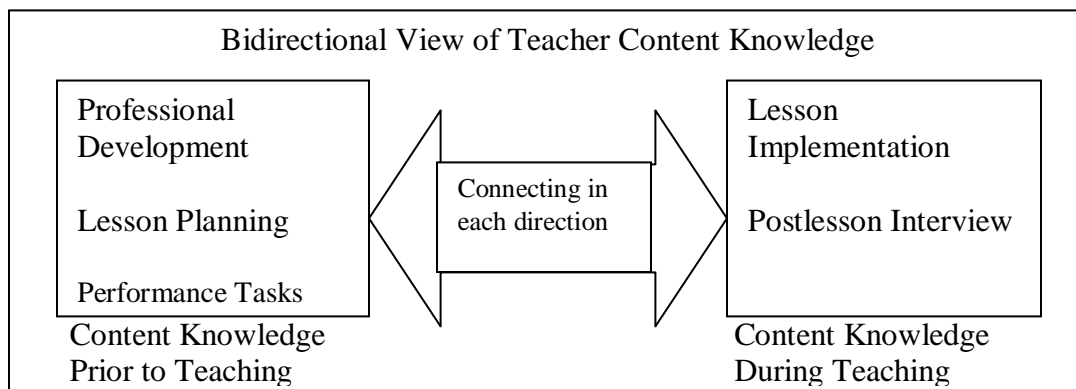
Teacher knowledge is important because it is considered essential for teaching (Shulman, 1987). As a result, research has studied teacher knowledge particularly of mathematics, in various ways. At first, teacher knowledge was studied using their mathematics coursework as a proxy for what they knew. Results from two important studies, Begle (1979) and Monk (1994), showed that advanced mathematical understanding

contributed little to teacher effectiveness and that there were some significant effects of courses in undergraduate mathematics pedagogy on student achievement. Another way teachers' knowledge was studied was to look at the nature of teacher's knowledge. This approach was a qualitative focus on the teacher's mathematical knowledge of specific topics. It involved interviewing teachers prior to teaching to explore their knowledge for teaching. One of the most significant contributions with this closer focus on teacher knowledge has been a new conception of subject matter knowledge for teachers called pedagogical content knowledge. This is a special kind of knowledge that links content with aspects of teaching and learning (Shulman, 1986). A third approach to studying teacher knowledge is viewing it as the knowledge that is pedagogically functional—what teacher's know, how they know it, and what they are able to mobilize when teaching (Ball et al., 2001, p. 451). This was also done through tasks that focused on this type of teacher knowledge.

All three perspectives on studying teacher knowledge offer important insights. However, they might not offer the most complete snapshot of what teachers know because they do not include studying the knowledge that is displayed in the moments of teaching. The knowledge that is displayed prior to teaching might not portray the knowledge that is displayed during teaching. They might be different, similar, or complementary to each other. Both are important in understanding teacher knowledge. As a result, I connected these two views—prior to and during teaching—and took a bidirectional view of teacher content knowledge in this dissertation study. (In contrast to some of the aforementioned studies that did not specifically focus on content knowledge as a dynamic system in the way this study does.)

Figure 8.1 illustrates this bidirectional view of teacher content knowledge. Similar to the studies discussed previously, a view of teacher content knowledge prior to teaching is important. For my study, the content knowledge exhibited prior to teaching included the knowledge exhibited in professional development experiences, lesson planning, and performance tasks with interview. Together the content knowledge the teacher exhibited across these experiences gave the view of the teacher's content knowledge prior to teaching. In contrast, the second view focused on the content knowledge the teacher exhibited in the other direction, that is, during lesson implementation including a postlesson interview.

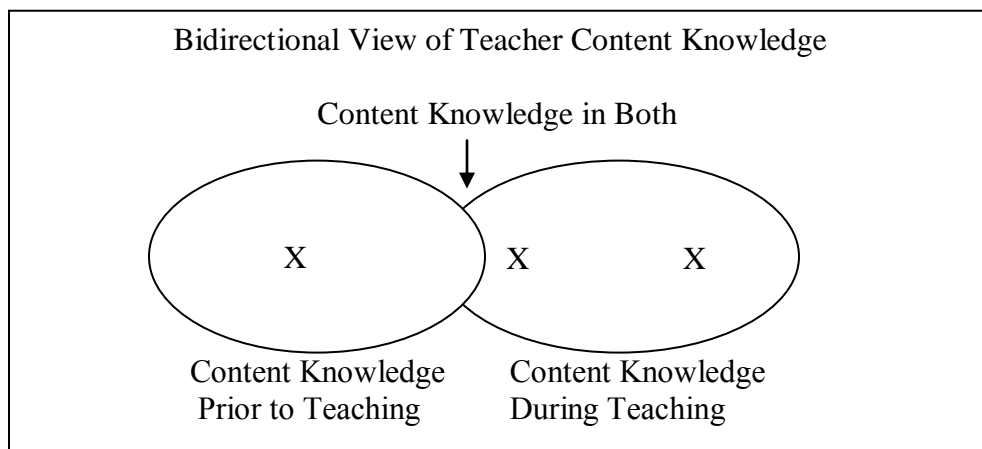
Figure 8.1 illustrates the bidirectional view of the teacher's content knowledge of variability that I discuss in this chapter. In this study, a bidirectional view of content knowledge means that the assumption of unidirectionality of content knowledge for teaching (knowledge prior to teaching is applied during teaching) is challenged and deemed incomplete. Meaning that at the moment of teaching, teachers not only use content knowledge they had prior to teaching but that they also might generate new knowledge during and after implementing lessons that help to then inform, elaborate, or revise their prior knowledge.



*Figure 8.1 Bidirectional View of Teacher Content Knowledge*

Figure 8.1 shows the view of teacher content knowledge exhibited prior to teaching connecting to the teacher content knowledge exhibited during teaching. The connecting arrow in the diagram indicates not just viewing teacher content knowledge in both directions. It also suggests that some of what the teacher exhibits knowing prior to teaching is connected to what she knows during teaching, and vice versa, some of what the teacher exhibits knowing during teaching connects to what the teacher knows prior to teaching.

This dissertation study used this bidirectional view of teacher content knowledge as a way of getting a more complete snapshot of teacher content knowledge. The diagram in Figure 8.2 depicts the results of its use.



*Figure 8.2 Results of Using Bidirectional View of Teacher Content Knowledge*

As the diagram in Figure 8.2 illustrates teacher content knowledge that is exhibited prior to teaching does connect with the teacher content knowledge exhibited during teaching. However, it also indicates how they are connected. The findings of this study showed that when viewing teacher content knowledge in both directions certain knowledge was made visible:

1. only prior to teaching,

2. only during teaching,
3. both prior to and during teaching.

One example of the content knowledge of variability Mae exhibited when she discussed range is shown in Table 8.1.

Prior to teaching	Both	During teaching
<ul style="list-style-type: none"> <li>Range might not be useful to describe overall consistency when an outlier is present</li> </ul>	<ul style="list-style-type: none"> <li>How to calculate range as a measure of variability</li> </ul>	<ul style="list-style-type: none"> <li>Conceptual underpinnings of calculating range</li> <li>Range is integral to describing variability</li> <li>Context aids in determining the significance of the magnitude of the range.</li> </ul>

*Table 8.1 Bidirectional View of Mae's Content Knowledge of Variability When Discussing Range*

Table 8.1 illustrates that Mae's content knowledge was exhibited prior to, during, and both prior to and during her teaching. This, in turn, indicates that Mae's knowledge prior to and during teaching is connected. It also indicates that a bidirectional view of teacher content knowledge can give a more complete snapshot of it. This is compared to a unidirectional view—only looking at the content knowledge exhibited either prior to or during teaching—in which case some of Mae's content knowledge would not have been visible.

Table 8.1 only represents Mae's content knowledge that emerged when she discussed the statistics topic of range. Across the board, Mae's content knowledge of variability emerged slightly differently depending on what she noticed when discussing variability. For Mae's other notices, sometimes her content knowledge was exhibited just

prior to teaching or sometimes in both. These differences might be based, for example, on the particular problems completed prior to or during teaching. Nonetheless, the results across all of what Mae noticed when making sense of variability indicate that Mae exhibited content knowledge prior to, during teaching, and both. Table 8.2 illustrates this for all of her notices.

<b>Mae's notices</b>	<b>Where Mae exhibited her knowledge of variability</b>		
<b>Sensemaking notices</b>	<b>Prior</b>	<b>Both</b>	<b>During Teaching</b>
Partitioning	X		(none exhibited)
Measures of center	X	X	(one piece of knowledge the same)
Perceived outliers	X	X	X (extended knowledge)
Range	X	X	X (extended knowledge and new knowledge exhibited)
Shape	X		(none exhibited)
Defines/describes variability	X	X	(same knowledge exhibited)

*Table 8.2 Bidirectional View of Mae's Content Knowledge of Variability*

As Table 8.2 shows, although the connection between each view of teacher knowledge—prior to and during teaching—might be particular to what is noticed in discussing variability, the results do suggest that there is a connection between each view of teaching. Future studies can investigate the use of this bidirectional approach to studying teacher content knowledge in the field of statistics and/or mathematics as well as in other school subjects. These studies can investigate if both a more complete view of content knowledge emerges, and the possible differences in, or the nature of, the connections in this knowledge.

The next section on the contributions of this study focuses on the construct of sensemaking (Weick, 1995; Drake, 2006) and the benefits of using it to study teacher knowledge as a dynamic phenomenon. That section describes how the cognitive practices

of sensemaking aligned with and helped to explain the connection made between each view of teacher content knowledge.

### **Contributions From the Study**

The results of this study contribute to and extend the research literature on teacher knowledge and statistics education research. For research on teacher knowledge, these results attest to the value of using the construct of sensemaking to study teacher content knowledge as a dynamic phenomenon. For statistics education researchers, the results of this study offer the following (a) an empirical test of Canada's (2004) Evolving Framework work with preservice teachers with an experienced middle school mathematics teacher; (b) a framework for analyzing middle school teachers' sensemaking of variability in data distributions, and (c) an assessment of the usefulness of a curriculum framework for pre-K–12 statistics education, the 2005 GAISE Report (Franklin et al., 2007), to analyze a teacher's content knowledge of variability.

### **Sensemaking as a Construct to Study Teacher Knowledge**

Teacher knowledge is an expansive concept that can be challenging to study. As previously discussed, teacher coursework has been used as a proxy for it. The mathematics and mathematics education courses teachers took stood for the knowledge the teacher had of mathematics (Monk, 1994; Begle, 1979). In other studies, teachers' performance on certain mathematical tasks with interviews represented the qualitative nature of teacher knowledge and/or their pedagogical content knowledge (e.g., Ma, 1999). And finally, other studies used teachers' performances on tasks to stand for the specialized knowledge that is needed for teaching, that is, pedagogically functional knowledge (Ball et al., 2001). In contrast, this dissertation study sought to look at teacher content knowledge as a

dynamic phenomenon represented by the sense a teacher made of it. This stance is different from the other studies.

Although other studies seek to know the sense that is made of mathematics (whether via coursework, performance on certain mathematical tasks, or on tasks that gauge pedagogically functional knowledge), the sensemaking model (Weick, 1995; Drake, 2006) used in this study takes another view of teacher content knowledge. Because this study seeks to connect the knowledge the teacher exhibits prior to teaching to what she exhibits during teaching, the sensemaking model that was used brings a more dynamic aspect to viewing the content knowledge.

In sensemaking, there are three cognitive practices—noticing, interpreting, and implementing (Weick, 1995; Drake, 2006). Each practice is a part of sensemaking, and when using sensemaking as a proxy for knowledge, as this study does, content knowledge can be culled from what one notices, interprets, and implements (Weick, 1995; Drake, 2006). This is analogous to a cartographer's sensemaking where the terrain is unknown. The sense he or she makes of the terrain (in this study variability) depends on and is indicated by what is noticed, how it is interpreted and, finally, how it is represented in a map (or, in this study, a lesson).

As previously discussed in Chapter 5, this study found that when making sense of variability Mae noticed a compilation of: a strategy, concepts, and definitions/descriptions used in the fields of statistics and statistics education. These constituted her *sensemaking notices* and emerged from her utterances in professional development, lesson planning, and in solving her performance tasks with an interview. How she discussed these notices when

interpreting variability gave insights into how she made sense of variability, and when implementing her lesson, her sensemaking was brought more into view.

By analyzing Mae's sensemaking practices, her content knowledge of variability was culled. Based upon her experiences in this dissertation study, Mae's content knowledge was not expected to match that of a statistician. At best the content knowledge she exhibited through her sensemaking might be considered a basis upon which to build. The informal or nonstandard language Mae used during her sensemaking is often not the language of statisticians. What statisticians might see as incomplete knowledge or possibly a misconception is considered a point in Mae's trajectory of knowing variability. As a coach and teacher, there is something to be said about Mae's content knowledge and how it points toward potential for more learning.

Three resources were used to identify Mae's content knowledge: Grossman, Wilson, and Shulman's (1989) definition of content knowledge, the concept map depicted in Chapter 2, and the 2005 GAISE Curriculum Framework (Franklin et al., 2007). Grossman, Wilson, and Shulman (1989) stated, "We use the term content knowledge to refer to the 'stuff' of discipline: factual information, organizing principles, central concepts. In addition to being able to identify and discuss concepts separately, an individual with content knowledge can identify relationships among concepts external to the discipline" (p. 27). The facts Mae exhibited knowing were not necessarily the major facts of knowing about variability or what a statistician would say it means to know variability. Nonetheless, the facts about variability that Mae exhibited knowing might be considered important points in her trajectory of knowing variability.

The concept map depicted in Chapter 2 also helped frame the content knowledge that Mae exhibited through her sensemaking practices. In this concept map, characterizing variability included discussing the distribution's clusters, gaps, as well as measuring variation through the interquartile range, range, outliers, and the Mean Absolute Deviation. In this dissertation study, Mae's utterances on range, outliers, and clusters to a lesser or greater degree were a part of her content knowledge. Finally, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) for statistics education was another tool in the analysis of Mae's content knowledge. This framework helped to identify what is important in the landscape of knowing variability in statistics education. This included knowing that variability can be based on natural variation or measurement error.

Table 8.3 gives a snapshot of the results of what Mae did during her sensemaking, and of the content knowledge of variability that was exhibited during this sensemaking. As stated above, based upon the work Mae did in this study, the content knowledge she exhibited might be considered a foundation upon which more knowledge can be built. (For a view of where this content knowledge was exhibited—prior to or during teaching—see Table 8.2.)

<b>What Mae did during sensemaking</b>	<b>Content knowledge of variability Mae exhibited</b>
<b>Partitioning</b>  a) Used percent of data appropriately when analyzing unequal size data sets b) Used multiple reference lines to make sense of variability c) Saw percent of data above and below the mean	Mae exhibited knowing that: <ul style="list-style-type: none"> <li>• proportional reasoning can be used to discuss variability</li> <li>• partitioning helps see the variability of a distribution</li> </ul>

*Table 8.3 A Snapshot of Mae's Sensemaking and Content Knowledge of Variability*

Table 8.3 (cont'd)

<p><b>Measures of Center</b></p> <ul style="list-style-type: none"> <li>a) Chose a problem with measurement data in order to see variability</li> <li>b) Saw reference lines as helpful to identify the main cluster or modal clump (Konold et al., 2002)</li> <li>c) Saw that the relative locations of mean/median could be based on the data's clustering and values perceived as outliers</li> </ul>	<p>Mae exhibited knowing that:</p> <ul style="list-style-type: none"> <li>• identifying the main cluster or modal clump (measure of center and dispersion of most of data around it) is a part of seeing variability</li> </ul>
<p><b>Perceived Outliers</b></p> <ul style="list-style-type: none"> <li>a) Defined her perceived outliers</li> <li>b) Saw that context influences including and excluding her perceived outliers</li> </ul>	<p>Mae exhibited knowing that:</p> <ul style="list-style-type: none"> <li>• the overall pattern in variability has deviations</li> <li>• the conceptual underpinnings of calculating outliers</li> <li>• variability can be based on natural variation</li> <li>• variability can be based on measurement error</li> </ul>
<p><b>Range</b></p> <ul style="list-style-type: none"> <li>a) Defined range</li> <li>b) Calculated range</li> <li>c) Determined the usefulness of range when to her an outlier is present</li> <li>d) Discussed magnitude of range in relation to context</li> </ul>	<p>Mae exhibited knowing:</p> <ul style="list-style-type: none"> <li>• how to calculate range to measure variability</li> <li>• the conceptual underpinning of calculating range</li> <li>• that the range is integral to describing variability</li> <li>• that the range is not necessarily a useful measure to describe variability when an outlier is present</li> </ul>

Table 8.3 (cont'd)

<p><b>Shape</b></p> <p>a) Discussed asymmetry b) Discussed shape of main cluster</p>	<p>Mae exhibited knowing that:</p> <ul style="list-style-type: none"> <li>the shape of the main cluster could indicate how the typical value varies</li> </ul>
<p><b>Defines and Describes Variability</b></p> <p>a) Defined variability b) Described variability c) Defined and described variability</p>	<p>Mae exhibited knowing that:</p> <ul style="list-style-type: none"> <li>clusters, gaps, and outliers are a part of characterizing variability</li> </ul>

*Table 8.3 A Snapshot of Mae's Sensemaking and Content Knowledge of Variability*

As indicated on Table 8.3, each particular sensemaking notice gave possibly new insights into Mae's content knowledge of variability. At the same time, each sensemaking notice was not equally fruitful in the content knowledge Mae had about variability.

The quality of the content knowledge Mae exhibited knowing through her sensemaking practices was based upon what she notices, and how she interprets and implements them. If any of her sensemaking notices are unconventional or lacking, then it could prove potentially less fruitful as an indicator of her knowledge of variability. This also holds for her interpreting and her implementing. The content knowledge Mae exhibits knowing through her sensemaking practices are as good as what she noticed and interpreted prior to and during lesson implementation. The results of this dissertation study indicated that not all of her sensemaking notices were equally fruitful for interpreting variability. For example, as previously discussed in Chapter 5, some of Mae's notices, such as locations of measures of center in connection to variability, were unconventional. Therefore, the content knowledge of variability that can be claimed as a result might not be considered as

solid as that of her other notices when compared to what the statistics community considers official knowledge of statistics.

Mae's informal language use combined with her unconventional connection of the locations of measures of center to variability added complexity to the claims of her knowledge. On the one hand, Mae's unconventional connection points to the possibility of her knowing that the locations of measures of center are influenced by aspects of the variability of the data, such as its outliers. However, on the other hand, with the nuances of the informal language she used, this was not explicit. Based on this, the claim to her knowledge of variability is less solid than the other claims that can be made when she uses more conventional and widely accepted official language of statistics. The complexity involved in analyzing Mae's informal language is discussed further in the limitations section.

### **Empirically Testing Canada's (2004) Evolving Framework**

Canada (2004) studied elementary preservice teachers' (EPSTs) expectations, displays, and interpretations of variation within three statistical contexts: repeated sampling, data distributions, and probability outcomes. His work produced an Evolving Framework that characterized EPSTs' thinking about variation. Table 8.4 is an outline of his framework. His framework is comprised of aspects, dimensions, and themes of his EPSTs' thinking about variation. (The bolded items represent where some of Mae's sensemaking touched upon his themes.)

Evolving Framework	
<p><b>[1] Expecting Variation</b></p> <ul style="list-style-type: none"> <li>A] Describing What is Expected <ul style="list-style-type: none"> <li>i) Concerning Expected Value</li> <li>ii) Concerning Repeated Values</li> <li>iii) Concerning Range or Extremes</li> </ul> </li> <li>B] Describing Why (Reasons for Expectations) <ul style="list-style-type: none"> <li>i) Involves Possibility or Likelihood</li> <li>ii) Involves Experiential Reasoning</li> <li><b>iii) Involves Proportional Reasoning</b></li> <li>iv) Involves Distributional Reasoning</li> </ul> </li> </ul> <p><b>[2] Displaying Variation</b></p> <ul style="list-style-type: none"> <li>A] <b>Producing Graphs</b> <ul style="list-style-type: none"> <li>i) <b>Technical Details</b></li> <li>ii) <b>Characteristics of the Distribution</b></li> </ul> </li> <li>B] <b>Evaluating and Comparing Graphs</b> <ul style="list-style-type: none"> <li>i) <b>Focus on Average</b></li> <li>ii) <b>Focus on Range or Extremes</b></li> <li>iii) <b>Focus on Shape</b></li> <li>iv) <b>Focus on Spread</b></li> </ul> </li> <li>C] <b>Making Conclusions about Graphs</b> <ul style="list-style-type: none"> <li>i) <b>Emphasizing Decisions in Context</b></li> <li>ii) <b>Emphasizing</b> Consistency or Reliability</li> <li>iii) Emphasizing Level of Detail &amp; Usefulness</li> </ul> </li> </ul> <p><b>[3] Interpreting Variation</b></p> <ul style="list-style-type: none"> <li>A] <b>Causes and Effects of Variation</b> <ul style="list-style-type: none"> <li>i) <b>Definitions &amp; Descriptions</b></li> <li>ii) Examples</li> </ul> </li> <li>B] <b>Influencing Expectations and Variation</b> <ul style="list-style-type: none"> <li>i) <b>Naturally Occurring Causes</b></li> <li>ii) <b>Physically Induced Causes</b></li> </ul> </li> <li>C] Effects of Variation <ul style="list-style-type: none"> <li>i) Effects on Perception</li> <li>ii) Effects on Decisions</li> </ul> </li> <li>D] Influencing Expectations and Variation <ul style="list-style-type: none"> <li>i) Quantities in Sampling</li> <li>ii) Number of Samples</li> </ul> </li> </ul>	

*Table 8.4 An Outline of Canada's (2004) Evolving Framework of Elementary Preservice Teachers' Thinking About Variation*

As previously explained in Chapter 5, Canada's (2004) Evolving Framework was an initial guide to identifying Mae's sensemaking notices. It was found that most of Mae's sensemaking notices aligned with Canada's (2004) themes of average (measures of center),

range, extreme values (which Mae called outliers), and shape. This might be based on his EPSTs' work in evaluating and comparing graphs that overlapped with Mae's work in analyzing or comparing data distributions.

The findings of this study indicate that Canada's (2004) Evolving Framework was a helpful tool in investigating Mae's sensemaking notices. Yet, it was not a perfect fit for a number of reasons. In contrast to Canada's (2004) work that involved repeated sampling, data distributions, and probability outcomes, the work in this dissertation study involved only analyzing and comparing data distributions. Thus, the amount of overlap in the findings of our respective studies was lessened. In addition, in contrast to his elementary preservice teachers' teaching experiences, or lack thereof, Mae's experiences as a middle school teacher for five to ten years might have led her to have different notices about variability.

In this vein, one of Mae's sensemaking notices—partitioning (with proportional reasoning)—was not a part of his EPSTs' reasoning when comparing graphs. Canada (2004) stated that this was probably based on his preservice teacher's lack of graph sense because they did have experiences with proportional reasoning in previous coursework (p. 130). Although his EPSTs' use of proportional reasoning did surface, however, it was in their work with expecting variation in probability outcomes and not analyzing distributions.

As a result of the less than perfect fit in the components and findings of our respective work, a beginning framework emerged from this dissertation study. Table 8.5 displays this framework for middle school teachers.

## Framework for Analyzing Middle School Teachers' Sensemaking of Variability of Data Distributions

Table 8.5 represents the framework for analyzing middle school teachers' sensemaking of variability that emerged from this dissertation study. It differs from Canada's (2004) Evolving Framework in that its scope is limited to statistical work involved in comparing and analyzing data distributions.

<b>Framework for analyzing middle school teachers' sensemaking of variability of data distributions</b>	
<b>When analyzing and comparing graphs, what teachers might use:</b>	<b>Examples of what teachers might do:</b>
a) Partitioning using proportional reasoning	a) Use multiple reference lines to make sense of variability
b) Measures of center	b) Identify modal clumps (Konold et al., 2002)
c) Perceived outliers	c) Define, include, and exclude values perceived as outliers
d) Range	d) Calculate range and evaluate its usefulness when determining consistency
e) Shape	e) Describe shape of clusters and the whole distribution
f) Defining and describing	f) Use concepts such as clusters, gaps, and outliers when defining and describing variability

*Table 8.5 Framework for Analyzing Middle School Teachers' Sensemaking of Variability of Data Distributions*

This framework for studying a middle school teachers' sensemaking of variability of data distributions is not meant to critique or to be considered equivalent to Canada's (2004) Evolving Framework. However, based on the different populations studied and the different types of statistical tasks used in each study, it might be an emerging tool to use for

analyzing middle school teachers' sensemaking of variability of data distributions. Its appropriateness for middle school teachers might be based on the proportional reasoning that is considered more developmentally suitable for the age group they teach.

Designing this framework was not the intent of this study and at most it represents a beginning framework. The usefulness of this framework can be investigated in future studies of middle school teachers' sensemaking of variability of data distributions.

Although the effectiveness of the *CMP Data Distributions* text (Lappan et al, 2006) that was used is not implicated in this study, this possible framework's usefulness might be investigated in conjunction with it. In this way, the framework might be critiqued and possibly more or different sensemaking notices might become evident indicating possibly new or different content knowledge.

### **Assessing the Usefulness of the 2005 GAISE Curriculum Framework for Discussing Teacher Content Knowledge of Variability**

The 2005 GAISE Report (Franklin et al., 2007) is a coherent curriculum framework for statistics education grades pre-K–12. The 2005 GAISE Curriculum Framework breaks down learning statistics into three levels: Level A, Level B, and Level C. These levels are not meant to be grade based. They are connected to the statistical investigation process where the depth of understanding and sophistication of methods required increases across the levels. Based on the nature of this study, only parts of the 2005 GAISE Curriculum Framework (Franklin et al., 2007) are applicable. These include the process component of the statistical investigation, analyzing the data, along with the nature of variability and the focus on variability.

Table 8.6 indicates how Mae's content knowledge aligned with the 2005 GAISE Levels of these process components. It also indicates where her content knowledge

emerged prior to or during teaching. In addition to these, Table 8.6 shows how Mae's sensemaking notices align with the 2005 GAISE Levels and, further, where these notices emerged—prior to or during her teaching.

Mae's sensemaking notices	GAISE Level A		GAISE Level B		GAISE Level C	
	Prior to lesson	During lesson	Prior to lesson	During lesson	Prior to lesson	During lesson
Partitioning	X		X			
Measures of center	X	X				
Range	X	X	X	X		
Perceived outliers	X	X				
Shape	X					
GAISE process component	Prior to lesson	Lesson	Prior to lesson	Lesson	Prior to lesson	Lesson
Analyze data	X	X				
Nature of variability -Measurement variability -Natural variability -Induced variability	X	X				
Focus on variability: -Variability within a group -Variability between groups		X	X			

*Table 8.6 Alignment of Mae's Content Knowledge with 2005 GAISE Levels and Bidirectional View of Teaching*

As indicated by Table 8.6, the findings of this study show that the bulk of Mae's content knowledge fell at Level A with some at Level B. As previously discussed, it was not expected that Mae would be at Level C. The nature of the work Mae engaged in this study did not match the curriculum expectations of Level C. Specifically, her work did not correspond with the content of the curriculum, the level of sophistication of methods, or the

depth of understanding that Level C requires. One example is the formal treatment of outliers, which is not required by the 2005 GAISE Curriculum Framework (Franklin et al., 2007) until Level C.

Nonetheless, this study found that the 2005 GAISE Curriculum Framework (Franklin et al., 2007) was helpful to analyze Mae's content knowledge of variability for two reasons. First, through articulating the levels of the pre-K–12 curriculum, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) aided in situating the sense Mae made of variability in its curriculum continuum for statistics education—Levels A, B, or C. In addition, the framework provided a way to discuss how Mae was making sense of variability. In this way, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) helped determine the pertinence of Mae's content knowledge to statistics education.

Because the 2005 GAISE Curriculum Framework (Franklin et al., 2007) aided in analyzing Mae's content knowledge of variability, future research can address its use in studying other concepts in statistics education. In addition, its use in combination with this study and other research (e.g., Makar and Confrey, 2005) could inform teacher educators of teachers' informal and nuanced language use as well as possible problem areas in teachers' sensemaking of variability.

### **Limitations**

While this study contributes to research literature, there are limitations to the generalizability of these results. More research is needed to determine whether these results about an in-service middle school teacher's content knowledge of variability in data distributions are valid only for a middle school teacher with 5–10 years of experience

working in a large urban setting. Perhaps they might apply to other populations of teachers, or to teachers with different levels of experiences with statistics.

In addition, my analysis of Mae's content knowledge relied upon her utterances throughout this study in her professional development, lesson planning, performance tasks with interview, lesson implementation, and postlesson discussion. There were complexities in the process that included my relationship with Mae, the natural effects of being studied, and her use of informal language. These are discussed in the following.

First, as a participant observer my relationship with Mae was implicated in each site of data collection—professional development, lesson planning, performance task interview, lesson implementation with postlesson discussion. Although efforts were made to mitigate any power relationship that might have ensued based on my role as the school's mathematics coach, it is still possible that her utterances and behavior were more constrained because of it. Mae also might have been nervous and possibly more reserved based on the fact that her words were being recorded and her actions videotaped. Yet, the effects of these potential limits might have been lessened for a few reasons.

For one, efforts were made to make the atmosphere in all sites of data collection friendly and accommodating. Also, Mae is by her nature a very outgoing and verbal person who spoke much during the study. This was precisely the reason she was chosen for this dissertation study. Speaking much is not a factor of what is being said. Nonetheless, through data collection over time and across different sources, it is reasonable to argue that an authentic snapshot of Mae's content knowledge emerged. In spite of some of the possible drawbacks of the power relationship between me and Mae and the effects of being studied, studying teacher content knowledge bidirectionally—prior to and during

teaching—afforded more access to her content knowledge. This is in comparison to what studying teacher content knowledge in only one direction (for example, prior to teaching with performance tasks) would make possible.

Furthermore, the analysis of Mae’s content knowledge relied upon my interpretations of her utterances throughout this dissertation study. There were complexities in this process due in part to Mae’s use of informal/nonstandard language when discussing variability. In contrast to standard language that tends to keep concepts separate, informal/nonstandard language is more complex and nuanced. Makar and Confrey (2005) found that when using informal language, notions of variability, shape, and central measures tend to be integrated. As a result, these statistical concepts were harder to separate and added complexity to the analysis of Mae’s sensemaking. For clarity and guidance, references from statistics education and statistics education research were used to aid in the analysis. Nonetheless, the results of this research represent one possible analysis of Mae’s content knowledge of variability of data distributions.

Lastly, the results of this study were affected by time in two ways. First, this study took place over approximately a two-month period and involved observing one of Mae’s lessons. Therefore, the findings are limited to a small part of Mae’s teaching. As a result, more research is needed to determine the content knowledge of variability Mae would exhibit across a number of lessons. Second, the constraints of the school year and the teacher’s personal calendar shortened the original professional development schedule. Less professional development led to a less than thorough treatment of the *CMP Data Distributions* text (Lappan et al, 2006). This, in turn, curtailed the scope and possible depth of experiences Mae had with variability. With more time to interact with the curriculum

and a colleague, a different snapshot of Mae's content knowledge might have emerged. Despite this, efforts were made to ensure certain parts of the curriculum were presented in professional development. As a result, Mae did have some foundational experiences with variability in data distributions, but much less than what was originally anticipated.

In closing, I would not interpret the results of this dissertation study as an evaluation of the effectiveness of the *CMP Data Distributions* text (Lappan et al. 2006) for two reasons. First, this research was not designed to do so. Second, as previously stated, in professional development the *CMP Data Distributions* text (and as a result the teacher's experiences with variability) did not receive the full treatment that the authors intended. Therefore, there is no claim either way that the text was or was not a major factor in Mae's sensemaking of variability in data distributions.

### **Implications**

While I thought about sharing the results of this study with teachers and teacher educators, I was not sure what parts would be helpful to them. I believe informing Mae of her content knowledge of variability might have both positive and negative effects. On the positive side, Mae might perceive that she knew more than she thought she did of this new statistics topic, or she might see the fruitfulness of using some of her *notices* over others in making sense of variability. On the negative side, Mae might feel that she was evaluated, and based on this might be uncomfortable with her content knowledge falling at, for example, Level A of the 2005 GAISE Curriculum Framework (Franklin et al., 2007). Nonetheless, a discussion combining the results with the use of this framework, might help her to realize the breadth of the landscape of statistics and how its many components unfold across the curriculum. Also, more importantly based on this discussion, she might decide

to use the framework perhaps along with the CMP *Data Distributions* text (Lappan et al., 2006) to help develop her content knowledge and her students' content knowledge of variability.

Teacher educators could also benefit from an introduction to statistics education documents such as the 2005 GAISE Curriculum Framework (Franklin et al., 2007) and the research that uses it. Together these could assist in their practice in two ways. First, the framework might be useful for both guiding instruction and assessment of the teachers they might mentor, coach, or instruct. Also, as previously stated, the framework could be used along with curriculum materials, such as the CMP text *Data Distributions* (Lappan et al., 2006), to develop teachers' content knowledge together with their knowledge of the framework's continuum for pre-K–12 statistics education. Furthermore, research that used this Framework, such as this study, could help to inform teacher educators of possible problem areas in teacher sensemaking of variability.

Furthermore, the insights into Mae's content knowledge reported in this study indicate that still more is needed to teach it. Statistics education research has suggested ways that teachers' content knowledge of variability can be developed. Researchers emphasize that teachers need to have opportunities to learn statistics similar to their students, such as through many experiences with collecting and analyzing their own data (Makar and Confrey, 2005, p. 30). This is in contrast to using the precollected data in texts that have, for example, precalculated outliers and, therefore, might limit the necessary and rich discussion involved in making sense of them. Another benefit is that collecting and analyzing data over a number of experiences can be helpful to develop teachers' expectations of natural variability, for example, in a class's head sizes, or for making sense

of repeated measures of an individual student's head. Together, over time, these experiences have the potential to develop the intuitive knowledge that teachers seem to lack about stochastics (statistics and probability) (Shaughnessy, 2007).

### **Future Research**

While this dissertation provided new approaches to studying teacher content knowledge as a dynamic phenomenon, and some insights into a middle school teacher's content knowledge of variability of data distributions, more research is needed. This research extends into the fields of education, mathematics, and statistics education.

#### **Future Research for Education and Statistics and/or Mathematics Education**

The results of this dissertation study might contribute to research literature.

However, the results also suggest more research is needed regarding:

1. extending the generality of this study including
  - a. a framework for analyzing teacher sensemaking of variability of data distributions,
  - b. new content knowledge
2. extending the domain of this study including
  - a. the bidirectional view of teacher content knowledge,
  - b. sensemaking (Weick 1995; Drake, 2006) as a construct to study teacher knowledge,
  - c. the 2005 GAISE Curriculum Framework (Franklin et al., 2007) for analyzing a teacher's content knowledge of variability, and
3. studying variability and its entailments.

### **Extending the Generality of This Study**

As previously mentioned, this study was conducted in a large urban middle school with one middle school teacher. These results are not necessarily generalizable to all populations. A small-scale descriptive study allows for developing models and new hypothesis to test in larger samples. It is possible that results from this dissertation will hold in other settings, but investigations of larger populations of teachers from different locations would expand the framework for analyzing a middle school teacher's sensemaking of variability in light of new data. Additionally, new content knowledge of variability might manifest with more teachers in more classrooms.

### **Extending the Domain of This Study**

The results of this study could be extended by conducting similar analysis of other statistical concepts, other fields of mathematics and/or other academic domains, such as social studies, English, or science, to contrast with students' experiences in mathematics classrooms. This includes extending the following:

**Bidirectional View of Teacher Content Knowledge.** This study discussed a bidirectional view of studying teacher content knowledge—prior to and during teaching. This is in contrast to what is typically done in education research—studying it in one direction either prior to or during teaching. The results determined that there is a connection between teacher content knowledge exhibited in each direction, and that a more complete snapshot of the teacher's content knowledge was attained. Although promising results were obtained regarding this approach to studying teacher content knowledge of variability, its usefulness and applicability are yet to be investigated in other academic subjects, such as social studies, English, or science. These studies could determine

whether, through the use of this bidirectional approach, a more complete view of content knowledge emerges along with possible differences in, or significance of, the connections between the content knowledge exhibited in each direction.

**Sensemaking as a Construct to Study Teacher Knowledge.** This study used the construct of sensemaking (Weick, 1995; Drake 2006) to study teacher knowledge as a dynamic phenomenon. Findings from this study demonstrated that it aligned with the bidirectional view of studying teacher content knowledge. Findings also indicated that the fruitfulness of the teacher's cognitive sensemaking practices—noticing, interpreting, and implementing—could affect the strength of the content knowledge that she exhibits. Future research could investigate the use of sensemaking to study other concepts in statistics or mathematics, or in other academic subjects.

**2005 GAISE Curriculum Framework.** In this dissertation study, the 2005 GAISE Curriculum Framework for pre-K–12 (Franklin et al., 2007) was useful in determining and discussing a middle school teacher's content knowledge of variability. For this study, it aided in providing the levels of development for statistical work that involved analyzing the data, nature of variability, and focus on variability. Also, the framework's level-based curriculum requirements assisted in placing the teacher's knowledge of variability along the continuum of pre-K–12 statistics education. Future research could investigate the usefulness of this framework in studying other statistical concepts such as covariation.

In addition, teacher educators could use the 2005 GAISE Curriculum Framework (Franklin et al., 2007) in alignment with research, such as this study and other studies on teachers' informal language use (Makar and Confrey, 2005). In this way, the 2005 GAISE Curriculum Framework (Franklin et al., 2007) could guide the teacher educator's

curriculum and assessment, while research could help inform them of teachers' informal and nuanced language use and the potential problem areas in making sense of variability.

### **Studying Variability and Its Entailments**

In this study various insights were made of the content knowledge of variability exhibited by a teacher. Of this content knowledge some was possibly more problematic than others. One example was Mae's connection of variability (e.g., its deviations as in the values she perceived as outliers) of the distribution to the relative locations of measures of center. Because this connection was somewhat unconventional, claims of Mae's knowledge of this was less than solid and as a result warrant further studies to determine how and if this informal knowledge could lead to developing a robust knowledge of outliers.

The particulars of how teachers build conceptions of statistics needs more study (Makar and Confrey, 2005). Therefore, these future studies could investigate, for example, the ways teachers come to make sense of variability and all of its entailments—such as measures of center in relation to it. This would in part answer the call of Shaughnessy (2007) who in his review of research in statistics education suggested that students' intuitive notions of center and variability be built upon. In this regard, I am suggesting the same for teachers who are just beginning to make sense of measures of center in connection to variability. More research on teachers in this area can continue to build a base of teacher sensemaking of variability from which, for example, teacher educators could draw upon.

In addition, future research could investigate ways teachers make sense of outliers. Because outliers require informal treatment at the 2005 GAISE Curriculum Levels A and B (Franklin et al., 2007), informal language will most likely be used when teachers make

sense of them. Therefore, these studies could also investigate ways of analyzing teachers' nuanced and complex informal talk regarding outliers that is intimately tied to their sensemaking. In this way, a growing base of teacher sensemaking and/or continuum of teacher content knowledge in these areas could prove valuable to both the research community and teacher education.

### Summary

In order to prepare our students to be statistically literate in the 21<sup>st</sup> century, our teachers need the content knowledge of statistics necessary to teach them. To this end, this study attempted to capture as complete a snapshot as possible of a middle school teacher's content knowledge of a foundational statistical topic—variability. In doing so, it took a bidirectional approach to studying teacher content knowledge—prior to and during teaching—along with a sensemaking model (Weick, 1995; Drake, 2006) to study teacher content knowledge as a dynamic phenomenon. Together they served to give a more complete and multifaceted snapshot of teacher content knowledge.

The findings of this study showed that there was a dynamic connection between teacher content knowledge exhibited prior to or during teaching. Teacher knowledge exhibited in one direction—prior to teaching—was either confirmed, extended, or not made visible during teaching. And in the other direction—during teaching—new content knowledge was exhibited that was not seen before.

This study also found that the teacher's cognitive sensemaking practices—noticing, interpreting, and implementing—affected the content knowledge of variability. Specifically, each of the *sensemaking notices* found in this study—*partitioning, measures of center, values perceived as outliers, range, shape, or defining and describing*

*variability*—afforded varying degrees of fruitfulness as indicators of content knowledge. The quality and unconventional nature of these notices might have been factors. This, along with the informal language used to discuss variability, gave insights into the complexities and nuances involved in the teacher’s sensemaking of variability.

In this vein, the results of this study could possibly inform teachers and teacher educators of entry points as well as problematic areas for learners when making sense of variability. In addition, teachers and teacher educators might benefit from knowing about some of the complexities and nuances in the learner’s informal utterances of variability. They also might decide to use resources such as the 2005 GAISE Curriculum Framework (Franklin et al., 2007) to guide their instruction and place the learner’s sensemaking along the continuum of statistics education. Furthermore, statistics education research could begin to use some of the tools discussed in this study—bidirectional view of teacher content knowledge and sensemaking as a proxy for teacher content knowledge—to investigate other statistical concepts. By doing so, research can continue to build upon a base of teacher content knowledge of various statistical concepts, such as variability and its entailments. All this work needs to be done to help prepare our students to become statistically literate for the 21<sup>st</sup> century.

### **Closing**

This dissertation study began with an awareness of the importance of teacher content knowledge of statistics and my own interest in knowing more about the field. When these interests intersected with my experiences with middle school teacher development, I chose to study a middle school teacher’s content knowledge of variability of data distributions. Through the process of studying Mae’s content knowledge of

variability in data distributions, I came to see how multifaceted her content knowledge is. Yet, I also saw how she is just beginning to make sense of the vast terrain of statistics, especially variability.

With all this being said about Mae's content knowledge of variability, it might be easy to think that there is a great deal more for her to know about variability. Yet a key point needs to be acknowledged. The very fact that Mae noticed variability (to whatever extent) is considered essential in the field of statistics. As previously stated in Chapter 1, noticing variability is an important aspect of statistical thinking. In Pfannkuch's and Wild's (2000, 2004) model one type of fundamental statistical thinking includes attention to or consideration of variation. This is not the only type of thinking inherent to statistics. However, it is a foundational one that separates it from the general types of thinking that are the hallmarks of mathematical thinking, such as looking for patterns, abstracting, generalizing, specializing, and generating and applying algorithms (Shaughnessy, 2007). Thus, Mae's noticing of variability in data distributions is significant, especially since it is not necessarily a mind-set developed in a teacher who, like herself, majored in mathematics.

In this vein, this dissertation study gave some insight into a middle school teacher's content knowledge of variability. It also discussed possible ways of studying teacher content knowledge as a dynamic phenomenon. All this was done with the intent to learn more about teacher sensemaking of this important statistical topic with the ultimate goal of preparing students to be statistically literate. In doing so, just like their students, our teachers need experiences with analyzing data that is given to them and data that has been collected or generated by them. In addition, we need research to continue building a

knowledge base of what teachers know about statistics and ways of studying it. Further, we need research to help inform ways of developing teacher content knowledge in teacher education coursework, professional development, and/or coaching. A growing knowledge base, such as this, could help ensure that teachers, even those like Mae who majored in mathematics, will be equipped to bring their students to a 21<sup>st</sup> century level of statistical literacy.

## Appendices

## APPENDIX A

### Interview Performance Tasks

1. A class of twenty-one 6th-grade students wanted to find out some information about MAX train rides. Their first goal was to find out the duration of a ride from Washington Park to Gresham. They all got on the same train, but they sat separately and kept track of the time on their own. Later in class, they were surprised to find that they did not have the same results:

<u>Duration of Ride (min:sec, to the nearest second)</u>				
58:36	58:36	58:44	58 :51	58:51
58:50	58:49	58:50	58:56	59:01
59:02	59:06	59:11	59:09	59:16
59:14	59:15	59:19	59:21	59:20
59:24				

What are some possible reasons for why the class did not get the same result?  
(Canada, 2004)

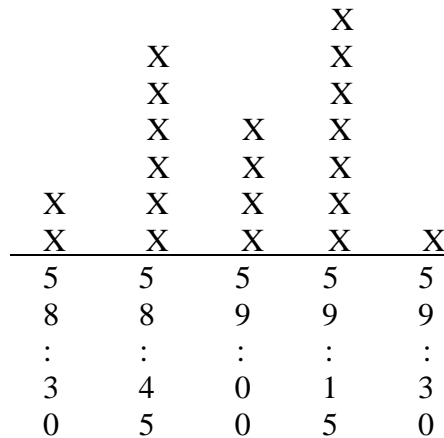
2. The class was deciding how to display their data. In Graph 1, they rounded to the nearest 15 seconds. In Graph 2, they rounded to the nearest 5 seconds.

**Graph 1** (Rounded to nearest 15 sec.)

Mean = 59:00

Median = 59:00

Mode = 59:15



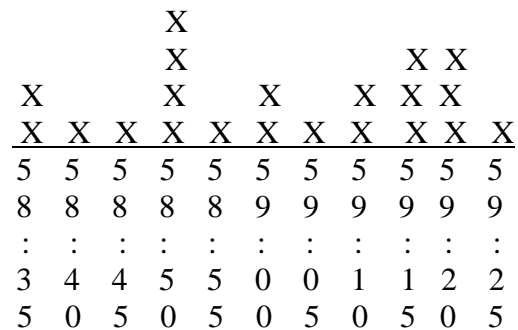
Duration of Trip  
(Minutes and Seconds)

**Graph 2** (Rounded to nearest 5 sec.)

Mean = 59:01

Median = 59:00

Mode = 58:50



Duration of Trip  
(Minutes and Seconds)

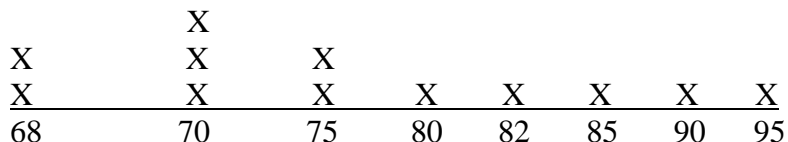
- How do these graphs differ in the stories they tell about the duration of the trip?
  - Some members of the class argue that the trip was really under 59 minutes, while some argue that it was over 59 minutes. Others claim it was exactly 59 minutes. What do you think about the true duration of the trip, and why do you think this?
  - Does one graph help you more than the other in making your conclusion?
- (Canada, 2004)

3. A new car was being tested to see how well the brakes worked. The test engineer measured how many inches the car took to slow from 40 mph to 0 mph; the fewer inches taken, the better the braking power. Twelve trials were run, under the same road conditions and with the same test driver. Here were the results (to the nearest inch):

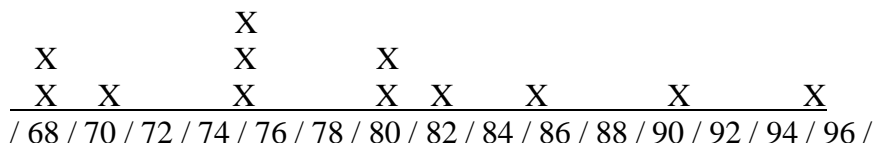
<u>Stopping Distance (in.)</u>			
68	68	70	75
75	75	80	80
82	85	90	95

The engineer was then trying to decide how to graph the results. She came up with the following three graphs for representing the data:

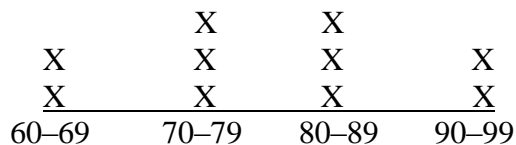
Graph 1



Graph 2



Graph 3



- Do these graphs differ in the way they show the braking power? If so, how?
- Do you think one graph shows more variability in the results than the others? Explain.
- If the engineer wanted to suggest that the car was fairly consistent in its braking power, which graph would you suggest she use, and why?

(Canada, 2004)

4. The Wait-Time for the MAX is defined as the interval of time which starts when one train leaves and ends when the next train arrives. In other words, the Wait-Time is how long there's no train at the station.

A class of twenty students wanted to find out if there was a difference in Wait-Times between Westbound and Eastbound MAX trains. They went and got the following ten Wait-Times for different Westbound trains and ten Wait-Times for different Eastbound trains (rounded to the nearest half-minute):

**Data: (Wait-Times in Minutes)**

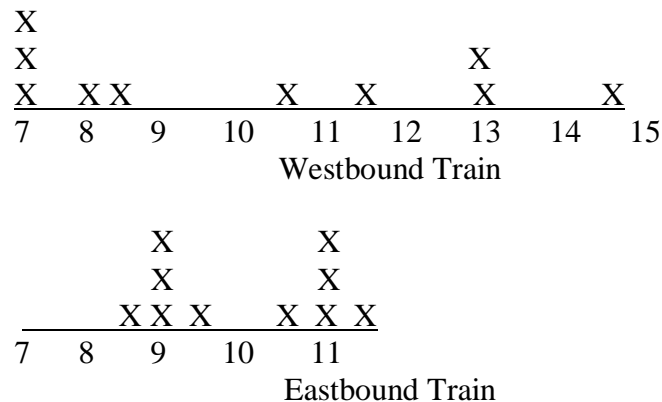
**Westbound**

7.0	7.0	7.0	11.5	10.5
8.5	8.0	13.0	14.5	13.0

**Eastbound**

8.5	9.0	9.0	11.0	11.0
9.5	9.0	11.0	10.5	11.5

Wait-Times for MAX Trains  
(In Minutes)



- What can you conclude about the Wait-Times for the two trains?
- One student in class argues that there's really no difference in the Wait-Times of the two trains, since the averages are the same. Do you agree?

(Canada, 2004)

5.

- [6] The following graphs describe some data collected about Grade 7 students' heights (measured in centimeters) in two different schools:

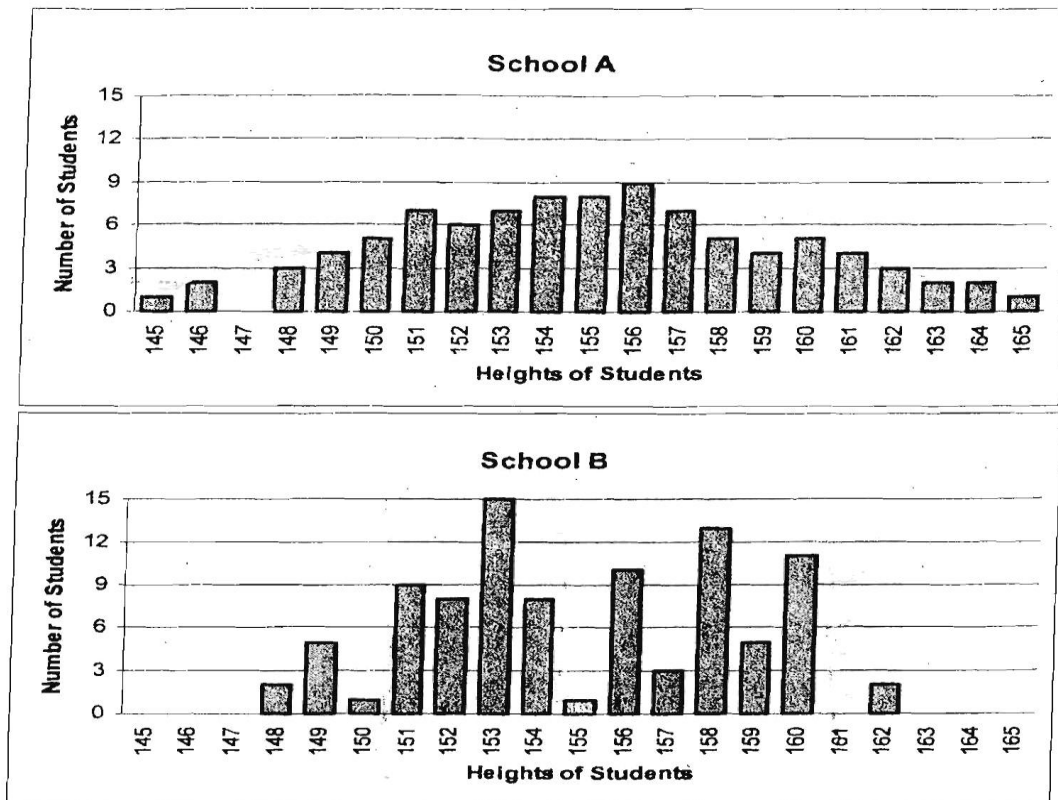


Figure 9.1 Graphs of School A and School B

Which graph shows more variability in students' heights?  
Explain why you think this.

(Canada, 2004)

6. Now consider two more classes, the PINK class and the BLACK class. The scores for the two classes are shown below, and once again each box is one person's test score.

# of people				PINK CLASS				
				X	X			
				X	X	X		
			X	X	X	X		
		X	X	X	X	X	X	
		X	X	X	X	X	X	
		X	X	X	X	X	X	
	X	X	X	X	X	X	X	X
1	2	3	4	5	6	7	8	9

Number Correct

# of people				BLACK CLASS				
						X		
						X		
					X	X	X	
					X	X	X	
			X	X	X	X	X	
	X	X	X	X	X	X	X	X
1	2	3	4	5	6	7	8	9

Number Correct

Again look at the scores of all students in each class, and then decide: Did the two classes do equally well on the test, or did one of the classes do better than the other? Explain how you decided.

(Watson & Moritz, 1999; Watson & Shaughnessy, 2004)

APPENDIX B

Process Component	Level A	Level B	Level C
<b>I. Formulate Question</b>	<p><b>Beginning awareness of the <i>statistics question distinction</i></b></p> <p>Teachers pose question of interest</p> <p>Questions restricted to the classroom</p>	<p><b>Increased awareness of the <i>statistics question distinction</i></b></p> <p>Students begin to pose their own questions of interest</p> <p>Questions not restricted to the classroom</p>	<p><b>Students can make the <i>statistics question distinction</i></b></p> <p>Students pose their own questions of interest</p> <p>Questions seek generalization</p>
<b>II. Collect Data</b>	<p>Do not yet <i>design for differences</i></p> <p>Census of classroom</p> <p>Simple experiment</p>	<p>Beginning awareness of <i>design for differences</i></p> <p>Sample survey; begin to use random selection</p> <p>Comparative experiment; begin to use random allocation</p>	<p>Students make <i>design for differences</i></p> <p>Sampling designs with random selection</p> <p>Experimental designs with randomization</p>
<b>III. Analyze Data</b>	<p>Use particular properties of <i>distributions</i> in the context of a specific sample</p> <p>Display variability within a group</p> <p>Compare individual to individual</p> <p>Compare individual to group</p>	<p>Learn to <i>use</i> particular properties of <i>distributions</i> as tools of analysis</p> <p>Quantify variability within a group</p> <p>Compare group to group in displays</p>	<p>Understand and <i>use distributions</i> in analysis as a global concept</p> <p>Measure variability within a group; measure variability between groups</p> <p>Compare group to group using displays and measures of variability</p>

Table 9.1 2005 GAISE Curriculum Framework

Appendix B (cont'd)

	Beginning awareness of group to group	Acknowledge sampling error	Describe and quantify sampling error
	Observe association between two variables	Some quantification of association; simple models for association	Quantification of association; fitting of models for association
<b>IV. Interpret Results</b>	Students do not look <i>beyond the data</i>	Students acknowledge that <i>looking beyond the data</i> is feasible	Students are able to <i>look beyond the data</i> in some contexts
	No generalization beyond the classroom	Acknowledge that a sample may or may not be representative of the larger population	Generalize from sample to population
	Note difference between two individuals with different conditions	Note the difference between two groups with different conditions	Aware of the effect of randomization on the results of experiments
	Observe association in displays	Aware of distinction between observational study and experiment	Understand the difference between observational studies and experiments
		Note differences in strength of association	Interpret measures of strength of association
		Basic interpretation of models for association	Interpret models of association
		Aware of the distinction between association and cause and effect	Distinguish between conclusions from association studies and experiments

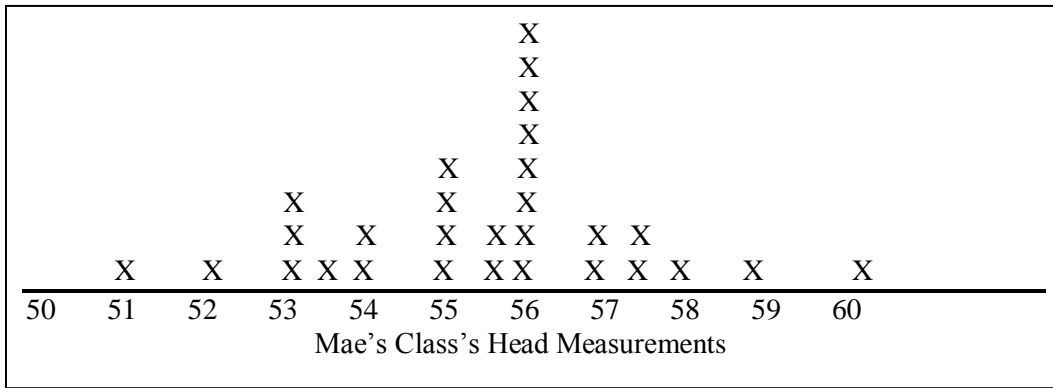
APPENDIX B (cont'd)

<b>Nature of Variability</b>	Measurement variability	Sampling variability	Chance variability
	Natural variability		
	Induced variability		
<b>Focus on Variability</b>	Variability within a group	Variability within a group and variability between groups	Variability in model fitting
		Covariability	

*Table 9.1 2005 GAISE Curriculum Framework*

*Source:* Franklin, C., Kadar, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria, VA: American Statistical Association.

# APPENDIX C



*Figure 9.2 Mae's Class's Head Measurements*

## Bibliography

## BIBLIOGRAPHY

- Aczel A. (1996). *Complete business statistics*. Chicago, Robert. D. Irwin.
- Bakker, A., Gravemeijer, K. J. & Pollack. E. (2004). Learning to reason about distribution. In J. Garfield & D. Ben Zvi (Eds). *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147–168). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Ball, D. L., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: the unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed). *Handbook of research on teaching* (pp. 433–456) New York: McMillan.
- Begle, E. G. (1979). *Critical variables in mathematics education: Findings from a survey of the empirical literature*. Washington, DC: Mathematical Association of America and National Council of Teachers of Mathematics.
- Canada, D. (2004). Preservice elementary teachers conceptions of variability. Unpublished Doctoral Dissertation, Portland State University, Portland, OR.
- College Board (2006). College Board standards for college success: Mathematics and statistics. Retrieved from <http://www.collegeboard.com>
- Common Core State Standards Initiative (2010). The standards: Mathematics. Retrieved from <http://www.corestandards.org>
- Connected Mathematics Project* (1995). Palo Alto: Dale Seymour Publications.
- delMas, R.C. (2002). Statistical literacy, reasoning, and learning: a commentary. *Journal of Statistics Education*, 10 (3). Retrieved from [http://www.amstat.org/publications/jse/v10n3/delmas\\_discussion.html](http://www.amstat.org/publications/jse/v10n3/delmas_discussion.html)
- Drake, C. L. (2006, Spring). Turning points: Using teachers' mathematics life stories to understand the implementation of mathematics education reform. *Journal of Mathematics Teacher Education* 9, 579–608.
- Drake, C. L. & Sherin, M. G. (2006). Practicing change: Curriculum adaption and teacher narrative in the context of mathematics education reform. *Curriculum Inquiry* 36 (2), 153–187.
- Franklin, C., Kadar, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria, VA: American Statistical Association.

- Friel, S. N., Curcio, F., & Bright, G. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education* 32 (2), 124–158.
- Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education* 10 (3), Retrieved from <http://www.amstat.org/publications/jse/v10n3/garfiled.html>
- Grossman, P. L., Wilson, S. M., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. In M. Reynolds (Ed.), *The knowledge base for beginning teachers* (pp. 23–36). New York: Pergamon.
- Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics* 11, 1–21.
- Hammerman, J., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal*, 3 (2), 17–41.
- Jacobs, J.K., & Morita, E. (2002) Japanese and American teachers' evaluation of videotaped mathematics lessons. *Journal for Research in Mathematics Education*, 33 (3), 154–175.
- Konold, C. & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W. Martin & D. Schifter (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp.193–215). Reston: National Council of Teachers of Mathematics.
- Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education*, 33 (4), 259–289.
- Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A., Wing, R., & Mayr., S. (2002). Students' use of modal clumps to summarize data. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics, Cape Town, South Africa* [CD-ROM], Voorburg, The Netherlands: International Statistical Institute.
- Lappan, G., Fey J., Fitzgerald W., Friel S, Phillips E. (2002). *Getting to know connected mathematics: An implementation guide*. Upper Saddle: Prentice Hall.
- Lappan, G., Fey J., Fitzgerald W., Friel S, Phillips E. (2006). *Data distributions: Describing variability and comparing groups*. Upper Saddle: Pearson.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah: Lawrence Erlbaum.

- Makar, K., & Confrey, J. (2004). Secondary teachers' statistical reasoning in comparing two groups. In J. Garfield and D. Ben-Zvi. (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 353–373). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Makar, K. & Confrey, J. (2005). 'Variation talk': articulating meaning in statistics. *Statistics Education Research Journal*, 4 (1), 27–54.
- McGraw-Hill Staff. (2008). *Quick Review Math Handbook* (1st ed., Book 1, Student ed.). Columbus, OH: Glencoe/McGraw-Hill.
- Mickelson, W. T., & Heaton, R. M. (2004). *Primary Teachers' Statistical Reasoning About Data* (D. Ben-Zvi, Ed.). Dordrecht: Kluwer Academic Publishers.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13 (2), 125–145.
- Moore, D. S. (1990). *Uncertainty. On the Shoulders of Giants: New Approaches to Numeracy* (L. A. Steen, Ed.). Washington, D.C.: National Academy Press.
- Moore, D. S. (2000). *The basic practice of statistics*. New York: W. H. Freeman.
- National Council of Teachers of Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston: Author.
- National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston: Author.
- Outlier. (n.d.). In *Wikipedia the free online encyclopedia*. Retrieved from <http://en.wikipedia.org/wiki/Outlier>
- Pfannkuch, M., & Wild, C. J. (2000). Statistical thinking and statistical practice: Themes gleaned from professional statisticians. *Statistical Science*, 15, 132–152.
- Pfannkuch, M., & Wild, C. J. (2004). Towards an understanding of statistical thinking. In J. Garfield & D. Ben Zvi (Eds). *The challenge of developing statistical literacy, reasoning and thinking* (pp. 17–46). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Reading, C. & Reid, J. (2006). An emerging hierarchy of reasoning about distribution: From a variation perspective. *Statistics Education Research Journal*, 5(2), 46–68, <http://www.stat.auckland.ac.nz/serj>

- Shaughnessy, M. (1992). Research in probability and statistics: reflections and directions. In D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465–494). New York: Macmillan.
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *The Second Handbook of Research on Mathematics* (pp. 957–1010). Reston: National Council of Teachers of Mathematics (NCTM).
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman., R. Pea, B. Barron, & S. Derry (Eds.) R. Goldman, R. Goldman, R. Pea, B. Barron, & S. Derry, *Video research in the learning sciences* (pp. 383–395). Hillsdale, NJ: Erlbaum.
- Sherin, M. G., & Van Es, E. (2009). Effects of Video Club Participation on Teachers' Professional Vision. *Journal of Teacher Education*, 60 (1), 20–37.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), pp. 4–14.
- Shulman, L. S. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review*, 57, 1–22.
- Skemp, R. R. (1987). *The psychology of learning mathematics*. Hilldale: Lawrence Erlbaum.
- Snee, R. (1990). Statistical thinking and its contribution to total quality. *The American Statistician*, 44 (2), 116–121.
- Sorto, M. A. (2004). *Prospective middle school teachers' knowledge about data analysis and its application to teaching*. Unpublished doctoral dissertation, Michigan State University, East Lansing, MI.
- Tversky, A. & Kahneman, D. (1973). Availability: a heuristic of judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- U.S. Department of Education's Mathematics and Science Expert Panel. (1999). *Exemplary & Promising Mathematics Programs*. Jessup, MD: U.S. Department of Education.
- Watkins, A. E., Schaeffer, R., & Cobb, G. (2003). *Statistics in action: Understanding a world of data*. Emeryville: Key Curriculum Press.
- Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks: Sage Publications.