

ASSESSMENT PRACTICES OF MATHEMATICS TEACHERS WHO ALSO TEACH
AP STATISTICS

by

SHADRECK SONES CHITSONGA

(Under the Direction of Jeremy Kilpatrick)

ABSTRACT

The purpose of this study was to investigate the assessment practices of mathematics teachers who also teach AP Statistics. My focus was to explore the teachers' conceptions of assessment in mathematics and AP Statistics, the way that the teachers carried out assessment in the two courses, and the factors that influenced the teachers' classroom assessment practices in the two kinds of courses. Though there has been a lot of research on assessment practices, there has not been a study that has looked at the assessment practices of teachers in both mathematics and AP Statistics. For example, we do not know whether the way teachers view assessment in mathematics is the same as the way they view assessment in AP Statistics.

Two teachers who were teaching mathematics and AP Statistics participated in this study from November 2008 to February 2009. Each teacher was interviewed two times for about 90 minutes each. Interviews focused on the teachers' conceptions of mathematics, statistics and assessment. Each teacher was observed 12 times in class. Artifacts were also collected from each teacher. Three frameworks were used to analyze the data—a theory of intellectual development, a mathematics taxonomy framework, and

the six assessment standards by (NCTM, 1995). Data were analyzed by using an analytic induction method.

The results of the study indicated that there was no difference between the teachers' conceptions of assessment in mathematics and AP Statistics. The main aim of assessment was to promote students' learning. Two forms of classroom assessment were identified— formal and informal assessment. Formal assessment involved the use of homework, quizzes, and tests. Informal assessment involved oral questions and listening to students' conversations. Assessment questions (oral and written) in AP Statistics and mathematics mostly assessed recall of information, comprehension of information and routine use of procedures. Factors that influenced the teacher's assessment practices included time, textbook publishers' tests, and teachers' pedagogical content knowledge. The findings in the study suggest that there is need to bring awareness to the teachers to think about the types of assessments they use and the type of thinking skills that they assess.

INDEX WORDS: Mathematics, AP Statistics, Teacher conceptions, Formal assessment, Informal assessment, Assessment tasks, Thinking skills

ASSESSMENT PRACTICES OF MATHEMATICS TEACHERS WHO ALSO TEACH
AP STATISTICS

by

SHADRECK SONES CHITSONGA

B. Ed., University of Malawi, Malawi, 1991

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILISOPHY

ATHENS, GEORGIA

2010

© 2010

Shadreck Chitsonga

All Rights Reserved

ASSESSMENT PRACTICES OF MATHEMATICS TEACHERS WHO ALSO TEACH
AP STATISTICS

by

SHADRECK SONES CHITSONGA

Major Professor: Jeremy Kilpatrick

Committee: Nicholas Oppong
 Patricia S. Wilson

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2010

DEDICATION

To my brother Henry and my sister Linley who never lived to see this day.

ACKNOWLEDGEMENTS

My sincere gratitude goes to Mary and John who accepted to participate in this study. Without them this dissertation would not have been possible. I thank them for sacrificing their time and for allowing me in their classrooms.

My major professor Dr. Kilpatrick spent countless hours reading my drafts, editing and offering me insightful suggestions. I do not know how to thank him. I have learned a lot working with him. Without him I would not have been able to complete this dissertation. I thank him for his patience, thoughtfulness and words of encouragement.

I express my gratitude to the members of my committee Dr. Oppong, who was also my academic advisor and Dr. Patricia S. Wilson. I thank them for their time, their helpful suggestions and support.

To my parents, Sones and Fainess Chitsonga, all I can say is that your words of encouragement have carried me through. Thank you for believing in me. Now, I can say, it is finally over. Thanks to all my brothers and my only sister Felistas, for all their support.

Finally, I would like to thank my lovely lady, my dear friend, Muta. She has been there with me throughout the whole process. She lifted my spirits when I seemed to doubt myself.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	x
LIST OF FIGURES.....	xi
CHAPTER	
1 INTRODUCTION AND BACKGROUND.....	1
Background	1
Difference Between Mathematics and Statistics	3
Rationale.....	4
Research Questions	9
Overview of the Report.....	10
2 REVIEW OF RELEVANT LITERATURE	11
Teachers' Conceptions of Mathematics	11
Teachers' Conceptions of Assessment.....	13
Characteristics of Assessment Items	14
Assessment Instruments and Grading	16
Factors Influencing Assessment and Grading Practices	17
Theoretical Framework	20
Teachers' Conceptions of Mathematics and AP Statistics.....	20

Teachers' Assessment Practices	24
3 DESCRIPTION OF THE STUDY	28
Participants	28
Description of the Schools	30
Procedure	33
Data Collection Procedures	35
My Position in the Study	37
Phenomenology	38
Data Analysis	39
4 THE CASE OF JOHN	42
General Background	42
John's Conceptions	44
John's Instructional Practices in Mathematics	52
John's Assessment Practices in Mathematics	54
John's Instructional Practices in AP Statistics	62
John's Assessment Practices in AP Statistics	65
Factors Influencing Assessment in Mathematics	71
Factors Influencing Assessment in AP Statistics	75
Summary of John	79
5 THE CASE OF MARY	81
General Background	81
Mary's Conceptions	83
Mary's Instructional Practices in Mathematics	92

Mary's Assessment Practices in Mathematics	95
Mary's Instructional Practices in AP Statistics	102
Mary's Assessment Practices in AP Statistics	104
Factors Influencing Assessment in Mathematics	111
Factors Influencing Assessment in AP Statistics	114
Summary of Mary	120
6 CROSS-CASE ANALYSIS	121
Conceptions of Assessment	121
Assessment Practices	125
Overall Analysis of Assessment Practices	131
Mathematics Standard	132
Learning Standard	134
Equity Standard	136
Openness Standard	139
Inferences Standard	141
Coherence Standard	144
Factors Influencing Assessment	146
7 SUMMARY AND CONCLUSIONS	151
Research Related to This Study	152
Teachers' Conceptions of Mathematics and AP Statistics	152
Teachers' Conceptions of Assessment	153
Characteristics of Assessment Items	154
Assessment Instruments and Grading	156

Factors Influencing Assessment and Grading Practices	156
Limitations of the Study.....	158
Implications for Teacher Education Programs.....	159
Recommendations for Future Research	163
Closing Remarks	166
REFERENCES.....	171
APPENDICES	
A INTERVIEW PROTOCOLS	176
B EXAMPLES OF MATHEMATICS AND AP STATISTICS ASSESSMENT TASKS FOR EACH TAXONOMY	179

LIST OF TABLES

	Page
Table 1: Taxonomy of Student Understanding in Mathematics	40
Table 2: Percent of John's Oral and Written Assessment Items Assessing Thinking Skills in Mathematics	60
Table 3: Grading Schedule for Mathematics.....	61
Table 4: Percent of John's Oral and Written Assessment Items Assessing Thinking Skills in AP Statistics	70
Table 5: Percent Distribution of Grades for AP Statistics	71
Table 6: Percent of Mary's Oral and Written Assessment Items Assessing Thinking Skills in Mathematics	101
Table 7: Percent Distribution of Grades for Honors Algebra II.....	102
Table 8: Percent of Mary's Oral and Written Assessment Items Assessing Thinking Skills in AP Statistics	110
Table 9: Percent Distribution of Grades for AP Statistics	111
Table 10: Number of Quizzes and Tests Used by Mary and John in Mathematics and AP Statistics	130

LIST OF FIGURES

	Page
Figure 1: Topics of Lessons During Classroom Observations.....	34
Figure 2: Sample Homework Problems Similar to Those John Assigned in His Precalculus Course.....	57
Figure 3: Sample Problems From the Ellipse Quiz Taken From John’s Test Bank	58
Figure 4: Sample Test Problems for John’s Analysis Course Taken From John’s Test Bank... ..	59
Figure 5: Homework Problem Similar to the One that John Assigned.....	63
Figure 6: John’s Drawing of Normal Curve.....	64
Figure 7: Sample Homework Problems Similar to Those John Assigned for His AP Statistics Class.....	67
Figure 8: Sample Quiz Problems Similar to Those John Used for His AP Statistics Class	68
Figure 9: Sample Test Problems Similar to Those John Used for His AP Statistics Class	69
Figure 10: Example Problem on Functions Similar to the One Discussed on January 14	95
Figure 11: Sample Homework Problems Similar to Those Mary Assigned in Honors’ Algebra II.....	98
Figure 12: Sample of Quiz Questions Similar to Those Mary Used in Honors Algebra II	99
Figure 13: Sample Test Questions Similar to Those Mary Used in Honors Algebra II .	100
Figure 14: Example of a Problem on Finding Probability Distribution.....	103

Figure 15: Probability Distribution of Number of Heads in Three Tosses of a Coin	103
Figure 16: Sample Homework Problems Similar to Those From Mary's AP Statistics Class.....	107
Figure 17: Quiz Problems Similar to Those Mary Used in AP Statistics	108
Figure 18: Sample Test Problems From Mary's AP Statistics Class.....	109
Figure 19: Example on the Use of the Rule of Means and Variances	119

CHAPTER 1

INTRODUCTION AND BACKGROUND

This report is about a study of two high school teachers of mathematics who also taught Advanced Placement Statistics (AP Statistics)¹ and their assessment practices in the two kinds of courses. The two teachers were very experienced in the teaching of mathematics. In teaching AP Statistics, however, one teacher was much more experienced than the other. To understand the assessment practices of these teachers, I interviewed them, observed them in class, and also collected artifacts (assessment documents).

Background

The inclusion of data analysis and probability in the National Council of Teachers of Mathematics (NCTM) 1989 and 2000 standards documents has been a welcome development insofar as the recognition of the importance of statistics in the PreK–12 curriculum is concerned. Apart from the NCTM, a number of organizations have undertaken projects to improve the teaching of statistics in Grades PreK–12. Among those is the project on Guidelines for Assessment and Instruction in Statistics Education (GAISE; Franklin et al., 2007), which has been recognized as a welcome development by the American Statistical Association (ASA, 2005).

¹ The Advanced Placement Program offers a course description and exam in statistics to secondary school students who wish to complete studies equivalent to a one-semester, introductory, non-calculus-based, college course in statistics (College Board, 2007, p. 3).

Endorsing GAISE, the ASA said:

The ASA Board of Directors appreciates and supports the efforts of the Pre K–12 GAISE writing group and endorses its recommendations for Pre K–12 statistical education in the document, *A Curriculum Framework for Pre K–12 Statistics Education* as an aid to enhancing statistics education at the K–12 levels. (p. 1)

In June 2010, the National Governors Association (NGA) and the Council of Chief State Officers (CCSSO) launched Common Core State Standards (CCSS) for mathematics and English. So far, 37 states and territories and the District of Columbia have adopted these standards. Of particular interest to the present study are the Mathematics Standards. A number of prominent organizations have issued statements in support of the CCSS. For example, ASA has welcomed these standards, more specifically the statistics and probability standards:

Several prominent ASA members were involved in the writing and reviewing process of the statistics and probability standards, and several of their recommendations were incorporated into the final version. If adopted by the states, these standards would provide millions of school children across the U.S. with a greater understanding of statistics and better preparation for college-level courses. In addition, ASA recognizes the need for high-quality professional development in statistics and probability and encourages statisticians and statistics educators to work with state and local education departments on the implementation of the standards. (ASA, 2010, p. 1)

The National Council of Teachers of Mathematics (NCTM), the National Council of Supervisors of Mathematics (NCSM), the Association of State Supervisors of Mathematics (ASSM), and the Association of Mathematics Teacher Educators (AMTE), released a joint public statement on the common core standards in June 2010.

The release of the Common Core State Standards (CCSS) is a welcome milestone in the standards movement that began more than 20 years ago when the National Council of Teachers of Mathematics published *Curriculum and Evaluation Standards for School Mathematics*. By initiating the development of the CCSS, state leaders acknowledged that

common K–grade 8 and high school standards culminating in college and career readiness would offer better support for national improvement in mathematics achievement than our current system of individual state standards. (p.1)

Historically, statistics in many schools and colleges has been housed in the mathematics department. It is also true that most PreK–12 teachers of statistics are, first, mathematics teachers and that too few teachers in PreK–12 know statistics and how to teach it. The GAISE Report (Franklin et al., 2007) gives a clear picture of the status of statistics education in the secondary school in terms of the teachers' knowledge of statistics:

Statistics . . . is a relatively new subject for many teachers, who have not had an opportunity to develop sound knowledge of the principles and concepts of data analysis that they are not called upon to teach. These teachers do not clearly understand the difference between mathematics and statistics. (p. 5)

Difference Between Mathematics and Statistics

It is important to note that statistics does not originate from within mathematics. Statistics is concerned with gathering, organizing, and analyzing data, and with inferences from data to the underlying reality (Moore, 1988, p. 4). The distinction between mathematics and statistics may seem trivial. But the fact that some teachers cannot make the distinction between the two is a huge problem because those teachers teach statistics as a mathematical topic, putting a lot of emphasis on computations, formulas, and procedures (Gal & Garfield, 1997). A numerical answer is not sufficient in a statistics course until that answer is related back to the context, to the original question posed (Chance, 2002). Real statistics is less about the pursuit of the “correct” answer in some ideal sense than about doing the best one can within constraints (Wild & Pfannkuch, 1999). The knowledge needed for teaching statistics is different from the

knowledge needed for teaching mathematics. Groth (2007) recently attempted to provoke a debate on the conceptualization of statistical knowledge for teaching. He argued:

If an explicit consideration of the differences between mathematics and statistics is not undertaken as research on teachers' knowledge is carried out, then important differences between the knowledge needed for teaching each discipline are likely to be overlooked. In turn, the gap between school statistics and the actual discipline of statistics is likely to widen because of an implicit errant assumption that statistics is a branch of mathematics. (p. 428)

Groth hypothesized that statistical knowledge for teaching includes both specialized knowledge and common knowledge. These components are both mathematical and nonmathematical. A statistics teacher should activate both common knowledge and specialized knowledge for pedagogical activities, and mathematical and nonmathematical knowledge should be activated for statistical activities.

Rationale

Though there have been a lot of studies on assessment, there is still much to learn about teachers' assessment practices. Researchers agree that assessment remains one of the least understood parts of the teaching and learning process. Stiggins (2001) lamented the disregard of classroom assessment over the past century by policy makers, school leaders, and the measurement community. He said that this neglect has resulted in unacceptably low levels of assessment literacy among practicing teachers, resulting in inaccurate assessment of the achievement of many students and giving rise to ineffective feedback to students about their achievement as well as the failure of immense numbers of students to reach their full academic potential. Therefore, there is still a need to continue studying the assessment practices of teachers. Garfield (1994) observed that

most college teachers of statistics look at assessment in terms of testing and grading—scoring quizzes and exams—assigning course grades to students. She continues to say that using only traditional methods of assessment such as quizzes and tests “rarely reveals information about how students actually understand and can reason with statistical ideas or apply their knowledge to solving statistical problems” (p. 1).

Though there is plenty of literature on assessment practices of secondary school teachers of mathematics, there is no study that has focused on the secondary teachers of mathematics who also teach statistics. The fact that many teachers of mathematics are being called upon to teach statistics means that studies related to the teaching of statistics deserve serious consideration. More and more statistics topics are finding their way into the K–12 mathematics curriculum as calls to integrate statistics with other subject areas become more numerous. The GAISE framework suggests the following: “The traditional mathematics strands of algebra, functions, geometry, and measurement all can be developed with the use of data. Making sense of data should be an integrated part of the mathematics curriculum, starting in pre-kindergarten” (Franklin et al., 2007, p. 35). With all these developments, it is also imperative to examine the assessment practices of these teachers who are teaching both mathematics and statistics. Specifically, the present study looked at teachers of mathematics who are also teaching AP Statistics. Though studying the assessment practices of mathematics teachers who also teach statistics would have been a better option, circumstances did not allow for such a study. Instead the current study looked at the assessment practices of the teachers who taught advanced mathematics courses (Precalculus and Honors Algebra II) and AP Statistics.

Honors Algebra II is an algebra course that is offered to students who have successfully completed an Algebra I course. School districts require that students who want to enroll for Honors Algebra II meet the state requirements for gifted students classification. The Honors Algebra II is a more rigorous curriculum than the ordinary Algebra II. Analysis (Precalculus) is taught in some states as an honors or advanced version of advanced Algebra and Trigonometry. In the present study, this course was taught as an elective and it was referred to as Precalculus. AP Statistics is different from the ordinary statistics that is offered in most high schools. AP Statistics is offered by the College Board. Only students who have successfully completed Honors Algebra II or Algebra II with teacher recommendation are allowed to take the course. Students who enroll for AP Statistics are expected to take an examination that is offered by the College Board in May of every year. Most school districts require teachers to attend workshops and or summer courses and workshops/institutes offered by the College Board before they can teach AP Statistics. The popularity of AP Statistics has been in ascendancy from the time the first AP Statistics examination was given in 1997.

Realizing the need to develop new assessment strategies and practices that would “enable teachers and others to assess students’ performance in the manner that reflects the NCTM’s vision for school mathematics” (NCTM, 1995, p. 1), NCTM published the *Assessment Standards for School Mathematics*. There are six standards pertaining to mathematics assessment:

1. THE MATHEMATICS STANDARD: *Assessment should reflect the mathematics that all students need to know and be able to do.*
2. THE LEARNING STANDARD: *Assessment should enhance mathematics learning.*
3. THE EQUITY STANDARD: *Assessment should promote equity.*

4. THE OPENNESS STANDARD: *Assessment should be an open process.*
5. THE INFERENCE STANDARD: *Assessment should provide valid inferences about mathematics learning.*
6. THE COHERENCE STANDARD: *Assessment should be a coherent process.* (pp. 11–21)

NCTM defines assessment as “the process of gathering evidence about the student’s knowledge of, ability to use, and disposition toward, mathematics and making inferences from that evidence for a variety of purposes” (p. 3). For the present study, I used this definition of *assessment*. Though there are a number of uses of assessment, NCTM put the purposes of assessment into four broad categories: monitoring students’ progress towards learning goals, making instructional decisions, evaluating students’ progress at a particular time, and evaluating programs. In brief, the purposes of assessment suggest that there is more to assessment than just giving students tests or quizzes for a grade.

There have been calls (NCTM, 1995; Stiggins, 1998) to integrate assessment with instruction so that assessment will be more effective. There have been major shifts in the way assessment is perceived. NCTM (p. 83) outlined some of the changes as follows: (a) Assessment was treated as independent of curriculum, but now assessment is aligned with instruction and curriculum; (b) students were viewed as objects of assessment, but now students are viewed as active participants in the assessment process; and (c) assessment was regarded as sporadic and conclusive, but now it is regarded as continual and recursive. Gronlund (2006) comments on the relationship between assessment and instruction: “The typical procedure of limiting instructional planning to the teaching-learning process is inadequate. Effective instruction requires that we expand our concern to a teaching-learning-assessment process with assessment as a basic part of the instructional program” (p. 3). This statement underscores the close relationship between

assessment and instruction. Appropriate assessment needs to be incorporated into the learning process so that teachers and students can determine whether the learning goals are being achieved (Garfield, 1995).

The *Professional Standards for Teaching Mathematics* (NCTM, 1991) presented six standards for the teaching of mathematics, organized in four categories: (1) task, (2) discourse, (3) environment, and (4) analysis. Though all these categories are related in one way or another to classroom assessment, one that clearly stands out in relation to the present study is tasks. It is the responsibility of teachers to choose quality mathematics tasks that students can engage in (NCTM, 1991). The Task Standard defines good tasks as

ones that do not separate mathematical thinking from mathematical concepts or skills that capture students' curiosity and that invite them to speculate and to pursue their hunches. Many of such tasks can be approached in more than one interesting and legitimate way; some have more than one reasonable solution. These tasks, consequently, facilitate significant classroom discourse, for they require that students reason about different outcomes, weigh the pros and cons of alternatives, and pursue particular paths. (p. 25)

Tasks are important in the learning of mathematics—they send messages to students about what mathematics is and what is involved in doing mathematics (NCTM, 1991).

Some researchers have commented on the on the creation of assessment tasks. For example, Senk, Beckmann, and Thompson (1997) have proposed that

future research investigate effective models for preservice and in-service education to help teachers address issues in assessment and that future curriculum development efforts include the creation of assessment tasks. In particular, teachers need examples of worthwhile assessment tasks geared to specific courses they are teaching, rather than examples that are meant to assess general levels of mathematical performance. (p. 210)

From the discussion above, it is clearly evident that the choice of assessment tasks demands a lot of attention and consideration from teachers of mathematics. It is also an area in which a lot of teachers need help. Stiggins (2001) reported that teachers believed in the importance of teaching their students to think critically and that thinking can be taught. However, the teachers in Stiggins's study were not sure about how to assess thinking skills. Some teachers have resorted to using published assessment materials because they consider them to be superior to any they can create themselves (Senk et al., 1997). Among other things, the present study examined the type of tasks teachers use in their classrooms. Specifically, I looked at the thinking skills that those tasks assess. Do teachers create their own assessment tasks?

Choosing or adopting a particular assessment strategy is a decision-making process on the part of the teacher. The literature on assessment observes that there are a number of factors that influence teachers' assessment practices. Although I knew some of the factors (based on the literature) that influence teachers' assessment practices, I thought that it was important to explore those factors with the teachers. I know that there is no documentation of the factors that influence the assessment practices of mathematics teachers who also teach AP Statistics.

Research Questions

This was a qualitative study based on phenomenology as the theoretical perspective. In particular, I deemed a case study appropriate for the purpose of the study, which was to investigate the assessment practices that are carried out in mathematics and AP Statistics classrooms. In particular, I wanted to know the teachers' conceptions of assessment, their assessment methods, and also the factors that influenced their

assessment practices. To explore those ideas, I also wanted to know the teachers' conceptions of mathematics and statistics. In the study, I addressed three major research questions:

1. What are teachers' conceptions of assessment in mathematics and AP Statistics?
2. How do teachers assess students in mathematics and AP Statistics?
3. What factors influence teachers' assessment practices in mathematics and AP Statistics?

Overview of the Report

In chapter 2, I review relevant literature, the Perry (1970) theory and intellectual development and the six assessment standards (NCTM, 1995). Included in the review are topics such as teachers' conceptions of mathematics and assessment, characteristics of assessment items, and factors influencing teachers' assessment practices. The description of the study, which includes the methods and also the theoretical perspective that guided the study, is provided in chapter 3. In chapters 4 and 5, I present case studies of the two participants in the study, and in chapter 6, I present a cross-case analysis. Finally, chapter 7 gives a summary and conclusions; it also includes implications and recommendations for future research.

CHAPTER 2

REVIEW OF RELEVANT LITERATURE

In this chapter, I present research literature related to the present study.

Specifically, I look at studies on teachers' conceptions of mathematics and assessment, teachers' assessment practices (e.g., assessment items, assessment instruments, and uses of assessment) and also the factors that influence teachers' assessment and grading practices. The literature I review is dominated by mathematics because there have not been any studies on the assessment practices of high school teachers of mathematics teachers who also teach AP Statistics.

Teachers' Conceptions of Mathematics

Thompson (1984) investigated the conceptions of mathematics and mathematics teaching of three junior high school teachers. One of the questions she examined was whether what teachers said were their beliefs, views, and preferences about mathematics and mathematics teaching matched what they did in class. She reported that there was some consistency between the teachers' "professed" conceptions of mathematics and the way they presented the mathematics content. Though Thompson acknowledged that the study of the relationship between the teachers' conceptions of mathematics and mathematics teaching is complex—and cautioned against "making conclusive statements"—she indicated that "teachers' beliefs, views, and preferences about mathematics and its teaching, regardless of whether they are consciously or

unconsciously held, play a significant, albeit subtle role in shaping the teachers' characteristic patterns of instructional behavior" (p. 125).

Kesler (1986) conducted a similar study with four algebra teachers. He investigated the relationship of teachers' conceptions of mathematics and level of dogmatism to their instructional behavior. To understand the teachers' conceptions of mathematical knowledge, he used Perry's (1970) scheme of intellectual and ethical development and the application of Perry's scheme to mathematical knowledge by Copes (1979). Findings of the study indicated that some of the teachers' conceptions of mathematics and teaching were related to their instructional behavior. For example, one teacher defined *mathematics* as "a study of numbers and symbols coupled with a set of rules and procedures for performing certain operations on them to obtain the correct answers to problems" (p. 156). Kesler described such a conception of mathematics as dualistic. Classroom observations of this teacher indicated that he exhibited the following instructional behaviors:

- a. Heavy emphasis on right versus wrong answers,
- b. Only one way to approach a problem (the textbook's way), and
- c. Mathematics, in this class, was either "basics or non-basics." (p.

157)

But also in the same study, Kesler reported that sometimes the teachers' conceptions of mathematics did not necessarily translate into corresponding instructional behavior. For example, there were some teachers whose conceptions of mathematics were multiplistic and relativistic, but the teachers "wandered back to the dualistic nature of mathematics by stressing the traditional importance of

correct answers, but not at the expense of students' creativity" (p. 159). This finding is not new (see Thompson, 1984). However, both studies indicate the importance of understanding the teachers' conceptions of mathematics and its relationship to instructional behavior. In most cases, there is a relationship between the two variables.

Teachers' Conceptions of Assessment

Brown (2004) studied the conceptions of primary school teachers and their managers using a 50-item Teachers' Conceptions of Assessment (COA-III) questionnaire.² Brown reported that the majority of the teachers agreed with the improvement conception, which states that "assessment is related to improvement of student learning and teachers' instruction" (p. 302) and the school accountability conception, which states that "assessment evaluates the quality of schools and teachers" (p. 302), but they disagreed with the view that assessment was irrelevant. Brown also examined the relationships between the different conceptions of assessment and reached the following conclusion:

When teachers think assessment is about Student Accountability, it is moderately likely they will also consider assessment to be Irrelevant, because it is bad for students or inaccurate, such that they can safely ignore it. It is possible that this conception is related to strong student-centered learning beliefs or humanistic curriculum or nurturing teaching beliefs. Teachers who conceive of assessment as Student Accountability are likely to have only a weak relationship to Improvement. In other words, assessment of students is likely to be Irrelevant when it is connected to Student Accountability but is more likely to be acceptable if it is related to Improvement of teaching and learning. (pp. 313–314)

² The questionnaire is based on the hypothesis that teacher's conceptions of assessment can be understood by examining how they agree or disagree with four purposes of assessment: (a) improvement of teaching and learning, (b) school accountability, (c) student accountability, or (d) treating assessment as irrelevant.

Susuwele-Banda (2005) reported mixed results on the connection between the way teachers perceive assessment and the way they carry out assessment. For example, one teacher perceived “classroom assessment as the tool that teachers use to inform teaching and learning” (p. 123). His lessons were “teacher-centered,” however, and there was no evidence that the lessons were integrated with assessment. Three out of four teachers who perceived assessments as only tests gave their students tests, and also their instructional practices were mainly teacher centered and did not integrate assessment into the lesson. Susuwele-Banda indicated: “This supports the findings by Mulhall and Taylor (1998) who reported that teachers may have theoretical knowledge but fail to translate it to practice” (p. 123). Despite reporting on this “mismatch” between what the teachers said they do and what they did in practice, Susuwele-Banda observed that for the majority of the teachers in his study, assessment mainly meant tests and examinations. Because of that conception of assessment, the teachers were limited in the number of methods and tools they could use to assess their students. The teachers used assessment mainly for the purpose of ranking students and not necessarily for diagnosing students’ problems or determining the capabilities of individual students.

Characteristics of Assessment Items

Stiggins, Griswold, and Wikelund (1989) examined the classroom assessment procedures of 36 teachers in Grades 2 to 12 to determine the extent to which the teachers in mathematics, science, social studies, and language arts measure higher order thinking skills in their respective subject areas. Stiggins et al. used a framework of thinking skills developed by Quellmalz (1985).³ They reported that in all the subjects but mathematics,

³ The framework includes five types of thinking skills: recall, analysis, comparison, inference, and evaluation. Recall is the lowest-level, and evaluation is the highest level.

the written assessment questions had a heavy reliance on recall items. In mathematics, the study found that 72 % of the items assessed inference, and only 19% of the items assessed recall. This result implies that only mathematics assessed high-level thinking skills. Stiggins et al. (1989) also observed that items assessing other thinking skills such as comparison and evaluation were rarely used. They suggested a reason that evaluation is largely ignored:

Evaluation may be viewed as difficult to address because there may be no right answer. Teachers may not feel secure enough in the subjects they teach to be able to guide and in fact evaluate answers to evaluation questions in terms of the strength of the defense provided for an opinion. But in general, both comparison and evaluation are very important thinking skills that need to be developed and assessed. (p. 243)

A similar observation was made by Sanchez (2002), who reported about a teacher, Todd, who was uncomfortable using open-ended assessment items because it was difficult for him to immediately determine whether a student's response to an open-ended item was correct. Consequently, he stopped discussing open-ended assessment items with his students.

One other interesting finding of the Stiggins et al. 1989 study was that there was a remarkable difference between oral questions and written questions in mathematics: Written assessments were predominantly inference, and oral questions were distributed evenly across recall, analysis, and inference.

Senk et al. (1997) studied the assessment and grading practices of high school teachers of mathematics in Algebra, Geometry, Advanced Algebra, Trigonometry, and Functions and Precalculus. They reported that in all the courses, the tests had a high percentage of low-level items (ranging from 53 to 90%). However, this finding is contrary to what Stiggins et al. (1989) found (as indicated above). Senk et al. also found

that there was not enough evidence to suggest that the test items required the students to use reasoning (justification, explanation, or proof): The mean usage was 5%, the highest percentage of the usage was by the teachers of Geometry courses, and there was no evidence that teachers of Algebra I or II courses used the items. Senk et al. (1997) also noted that the teachers seldom or never used any open-ended items on tests (percentages of open-ended items ranged from 0 to 10%). Sanchez (2002) reported a different result about the use of open ended assessments. However, she outlined the reasons why the teachers in her study used open-ended assessments:

The teachers in this study each used open-ended items that assessed higher-level thinking skills on every test. That was likely due to the teachers' participation in the assessment projects and the support they received from their school system for using open-ended items. Too, the system actually required that they use open-ended items on tests, so the teachers' orientation to authority also affected their use of open-ended items on tests. (p. 120)

Assessment Instruments and Grading

McMillan and Nash (2000) reported that teachers preferred assessments that allowed students to demonstrate their understanding of concepts instead of just concentrating on the answers. Such assessments were said to promote learning instead of just memorizing. Teachers preferred “constructed response” items to multiple-choice items. The teachers described their preferred assessments as follows:

The assessments where students actually have to show some work or write about something are more valuable for informing about how much students know. Because it is then you know they understand about the student than just grading a sheet of answers. (p. 13)

Though the teachers preferred these instruments, they sometimes had very little choice but to use multiple-choice items because of external pressures, such as those from parents.

Ohlsen (2007) studied NCTM members in nine U.S. states to investigate whether any relationships exist between the types of classroom assessments used in secondary mathematics classrooms and high-stakes state assessment programs. She reported that major exams and quizzes were the main choice of teachers as a student assessment instrument used to determine a student's semester grades. Senk et al. (1997) and Enderson (1995) reported a similar result about the teachers' dependence on written tests and quizzes to determine students' grades. Ohlsen also found that, despite the fact that teachers were familiar with the NCTM assessment standards, they did not often use oral presentations, essay questions, or team projects (from 40 to 50% of the teachers used these methods).

Factors Influencing Assessment and Grading Practices

Factors influencing teachers' assessment and grading practices have been studied quite a bit, sometimes yielding conflicting results. Textbook publisher's tests can have a significant influence on the tests teachers use as part of classroom assessment. Senk et al. (1997) found that of the 11 teachers who were asked about the source of their test questions, 6 indicated that they depended on the textbook publishers' tests. Even those teachers who wrote their own questions indicated that they still consulted the publishers' tests. However, this finding is contrary to what McMillan and Nash (2000) and Ohlsen (2007) found. They found that teachers preferred to write their own tests instead of depending on publishers' tests. The teachers in the Mc Millan et al. study indicated that the publisher's tests did not fit their local context. In a similar study, Stiggins (1985) reported that teachers had concerns about publishers' tests because of students' reactions and improvement of test use. Stiggins reported:

Those concerned about student reactions to published tests tended to view these tests as invalid, undependable, too long, and so on, and thus anticipated that the tests were not helpful to students. Those concerned about improving test use viewed published tests as time-consuming, not matching their instruction, failing to reflect true student characteristics, and generally not meeting important instructional needs such as identifying material to teach or reteach. (p. 278)

Senk et al. (1997) also reported that the majority of the teachers who used the textbook publishers' tests did not make any changes in them.

Senk et al. (1997) also reported the influence of standardized tests on mathematics teachers' teaching and testing: 75% of the teachers indicated that the standardized tests had very minimal or no influence on their teaching or testing. However, geometry teachers reported that standardized tests had a "very important influence" on some of the content they concentrated on. Though the teachers in the study said that they did not teach to the test, they paid special attention to questions that were typical of the questions asked in standardized tests.

McMillan and Nash (2000) and Sanchez (2002) found that teachers' knowledge and beliefs influenced their assessment practices. In her study, Sanchez observed that teachers had different rationales for the use of open-ended assessment items based on their beliefs. For example:

Robin believed that "open-ended items allow students who have a conceptual understanding of a topic but "miss all of these little details" to demonstrate their understanding. She asserted that not including open-ended items in assessment would be unfair for such students. She valued conceptual thinking and used open-ended items as a way to elicit that kind of understanding. (p. 100)

McMillan and Nash reported that teachers expressed a desire to make sure that students succeeded. McMillan and Nash described that as "pulling for students"—a situation in which some teachers gave multiple forms of tests to accommodate all their students. On

the issue of grading, one of the teachers said: “Everybody takes the quiz, but the way I record the grade is only the good grades. If you don’t get a B or better, then I do not record it” (p. 12). In general, the teachers said they were guided by their philosophy of teaching and learning on matters related to assessment.

A number of researchers have identified time as one of the factors that affects mathematics teachers’ assessment practices. Teachers indicated that “newer forms of assessment generally took, on average, twice as much time to prepare and twice as much time to grade as chapter tests” (Senk et al., 1997, p. 209). These “newer” forms of assessment might include assessments such as oral or written projects, or computer laboratory assignments. Sanchez (2002) reported that teachers did not have enough time to use open-ended assessments because they were time consuming. She also said that some teachers faced a dilemma between using open-ended assessments and finishing the curriculum. The teachers generally chose the latter.

Senk et al. (1997) reported that students’ attitudes can have an influence on the teacher’s assessment practices—students’ attitudes made it difficult for teachers to implement some assessment strategies. One teacher in the study talked about the issue of motivating students:

If you have better students who are self-motivated, you can have them working on things outside of class, special research projects, and computer work, and so on, I think that would be great. But we don't have that here. I'll tell you frankly, we have difficulty getting our kids to do anything out of the classroom. (p. 209)

The literature presented in this chapter is not exhaustive but suggests some of the issues facing classroom assessment today. There are many challenges to think about and a lot of questions to ask. For example, the literature has shown that teachers depend

heavily on written assessment instruments to determine students' grades. Has this situation changed since these studies were done? The literature explored here has demonstrated some conflicting results concerning, for example, teachers' dependence on textbook publishers' materials to form tests and their use of high-level assessment items. A number of factors have been identified that influence teachers' assessment practices. Are there others that might be identified?

In the present study, I explored some of the issues presented in this chapter. More importantly, I investigated the assessment practices of mathematics teachers who also teach AP Statistics. This review of literature provides a good basis for studying those practices.

Theoretical Framework

The literature has shown some connections between teachers' conceptions of mathematics and their instructional behavior, so I thought in my effort to understand the assessment practices of the teachers, it was important to look at their conceptions of mathematics, statistics, and assessment. I decided to use Perry's (1970) theory of intellectual and ethical development to help me understand the teachers' conceptions of mathematics and of AP Statistics. Because this study dealt with assessment, I thought it was important to examine the teachers' assessment practices using the *Assessment Standards for School Mathematics* (NCTM, 1995).

Teachers' Conceptions of Mathematics and AP Statistics

Though the Perry (1970) theory of intellectual and ethical development was not specifically developed for mathematics, Thomson (1992, p. 132) indicated that a number of researchers (Copes, 1979; Doherty, 1990; Helms, 1989; Kesler, 1986; McGalliard,

1983; Meyerson, 1978; Owens, 1978; Stonewater & Oprea, 1998) had used it or adaptations of it as a framework for analyzing and characterizing teachers' conceptions of mathematics. The Perry theory (or scheme) is a product of a study conducted by the staff of the Bureau of Study Counsel at Harvard College in 1953. The study was conducted "to document the experience of undergraduates in Harvard and Radcliffe over their four years of college" (Perry, 1999, p. 4). The way students addressed the challenges faced in such areas as academic work and the social life of the college seemed to "represent a coherent development in the forms in which the students functioned intellectually, in the forms in which they experienced values, and in the form in which they construed their world" (Perry, 1999, p. 9). In the scheme, the formal properties of the assumptions and expectancies a person holds at a given time regarding the nature of knowledge and value are referred to as *structures*. Such structures can be described as, for example, *dualistic*. The sequence of these structures can be called a "developmental pattern." According to the Perry study, the structures were arranged in a sequence of nine stages, commonly referred to as positions. These nine positions—(1) basic duality, (2) multiplicity pre-legitimate, (3) multiplicity subordinate, (4) multiplicity correlate or relative subordinate, (5) relativism correlate, competing, or diffuse, (6) commitment foreseen (7) initial commitment, (8) orientation in implications of , and (9) developing commitments--represent what is called "Perry's Scheme of Development." The nine positions are generally collapsed into four categories (Rapaport, 2006): In my analysis of the teachers' conceptions of mathematics and AP Statistics, I used those four categories. Below I present a brief overview of each of the nine positions.

Dualism

Basic duality. Basic duality makes the simplest set of assumptions about the nature of knowledge and values. In this position, the world of knowledge or conduct is divided into two distinct categories such as the division between the familiar world of “authority-right-we” against “illegitimate-wrong-others.” In educational terms this view allows for learning through memorization and hard work; learning is broken down into discrete items. The authority is the source of correct responses, answers, and procedures, and the teacher’s main duty is to teach these responses, answers, and procedures.

Multiplicity

Multiplicity prelegitimate. In the multiplicity prelegitimate position, there is a perception of diversity of opinion and uncertainty. As a departure from Position 1, the authority can present complexities with the sole purpose of getting people to “learn to think independently.” Answers are not just presented by the authority, but must be found by the learners themselves—authority is no longer viewed as the “source of all knowledge,” and it is also possible for the authority not to have all the answers. However, the authority and absolutes are always available, but one’s dependence on them is minimized.

Multiplicity subordinate. In the multiplicity subordinate position, there is an acceptance of diversity and certainty. Certainty and diversity are referred to as *multiplicity*. This legitimacy of multiplicity is only temporary, with the understanding that the authority has not yet found the answer. In relation to authority, one’s duty is to ask, What is it that “they want”? What is crucial is to pay more attention to what the authorities say.

Multiplicity correlate or relative subordinate. Instead of waiting for the time when the authority will bring forth solutions to the problems not yet answered in Position 3, individuals taking the multiplicity correlate or relative subordinate position can think about their own solutions, because “everybody has the right to his or her opinion.” Here there is an aura of independence of thought in which the capacity to compare different approaches kicks in.

Relativism.

Relativism correlate, competing, or diffuse. In the relativism correlate, competing, or diffuse position, all knowledge and values are viewed as relativistic. Perry (1999) considers Position 5, as the crucial stage in the development of values. He indicates that a “revolutionary structuring” takes place. In Position 4, multiplicity and relativism were viewed from the perspective of a dualistic world, whereas in Position 6; there is a beginning understanding of the implications of personal choice in a world that is relativistic. Though the authority is still there, there is a feeling that the authority might also be searching for answers in the relativistic world. In this position, one still has lingering doubts on what one truly believes. The belief is mainly shaped by the individual, however, and not necessarily “out of so much what people have been telling you” (Perry, 1999, p. 148).

Commitment

Commitment foreseen. Perry (1999) describes commitment foreseen, Position 6, as “the moment of realization” (p. 151). In this position, people choose those aspects of their life in which they want to invest their energies, care, and identity from “multiple

perspectives.” That is what Perry (1999) refers to as a personal commitment towards the relativistic world.

Initial commitment, orientation in implications of, and developing commitments.

In the three stages of initial commitment, orientation in implications of, and developing commitments, one takes the responsibility for the choices one has made in life. There is a commitment to one major area. Perry (1999) gives an example. Somebody can say, “I have decided on a career in medicine.” A number of things can contribute to the making of an initial commitment, such as one’s discovery of new interests. The person in this position experiences the “affirmation of identity” in a pluralistic society.

Teachers’ Assessment Practices

According to NCTM (1995), the six assessment standards can be used to critique the assessment of mathematics and statistics. For each standard, I examined two questions. I used these standards to help me answer the second research question (see p. 10). Next, I present a brief overview of these standards followed by the questions I examined.

The Mathematics Standard

Among other things, the mathematics standard discusses the need to consider mathematics that should be reflected in the assessment. For example, even though skills, procedural knowledge, and factual knowledge should be assessed, there is a need to include mathematical activities that provide students with the opportunity “to formulate problems, reason mathematically, make connections among mathematical ideas, and communicate about mathematics” (NCTM, 1995, p. 11). There is a need to consider the importance of mathematics in our society.

- What mathematics is reflected in the assessment?
- How does the assessment engage the students in realistic and worthwhile mathematical activities?

The Learning Standard

The learning standard emphasizes the importance of making sure that assessment is aligned with instruction. It also discusses the use of informal assessment such as observing and listening to students. Assessment should not only come at the end of instruction but should also be continuous. It is also important to consider assessments that “allow students to demonstrate what they know and what they can do in novel situations” (NCTM, 1995, p. 14).

- How does the assessment relate to instruction?
- How does the assessment allow students to demonstrate what they know and can do in novel situations?

The Equity Standard

The equity standard discusses the need to present all students with opportunities for them to demonstrate their “mathematical power.” Among other things, the standard indicates that the individual differences must be taken into account when thinking about assessments that will allow each student to demonstrate what he or she knows. The standard also states that “assessors should be open to alternative solutions” (NCTM, 1995, p. 15), taking into account students’ backgrounds and experiences.

- What opportunities has each student had to learn the mathematics being assessed?

- How does the assessment provide alternative activities or modes of response that invite each student to engage in the mathematics being assessed?

The Openness Standard

The openness standard talks about providing information to the students about the process of assessment—notifying them about what they need to know before they are formally assessed. Assessment process information should also be made available to the public. Finally, assessment should be reviewed continually and improved upon.

- How do students become familiar with the assessment process and with the purposes, performance criteria, and consequences of the assessment?
- How are teachers and students involved in choosing tasks, setting or creating criteria, and interpreting results?

The Inferences Standard

Apart from traditional assessment instruments like multiple-choice and short-answer tests, the inferences standards says that mathematics assessment should include other instruments such as “observations, interviews, open-ended tasks, extended problem situations, and portfolios” (NCTM, 1995, p. 19). These instruments are regarded as multiple sources of evidence that can help a teacher make a valid inference about students’ learning.

- What evidence about learning does the assessment provide?
- What multiple sources of evidence are used for making inferences, and how is the evidence used?

The Coherence Standard

According to NCTM (1995), the coherence standard “connects the other standards to assessment systems, assessment purposes, curriculum, and instruction” (p. 21). The standard outlines three assessment aspects: All phases of the assessment process must fit together; there must be a correspondence between assessment and its purposes; and that assessment must be aligned with the curriculum and instruction.

- How is professional judgment used to ensure that the various parts of the assessment process form a coherent whole?
- How does the assessment match its purposes with its uses?

In this chapter I reviewed literature related to assessment. I discussed the studies on teachers’ conceptions of mathematics and assessment, teachers’ assessment practices, and also the factors that influence teachers’ assessment and grading practices. The chapter also discussed the frameworks that I used to analyze the teachers’ conceptions of mathematics and statistics, and the teachers’ assessment practices.

CHAPTER 3

DESCRIPTION OF THE STUDY

This study consisted of two case studies. According to Romberg (1992), a case study enables the researcher to organize and report on information about the actions and perceptions of an individual or group under specific circumstances. Punch (1998) says, “Only the in-depth case study can provide understanding of a new or persistently problematic research area” (p. 156). I considered the case study design appropriate for this study because it gave me the opportunity to look closely at, understand, and report on the assessment practices of mathematics teachers who also teach AP Statistics. Furthermore, because there have not been any studies on assessment that focus on those mathematics teachers who also teach AP Statistics, a case study was ideal to gain a deeper understanding of the assessment practices of the teachers involved. With that in mind, I decided to make this case study interpretive. In an interpretive case study, “a researcher gathers as much information about a problem as possible with the intent of analyzing, interpreting, or theorizing about the phenomenon” (Patton, 1988, p. 38).

In this chapter, I describe the selection of participants in the study, the descriptions of participants’ schools, the data collection procedures, the theoretical perspective that guided the study, and the data analysis.

Participants

The participants in this study were two high school teachers named John and Mary (pseudonyms). At the time of the study, the participants were teaching both

mathematics and AP Statistics. Both were members of a local professional learning community⁴ for AP Statistics teachers, and as such they were also involved in a number of workshops aimed at improving the teaching of AP Statistics. These two participants provided the units of analysis for the study. The participants were identified based on purposeful sampling. This method is used when there are clear rationales or criteria for selecting the participants for the group to be studied or tracked (Champion, 2002). According to Patton (2002), the power of purposeful sampling lies in the selection of information rich cases. He says, “Studying information rich samples yields insights and in-depth understanding rather than empirical generalizations” (p. 230).

I consulted with the coordinators of the professional learning community for AP Statistics teachers to assist in the identification of participants who had experience in teaching both mathematics and AP Statistics. The teachers identified were members of this learning community for AP Statistics teachers. Their participation in the learning community was used as a criterion of selection for them to take part in this study. Once the participants were identified, I went to their schools to meet them and brief them about the study. They were both gracious and agreed to participate in the study. I also asked them if they could allow me to sit in their classes before beginning the study so that I could familiarize myself with the school environment. They agreed to do that, and I observed each teacher a number of times in his or her mathematics and AP Statistics classes. During that time I did not collect any data.

⁴ The AP Statistics professional learning community was a gathering of AP Statistics teachers in a Southeastern state. The main objective of the gathering was to help AP Statistics share ideas on the teaching of statistics.

Description of the Schools

The two schools used in the study were both in a Southeastern state county system. They were the only two public high schools in the county. The school year for both schools consisted of two 18-week terms, and the schools operated on a 4×4 modular or block schedule. All the classes were 90 minutes in duration.

John's School

The school had a population of slightly over 1000 students. About 90% of these students were White. The remaining 10% were Blacks, Asian, Hispanic, or American Indian/Alaskan. About 10% of the students were eligible for free lunches, and approximately 3% were eligible for a reduced-price lunch.

The school occupied one huge building in which all the classrooms were located. The mathematics department had its own section of the building for all the mathematics classes. All the mathematics classrooms were roughly identical in size. At the time of the study, John's school had 10 teachers of mathematics. The teachers shared a common staffroom where they did their planning and ate lunch together. John was the chairman of the mathematics department. His office was adjacent to the staffroom.

John's classroom was large enough to accommodate 30 desks and chairs. They were arranged in rows facing the green board that John used. Each student sat at his or her own desk, so when the students wanted to work together, they had to move the furniture around. There was another green board on one of the side walls. John rarely used that board. I saw him use it only when there was not enough space on the board in the front of the room. The room had one large bulletin board at the back of the classroom. That board was almost blank, with no posters. Another, much smaller

bulletin board was adjacent to the green board at the front of the room. Above that green board was a screen for an overhead projector.

John worked at three tables. One was at the front of the room, where he sat most of the time when he was using the overhead projector. The second table was across the room from the door to the classroom and was where John kept most of his teaching material as well as his desktop computer. The third desk, which was nearby, contained extra material such as a set of textbooks that students could use if they had not brought their textbook. Also available was a set of calculators and graphing utilities that students could borrow to use in class, though most had their own calculator. There was also a TV monitor, which was mainly used by the school administration to deliver messages to the students.

At the time of the study, John was teaching two classes of AP Statistics, one class of Mathematics Support I, and one class of Precalculus. I observed John in one AP Statistics class and his Precalculus class. John's AP Statistics class had 15 students, all of whom were White; there were 7 boys and 8 girls. John's Precalculus class had 19 students, 11 of whom were boys. Three of the students were Black, and the rest were White.

John taught all his classes in this room. The only time he used a different room was when he took his students to the computer laboratory. The lab was located a few meters away from his room but in a different section of the building. It was equipped with computers, about 20 terminals in all. It had also a digital overhead projector and a screen. During the time of the study, John took his AP Statistics students to the computer lab only once. I never saw him take his Precalculus students to the lab.

Mary's School

The school's student population was slightly less than 1000 students. Ninety percent of the students were White; the remaining 10% were Black, Asian, or Hispanic. The gender ratio in the student population was almost 1:1. Slightly more than 10% of the students were eligible for free lunch, and about 5% of the students were eligible for a reduced-price lunch.

The school building was a modern facility made up of four main blocks called *wings*, each of which had two floors that housed the classrooms. Mary's class was on the second floor. Adjacent to her classroom were two other mathematics classes. In fact, all the mathematics classes were located in this wing. Her class was directly opposite the computer laboratory.

Mary's classroom had 28 desks and chairs arranged in 4 rows, each with 7 desks and chairs. The room had two green boards on adjacent walls of the room. She used the first one for teaching and the second mainly to list the topics that would be covered in a given week. On the right hand side of the main board, Mary would put the dates of the tests that the students would take in a given week. That information applied to all the classes that Mary taught. The wall of the classroom opposite the board she taught from contained a bulletin board. On it were some posters about mathematics and AP Statistics. Some were printed, and others were done by either Mary or her students, displaying some mathematical concepts such as graphs.

Mary worked at two tables. The first one had a desktop computer that she used and was where she placed most of her teaching material. It was directly opposite the entrance to the room and was adjacent to the board she taught from. The other desk was

directly in front of the students' desks. Next to it was an overhead projector that Mary used most of the time while teaching. She also had some extra books that were used by the students if they had forgotten their textbook. These books were usually kept in the cabinet at the back of the room opposite the main board. Other materials in the room included a set of TI-83 graphing calculators and ordinary calculators that could be lent to students who did not have their own or had forgotten to bring theirs to class. There was also a class set of small boards that students could write on using markers. Students used the small boards to work out problems during group work. During the study, the classroom was undergoing a technology upgrade. Workers were in the process of putting a smart board in the classroom.

At the time of the study, Mary was teaching two classes of Honors Algebra II, one class of Mathematics Support I, and one class of AP Statistics. I observed Mary teaching one Honors Algebra II class and the AP Statistics class. Mary's AP Statistics class had 17 students, all of whom were White with an exception of one Black student. Eight of the AP Statistics students were girls; the Honors Algebra II classes had 18 students, 8 of whom were girls. All the students in Honors Algebra II were White.

Mary taught all her classes in the same room, with an exception of lessons that required the use of computers. The computer laboratory was equipped with about 20 computers. During the time of the study, Mary took her AP Statistics class to the computer laboratory once.

Procedure

The study was conducted over a 10- week period. Data collection began with interviews with the teachers, which were conducted in November and December 2008.

The interviews were followed by classroom observations that were conducted in Spring 2009. A detailed schedule of the observations is shown in Figure 1.

Date	John		Mary	
	AP Statistics	Precalculus	AP Statistics	Honors Algebra II
Jan 12	Sampling distribution			
Jan 13			Random variables (RV) introduction	
Jan 14	Inductive proof of Central Limit Theorem.	Rational functions		
Jan 15			Review of test on rules of probability	Solving inequalities
Jan 20			Means & variances of RVs	Operations with Functions
Jan 22			Means & variances of RVs	Quiz on functions; function notation
Jan 23	Introduction to Confidence intervals			
Jan 26			Quiz on RV and group quiz	Inverse functions
Jan 27	Confidence interval for population means	Complex numbers		
Jan 28			Review of graded quiz and extra problems on RV	Special functions
Jan 29	Confidence interval. for population proportion	Complex zeros:		
Feb 9		Parabolas		
Feb 10				Systems of linear inequalities
Feb 12	Z-test for population mean	Hyperbolas		
Feb 18		Hyperbolas (contd)		

Figure 1. Topics of Lessons During Classroom Observations

Data Collection Procedures

Data for the study were collected through multiple methods. Patton (1988) recommends the use of multiple methods of data collection because each data source has its strengths and weaknesses. He says that the use of multiple methods of data collection increases the reliability and the validity of the data collected. In this study, three methods were used to collect data: semi-structured personal interviews, observations, and collection of artifacts (documents).

The two participants were each interviewed two times. Punch (1998) comments on the interview as a method: “It is a very good way of accessing people’s perceptions, meanings, definitions of situations and constructive reality” (p. 175). He continues, “It is also one of the most powerful ways we have of understanding others” (p. 175). In the first interview, I asked the participants to give their background information, such as their educational background, with a special focus on their preparation to become teachers. The interview also dealt with the teachers’ conceptions about mathematics and AP Statistics. The second interview addressed issues related to assessment in both mathematics and AP Statistics. The questions asked in the interviews were the same for both teachers. The interview protocols appear in Appendix A. Each of the two interviews was more than 1 hour long. The interviews were recorded using a digital audio recorder. Additional informal interviews were also conducted at the end of each class observation whenever necessary—there were nine in total, four for John and five for Mary. These interviews were mainly to seek clarifications of observations made in class. These interviews were generally very brief, as immediately after the observation another

class was coming into the room, so there was not enough time to discuss many things.

Sometimes I obtained clarification through an exchange of emails with the participants.

I also made classroom observations. Though interviews are the primary source of data in qualitative research, Merriam (1998) cites two advantages of observations over interviews. She says:

First, observations take place in the natural field setting instead of a location designed for the purpose of interviewing; second, observational data represent a firsthand encounter with the phenomena of interest rather than the secondhand account of the world obtained in the interview. (p. 94)

I observed each teacher 12 times (6 times in the mathematics class and 6 times in the AP Statistics class). For convenience I observed each participant on the day that he or she taught both mathematics and AP Statistics. By coincidence, both participants taught mathematics and AP Statistics in consecutive periods. Each day, I planned to observe each participant in both classes. On certain days, I only observed one class. For example, on January 13, Mary's Honors Algebra II students were taking part in a mathematics competition, so she did not have a class on that day. Each observation lasted the entire teaching period, which was 90 minutes for both mathematics and AP Statistics. During the observations, I made field notes to record what I observed. Because the study concerned assessment, I paid particular attention to aspects of the lesson that I viewed as relevant to understanding the teacher's assessment practices. I observed the type of questions the teachers asked in class, the topics that they emphasized, how they responded to students' questions, the type of feedback they gave the students, and the type of mathematical or statistical knowledge their assessment practices conveyed. I also wanted to see if they used

such aspects of assessment as group work and how they conducted review for homework, quizzes, and tests. NCTM (2000) observes that there are many assessment techniques that mathematics teachers can use, such as open-ended questions, constructed-response tasks, portfolios, and performance tasks. So I also paid attention to the type of assessment instruments the teachers used in the classroom.

I also collected documents during the study. Merriam (1998) contends that that data in documents can be used in the same manner as data obtained from interviews or observations. Among other advantages of documents is that data obtained from them “can furnish descriptive information, verify emerging hypothesis, [and] advance new categories” (p. 126). The documents included the tests, quizzes, books, or other materials, such as worksheets that the teachers used in their classroom assessments.

My Position in the Study

Prior to the time I conducted the study, I had not met either of the two participants. The first time I met them was when I went to ask them to participate in the study, after exchanging a few email messages with them. Both teachers were forthcoming; however, I found Mary very outgoing, whereas John was a bit reserved. Before I started collecting data, I was simply an observer in John’s class, whereas in Mary’s class, she would sometimes ask me to comment on a few things or would let her students ask me questions if they were working on problems. Before I started collecting data, Mary invited me to attend one of the meetings of the AP Statistics learning community. I also helped her prepare a poster she was to use at a

presentation. My role did not change that much during the time I started collecting data—a complete observer in John’s class and an observer in Mary’s class, apart from one occasion when Mary asked me to participate in group work because one group did not have enough participants.

Phenomenology

The theoretical perspective for this study was phenomenology. I used the version of phenomenology that Crotty (1998) called “phenomenology as is usually presented today” (p. 83), contrasting it with earlier versions of phenomenology. I considered the daily activities that teachers engage in such as instruction, assessments, and interaction with their students as an “everyday” experiences for them. I considered their understanding or interpretations of these experiences to be subjective. From the phenomenology perspective, one is able to study “experience from the point of view or perspective of the subject.” Crotty looks at phenomenology as an “effort to identify, understand, describe, and maintain the subjective experiences of the respondents” (p. 83). Merriam and associates (2002) offer a similar definition: “The defining characteristic of phenomenological research is its focus on describing the ‘essence’ of a phenomenon from the perspectives of those who have experienced it” (p. 93). This approach allowed me to go into the study as a complete observer with the hope of identifying and understanding the teachers’ assessment practices from their perspective. One way of gathering data from the phenomenological perspective is by using open-ended interviews. For this study I used semi-structured interviews at the beginning of the study and also open-ended interviews after classroom observations.

Data Analysis

I analyzed the data using inductive analysis (Patton, 1988). “Inductive analysis means the patterns, themes, and categories come from the data rather than being imposed on them prior to data collection and analysis” (p. 306). Merriam (1998) contends that “data analysis is [an] interactive process throughout that allows the investigator to produce believable and trustworthy findings” (p. 151), and data collection and analysis occur simultaneously. I followed that strategy. The first step of analysis started immediately after the transcription of the interviews. I read the transcripts a number of times and started inserting comments in the documents. I then constructed a number of categories for each case based on the research questions and the theoretical framework. I did that by highlighting passages in the electronic documents using color coding for each category. Similar categories were highlighted in the same color. Any colored portions that contained repeated information were crossed out. The main categories that emerged from this analysis included the teachers’ conceptions of mathematics and AP Statistics, the teachers’ conceptions of assessment in mathematics and AP Statistics, and the factors that influenced the assessment practices in mathematics and AP Statistics. This initial analysis of the transcribed data was helpful during lesson observations. I used the analysis to check for consistency or inconsistency and to cross-validate the findings.

The second phase of analysis involved the examination of the assessment tasks that the teachers used in their mathematics and AP Statistics classes. These tasks comprised the oral questions that the teachers used in class and the written assessment tasks, which were mainly quiz and test items. These tasks were analyzed by using the

mathematics taxonomy (Table 1) developed by Smith, Wood, Compland, and Stephenson (1996), which is a modification of the 1956 Bloom taxonomy. The Smith et al. taxonomy was developed to make it compatible with the purpose of assessing students' understanding in mathematics. I chose to use this taxonomy for a number of reasons:

- It was specifically designed for analyzing mathematical tasks.
- The classification of the tasks in terms of thinking skills is in line with NCTM's (1989) *Curriculum and Evaluation Standards for School Mathematics*. For example, Standard 3 (mathematical reasoning) stipulates that students should “make and test conjectures, formulate counterexamples, [and] construct proofs for mathematical assertions, including indirect proofs and proofs by mathematical induction” (p. 143).
- Though it was constructed specifically for mathematics, the taxonomy is also applicable to statistics tasks. Of all the frameworks I looked at, this was the best for analyzing both mathematics and statistics tasks.

Table 1
Taxonomy of Student Understanding in Mathematics

Group A	Group B	Group C
Factual knowledge	Information transfer	Justifying and interpreting
Comprehension	Application to new situations	Implications, conjectures, and comparisons
Routine use of procedures		Evaluation

Note. From Smith et al. (1996, p. 67).

The taxonomy has three main groups—namely, A, B, and C—encompassing eight categories. Group A is the first level of thinking, and its tasks require students to

recall “previously learned information in the form it was given” (Smith et al., 1996, p. 68), or work with problems or tasks they have had plenty of practice with in class. Group B tasks ask students to use their knowledge beyond simple recall or working with routine problems. They require students to use their knowledge in unfamiliar or new situations. In Group C tasks, students must demonstrate abilities and skills like justifying, interpreting, comparing, and evaluating. Examples of the tasks for each group appear in Appendix B.

In the last phase of the analysis, I went back to the transcribed data and coded the data using the NCTM’s (1995) Assessment Standards as a framework. This analysis also involved the coding of the field notes from classroom observations. I compared the teachers’ statements from the transcript with what they did in class. This analysis helped me to answer my second research question.

During the first three phases of analysis, each case was treated separately. I then wrote the two cases. After that, I wrote a case study narrative for each participant and cross-analyzed the cases.

CHAPTER 4

THE CASE OF JOHN

General Background

John was a White male who at the time of the first interview was in his mid 50s and had taught mathematics in a Southeastern state for 17 years. He had a bachelor's degree in mathematics education from a university in a West Coast state. He had obtained a master's degree and a specialist degree in mathematics education from a Southeastern university in 1991 and 2003, respectively.

John was a family man. He had gone into teaching because he wanted to have time to spend with his family. He wanted to spend the vacation time with his children when schools closed. Prior to becoming a teacher, he had done several things like flying planes and sailing. Unlike his father, who had always been gone from home because of work, he wanted to be there for his children. That way, he would be in a good position to assist his children if they had any problems with their school work.

As an undergraduate, John had not started with mathematics education. He had initially started with a major in psychology, though he had always liked mathematics.

J: Oh, I think it was because when I went to a job fair, I was a junior in [the] psychology department [at a West Coast university]. And I went to a job fair, and it just scared me to death—what was available to science majors.

I: Oh, really?

J: Yeah, and I always liked math. So I changed my major to mathematics. And I moved from [one West Coast university to another] and finished my degree in math at the second. And I don't know why. I just wanted a degree. I didn't do anything. . . . What motivated me to mathematics was my fright at the possibilities of a job with a science degree.

I: OK. Were they scarce, or were they scary?

J: Well, it was, it was—. Well, to begin with, the pay was really low, and the status was pitiful. You know, [the] job status was just really down there. And it was all either working for government hospitals or government institutions; prisons, something like that. I mean just, the only people that—. I didn't really want to be a clinical psychologist, you know, but—. And so when you think about what else you can do, it had to do with prisons and its inner circles. (John interview, November 24, 2008)

John was very critical of the role the West Coast university had played in preparing him to become a teacher. He was particularly unhappy with the methods courses, saying “But there is nothing—. I mean, the first year teaching I just suffered. Wow! They didn't teach me anything.” However, he was happy with the work he did in mathematics courses. He said, “I took coursework, mathematics coursework, and that—. I would say out of all those, the coursework was the best for me because it gives you better content knowledge.” He believed having strong content knowledge enabled him to find alternative ways of teaching.

John was a National Board Certified teacher. He described the process of obtaining his certification as “the best in-service you can get.” He communicated that “it was all about reflection and introspection and reporting on what went on.”

From the early days of his teaching career, John had been ready to teach any level of mathematics. That was not possible initially, however, because he had taken a job at a school where a veteran teacher was already teaching the upper level mathematics (Calculus). The only option John had was to start something new, so he opted for AP Statistics. He started the AP Statistics program at his current school in 2000. John had opted for this course because of his experience as a graduate student. During his master's degree program, he had taken two statistics courses offered by the educational psychology department. He also took a statistics-for-teachers course, which was a course

offered by the statistics department to prepare high school teachers to teach AP Statistics.

John's motivation to teach statistics was also because he believed that there was a great need for statistics education. The following is his explanation:

And so I started doing research into the mathematics vs. statistics, you know, idea, and saw that there was a lot more need for statistics education than there was for trigonometry. For say, I mean, for engineers and those that, you know, you need that stuff. But journalism and art and humanities and psychology and every—almost every other field you look at—statistics is a way more utilitarian [than] mathematics, than what they were teaching at the high school at the time. I remember our department chair at that time in 1998–99 told me that nobody should graduate [from] high school without at least having passed trigonometry. And I said, “Well, I think that’s good.” But at the same time, you know, some of them are not going to be in this into science [so they can do] statistics. (John interview; November 24, 2008)

John's experience went beyond just teaching AP Statistics. He had, together with a professor from the statistics department at a Southeastern university, been involved in organizing a workshop for AP Statistics teachers.

John's Conceptions

Conceptions Related to Mathematics

John viewed mathematics as being made up of many different components. In general terms, however, mathematics mainly involved the development of models to answer real world problems. The following is his description of mathematics:

The definition of mathematics—. There are so many different parts of it. But I would think that if you were going to define *mathematics* to somebody that didn't know about it, it would be the ability then to take definitions in theory and use them to solve problems. And that's what mathematics, the study of mathematics, is all about. It is about modeling the real world and being able to use a mathematical model to answer questions about the real world. That's in my view what mathematics is. (John interview; November 24, 2008)

John viewed the essence of mathematics as residing in one's ability to ask oneself questions about certain phenomena. As an example, he said that long before mathematical models came into being, people would ask themselves questions like, "Why do boats float?" John indicated that it was important to teach theory, axioms, and theorems in mathematics, and that learning mathematics involved understanding of mathematical facts and building on them. However, that on its own was not enough. It was the application of mathematics that was really important. John viewed mathematics as a subject that was practical and thought that other countries do more with applications than the United States does.

It was obvious from the outset that John viewed mathematics as a subject that gave people a certain status. Mathematics is the subject for smart people. John said, "When you show somebody you have a math degree, they say, 'OK, well, you've got to be pretty smart.'"

John's conceptions of mathematics can also be inferred mainly from the comparisons he made with statistics:

Statistics is just more variable. You know, when you are solving a quadratic equation, there is one, none, or two solutions, you know. You know it's a function, and you are going to get it. And with statistics you have your variations. So there's, it is very little difference. (John interview; November 24, 2008)

In making these comparisons, John seemed to have a dualistic conception of mathematics. Unlike statistics, mathematics has no variations. His use of the example of the quadratic equation suggests that, for John, mathematics problems have a finite number of solutions. His explanation seems to suggest that in mathematics one will

always get the same solution. This notion suggests that one either gets that answer or not. This conception fits Perry's (1999) categorization of right or wrong answers.

It is also important to note, however, that John talked a lot about applications of mathematics—an indication that just learning theorems or theories is not enough. This emphasis shows that John's conception of mathematics went beyond dualism.

Conceptions Related to Statistics

According to John, the whole essence of statistics revolved around the idea of collecting data and also being able to analyze those data.

- I: You as a statistics teacher, how do you define *statistics* to someone who has never done statistics?
- J: Statistics is the art and science of learning from data. So if you want to teach statistics, the first thing you've got to be able to do is collect data, or if you are given data, . . . you've got to be able to analyze it. (John interview; November 24, 2008)

He reiterated the importance of data collection by distinguishing it from a mere exercise of collecting numbers. Statistics requires one to address some questions:

And then, of course, I tell [the students] that you can't make numbers say whatever you want. That's a myth. [Numbers] are what they are. And, you know, you get to analyze it. But you, you know, you can't change the way—the numbers that you got. So that's why data collection is so important. You know—. Who collected it? How did they collect it? For what reason was it collected? You know, some of those things. So then you can analyze it however you want. (John interview; November 24, 2008)

John also viewed statistics as a subject that was pure, with rules, and what was important was just to learn how to apply them. He said that is why some mathematicians enjoy teaching statistics. He said the following:

Because, you know, when I talk to other mathematicians—. You know, teaching of statistics is a breeze because it is so . . . pure, you know. The rules are just, you know, they are the same, they don't change. It's just, you know, you get to apply them. And so people that have used the math

part don't have [as] much trouble with statistics as people that don't do math—like your English teachers or social studies [teachers]. (John interview; November 24, 2008)

John talked about statistics as having *rules*; using Perry's (1970) framework, statistics may be regarded as being made up of *an array of discrete items* (the rules). Learning or teaching statistics is just a matter of applying these rules, which is equivalent to *simple obedience of the rules*, a characteristic that is typical of dualism.

John described people who understand statistics as those who understand the idea of variability in all the data they collect. To John, statistics was not just numbers but also the meaning attached to those numbers. Interpretation in statistics depends on the use of precise statistical language and also on the context of the problem.

J: I would be looking at [students'] interpretation of particular problems in the context of the problem. You know, you can always—. Well, let's say you are interpreting the slope of a regression equation. OK, now some people would think that it is just a direct proportionality, you know. You say, "Well, if my explanatory variable increases by one unit, then the response variable should increase by 0.83." OK, and they just leave it just like that. Well, if a person truly understands statistics, then they would say that if the explanatory variable increased one unit, then we would expect on average that the response variable would increase this much. Or, you know, that is what we would see most likely. Or sometimes this idea that there is variation in there. And so the people that really understand statistics are the ones that understand, I think, the variation that's involved in all of the data you collect and the way you collect your samples. And then they can—. So they understand the variability, and they are able to work within that idea that it varies a little bit.

I: So to you, interpretation is the main part of statistics?

J: Yeah. . . . If a person says, "The slope of the line is 0.83," then they are not going to do as well on the AP exam [as] someone [who] says that, you know, "Well, the slope of a regression line really measures the average change of the response variable with respect to the explanatory variable." Those people are getting a 5,⁵ you know. And so it has—.

⁵ AP Statistics examinations are graded on a five point scale: 5 Extremely well qualified*, 4 Well qualified*, 3 Qualified*, 2 Possibly qualified*, 1 No recommendation**, where * means qualified to receive college credit or advanced placement and ** means no recommendation to receive college credit or advanced placement. (College Board, 2007, p. 1)

Yeah, you can tell mostly by how they interpret their results, how well, you know, how well they understand statistics. So if all they can tell you is the answer, a number, they you've got to worry. And if they can explain what the number means, then you are probably in pretty good shape. (John interview; November 24, 2008)

When asked to compare mathematics and statistics, John gave a description that to a certain extent conflicts with his earlier characterization of statistics. In the responses (quoted on page 45) statistics is seen as variable. The implication here is that unlike mathematics, variability in statistics makes it possible to get different solutions to the same problem. John seemed to accept multiplicity in the number of solutions to statistics problems. Furthermore John seemed to consider mathematical problems as having a finite number of solutions, where there is restriction, whereas in statistical problems, he seemed to suggest, there are no set rules for the number of solutions one might get. In statistics, there is a diversity of solutions, so John's conception of statistics could be interpreted as multiplistic.

In summary, John presented a confused picture of his conceptions of statistics. On the one hand, he gave a dualistic picture of statistical knowledge by describing statistics as pure and saying that an understanding of statistics depends on the memorization and application of rules. On the other hand, he looked at statistics as more variable than mathematics, with multiple solutions to problems, which is a multiplistic picture.

Conceptions Related to the Difference Between Mathematics and Statistics

Though John acknowledged that a certain proficiency in mathematics was important in statistics, he also knew that there was a big difference between the two subject areas. I have already shown from the responses John gave how he differentiated

the two subjects. John also looked at the two subjects in terms of their usefulness to the lives of students. He considered statistics to be more useful than mathematics:

See one thing about stats [is that] we can bring it in to everyday from everywhere, and see [more] stuff than with math, you know. They bring in stuff about all new fractal research, you know. Well, these kids can't come close to [understanding] any of the cutting-edge stuff that you read about or see it in the newspaper about math. It is mostly stuff that they would never understand. But statistics, you know, . . . You know, you see it every day. (John interview; November 24, 2008)

Conceptions Related to Assessment in AP Statistics

John was well aware that there are different types of assessment. His first reaction to my question about his understanding of assessment was to ask me to specify whether I meant the daily assessment he did in his classroom or the assessment that is done at the end of a course. He also seemed to view assessment as a way of finding out what students knew by using multiple ways to establish that. In this case, if one used questions to elicit information from students, then there were a number of ways the same question could be asked.

I: What does *assessment*, *classroom assessment*, mean to you?

J: You mean like course assessment, or just class, just when I circulate around the classroom to see if they understand that day's topic, or whether it is the—?

I: Yeah, just assessment, assessment in general in statistics. What, what do you do?

J: Well, mostly it is just verbal response, you know. I'll ask questions; I'll talk about referring them, or I'll ask them to explain, you know—give me another example or give me a counterexample—stuff like that, to see if they understand what we are talking about. Draw a picture, you know. (December 8, 2008)

John believed that informal assessment could take place anywhere, even outside the classroom, where he would ask students questions casually. John also believed that assessment is not the sole responsibility of the teacher. Students should be a central part

of the assessment process. It was therefore important to give them the opportunity to help each other in trying to understand a particular concept or solve a problem. In that way, assessment serves two purposes: On the one hand, assessment improves the students' communication skills in transferring what they know to a fellow student; on the other hand, while that is taking place, the teacher has the opportunity to assess the student who is doing the job of explaining to another student. John believed that assessment should produce a "teaching moment" and give the students opportunities to correct their thinking. Formal assessment, for John, included tests and quizzes. For the statistics course, the tests were meant to "mimic AP Statistics."

- I: Because I know there is formal assessment, and there is informal assessment. So if you can just share with me some of the things that you do informally.
- J: Well, informally it happens in the hallways and at lunch and after school. When I would just stop them and say, you know . . . "What is a random variable?" You know, I see them somewhere, and we would be talking, and I say, "Oh yeah, that reminds me. What is a random variable?" So, you know, the informal assessment is that kind of thing. Or when in class I'll say, "Well, I'm stuck. . . . Somebody show us how to do this." You know, and they—I don't want to say every person's hand will go up—but, you know, there [are] plenty of volunteer[s] in there. So I mean by them working a problem, some other student is having trouble with, that's another way I informally assess—you know, both the person that is having the trouble and the other person—their ability to communicate and help the other people solve problems.
- I: OK, you said students do volunteer, or maybe—?
- J: Yeah. No, if I say, . . . if I was pretending to be stuck or just wanted them to do something, I would say, you know, "I need some help. Who can help me?" And so then, you know, the people, somebody would help me, you know. [They would] go up and say, "I'll do this one and that one." And then, you know, sometimes what they do is right, and other times what they do is wrong. And that leads to a great teaching moment so, you know, because well, so we fell in to this pitfall. There's, you know, there's "we set you up on purpose" [so] that we can show, you know, this is the common mistake a lot of people make, and we want to correct our thinking here.
- I: OK. Now on the formal part of it, what. . . ?

J: Well, [for the] formal part, I use tests generated by the [textbook] authors and different people that I know have been teaching stat a lot longer than I have that have really good success. So the formal assessment there is, it mimics the AP [Statistics] exam where they're the formal assessment. (December 8, 2008)

Though he used different assessment instruments, John had his preferences. He preferred projects because they assessed students in a number of areas, rather than questions that just addressed one thing only. He was not a big fan of multiple-choice questions, because he thought they did not truly reflect the students' abilities. He communicated this view in the assessment interview.

I: Do you have any particular method that you like in terms of assessment, that, "OK; I like this particular technique or method of assessing students?"

J: No, there's not, I mean, there's not one particular one. I know, well, I guess I kind of like projects in one class. And like in statistics, I like projects because it is a combination of a bunch of stuff. [The students] have investigations. They have to do, they have to write, you know, they have to do, you know, it is more than just a one-question thing. (December 8, 2008)

John also intimated that he liked the new Georgia Performance Standards (GPS) mathematics problems because they were "task oriented" instead of just asking students to solve routine problems like: "Solve a quadratic equation." He said:

Yes, I think it [GPS] is much better [than the previous curriculum], especially because when they [students] get to work, nobody is going to ask them to solve a quadratic. They are going to ask them to write it, or you are going to have, they are going to be presented with a problem, and they are going to have some tool, mathematical tool for you to model the situation and solve the problem. That's the stuff they are going to need, and the way we taught it traditionally was we give them all the tools, and never showed them how to use it. Okay, well, now they [GPS] are showing them [students] where to use the tool. (John interview; December 8, 2008)

It would appear that John liked problems that involved applications to real-life situations.

This preference was consistent with his conception of mathematics discussed on page 45.

Conceptions Related to Assessment in Mathematics

Because John was teaching both mathematics courses and the AP Statistics course, I wanted to find out whether his views of assessment in mathematics were different from those in statistics. His views seemed to be the same, apart from the fact that he had different expectations for the students in the two courses.

- J: My statistics classes are usually brighter students [than my mathematics classes]. They have gone through some of the math classes I have. So they are more able to just verbalize, to, you know, imagine, actually think more abstractly. So when we talk about a particular topic, . . . they can have an opinion and get an idea and respond correctly just with verbal cues—not necessarily needing so much visual stuff. Well, but I start with that in all my math classes, that the, you know, kids [should] be able to communicate, both written and verbal communication of what it is they know and what it is they don't know.
- I: So do you find yourself adjusting one way or another when you are, when you move from your statistics to mathematics?
- J: Well, I only adjust to the level. I know in my freshman classes, they tell me that I use words that are too big. OK, so I suggest that they bring in a dictionary with them because the words I use are SAT words. So I, . . . really, I try to maintain the same high standard. I have to adjust a little bit for their immaturity, but that would be about it. (John interview; December 8, 2008)

I was also interested in finding out whether there was any difference in the amount of work that students did in the AP Statistics course compared with the mathematics courses, say, in terms of number of tests given to students. The following is John's response:

I don't think that the number of tests is a lot different. But the type of answers that I am looking for and the amount of time it takes to grade and be careful of, you know, with it, that increases. (John interview; December 8, 2008)

John's Instructional Practices in Mathematics

John always started the Precalculus class by looking at homework problems that were assigned in the previous lesson. After discussion, he would then introduce the day's

lesson. John concentrated on making sure that his students understood the theorems. He made that point clearly in the conceptions interview. Whenever he taught, he encouraged students to take notes on the definitions, theorems, and examples that he discussed with them in class. When he asked or discussed problems, he would always ask: “Which theorem did I use?” Though he said that he never asked his students any questions on definitions in the exams, he still wanted them to understand the definitions, as definitions were crucial when it came to solving some mathematical problems.

John always tried to make sure that the students saw the connections between various topics. For example, in one lesson he asked students to write a polynomial of degree 6, given three of its complex zeros. He asked the students, “What does it mean to be the zero of a polynomial?” When the students did not answer, he asked them to go back to chapter 3, where he reminded them that in a polynomial if $f(c) = 0$, then $x - c$ is a factor. He gave them an example: If i is the zero of a polynomial, then $(x - i)$ is a factor, and so is its conjugate $(x + i)$. The students were then asked to write all the factors. It seems that John’s questions on the work the students had done in the earlier chapter enabled them to see the connection with the question he had asked them to do. The students were able to come up with the required factors.

In all the lessons, John tried hard to make sure that all the students got involved in the lesson. He did that in a number of ways, such as inviting the students to come to the board and show the rest of the class how they solved a particular problem. While teaching, he asked his students a lot of questions. John also wanted to make sure that his students saw that solving mathematical problems could be a tedious process and that

some problems could take a long time to solve. It was not uncommon for him to devote more than 15 minutes to going over a single problem.

John's Assessment Practices in Mathematics

Informal Assessment

John strongly believed in questioning as the backbone of informal assessment in his Precalculus class. At times, he would begin the lesson by asking students questions as a quick review of the previous day's work. He also asked a lot of questions while going over homework problems or going over class examples and exercises. In general, he asked questions throughout the lesson. Whenever students responded to John's questions, he would acknowledge each response in any of a number of ways such as, "That is very good." "You got that right." At times, he would go on to ask another student: "James, do you think that is right?" "Martin, what do you think?" "Does everyone agree?"

John was always willing to respond to students' questions. In some cases when students asked him questions, he would ask the class to respond. Sometimes instead of answering the student's question directly, John would ask the student questions that contained basic information that the student should have, to guide the student to answer his or her own question by applying that knowledge to a problem. For example, when the class was doing the problem on polynomials presented on page 53, a student asked the question: "How can I multiply the factors $[x - (4 - i)][x - (4 + i)]$?" John responded by asking the student: "Can you multiply $(a + b + c)(d + e + f)$?" The student replied that he could. John told him to apply the same method to the question he had asked.

John was disciplined in the way he conducted the questioning in class. He normally directed the questions to individual students, most especially to those he found not to be fully engaged. He usually discouraged the students from giving choral answers. He made sure that all the students were given equal opportunities to respond to his questions. In most cases, even when one student had given the correct answer, John would ask other students if they agreed with that answer. For example, while discussing hyperbolas, John asked, “If $y^2 - 4y - x^2 + 4x - 1 = 0$ is a hyperbola, can somebody give me the center of this hyperbola?” One student answered, “It is (2, 2).” John asked the class, “Does anybody agree?” The class almost in unison said yes. John then picked one student and asked, “Jane, did you get that?” The student said yes, and John said, “You do not sound convincing.” He then asked this student to explain how she got her answer.

John also made sure that he asked follow-up questions: For example, the class had just found one of the factors of the polynomial $x^4 - 9x^3 + 21x^2 + 21x - 130$. John asked the class a series of questions:

J: Can anyone tell me what remainder I will get when I divide $x^4 - 9x^3 + 21x^2 + 21x - 130$ by $x^2 - 6x + 13$?

S: You get 0.

J: Why is the remainder 0?

S: Because $x^2 - 6x + 13$ is a factor.

At times, John would not resist the temptation of moving quickly through the questions he asked. Sometimes he did not give the students time to reflect on the question. For example, John asked his students to find zeros of a quartic function, given $3 - 2i$ as one of its zeros. The question was meant for practice. John asked: “If $3 - 2i$ is a zero, can anyone tell me another zero? A student replied, “ $3 + 2i$.” John then said, “You guys should go ahead and find the other factors.” The moment the students started to work on

the problem, John asked them, “Can anyone tell me one factor?” A student said, “[$x - (3 + 2i)$].” John then gave the students a hint: “Build those two and multiply them out, and divide into the original polynomial.” In this case, John ended by going over the problem immediately, before the students could try it individually or in groups.

John also used monitoring in class for informal assessment. He would walk around the class to see what the students were doing and also to respond to their questions. He never graded any work in class that was part of his informal assessment.

This practice was in line with what he indicated in the interview on assessment:

I usually go round and just say, you know, “You got that right, you got those right, you got that right.” I don’t give them—. In class, we do very little where I grade it; the formal assessment. But I do move around, you know, and answer their questions. And most of the time, if they have time in class to work on problems, then I’m circulating and answering their questions; keeping kids on task. I try anyway. (John interview; December 8, 2008)

During class time, I never saw John assign any group work. But what he would do at times, if he asked the students to do some work in class, was to tell them that they were free to work together. Those students wanting to do that would normally pull their desks together and work on the problems. However, there was no class discussion involving what various groups found.

Formal Assessment

John used three main types of assessment instruments for his formal assessment in mathematics. These were (a) homework, (b) quizzes, and (c) tests and examinations.

Homework. Almost all the questions used in homework assignments came from the textbook, apart from a few questions John gave from worksheets. The students had an agenda—a sheet containing an outline of the topics to be covered and homework

problems for a given period like 3 weeks. On average, John assigned 6 homework problems each day. Sample homework problem similar to those assigned in his Precalculus class appear in Figure 2. Students were expected to complete all homework assignments before coming to class. John would normally ask the students at the beginning of the period if they had any questions from the homework problems. If there were any questions, he would go over them with the class or ask some students who got them right to explain to the rest of the class. He always encouraged students to write down the correct solutions during the review. On some days, he would collect homework. He graded homework for completion and for correct solutions of those problems discussed in class. He told me that he collected homework at random times to make sure that students were always prepared with their homework. He graded all the homework assignments he collected and recorded the grades.

- In Problems 1-2, find the center, transverse axis, vertices, foci, and asymptotes. Graph each equation using a graphing utility.
 1. $\frac{(x-2)^2}{4} - \frac{(y+3)^2}{9} = 1$
 2. $(y-2)^2 - 4(x+2)^2 = 4$

- **Suspension Bridge.** The cables of a suspension bridge are in the shape of a parabola. The towers supporting the cable are 500 feet apart and 200 feet high. If the cables are at a height of 20 feet midway between the towers, what is the height of the cable at a point 100 feet from the center of the bridge?

Figure 2. Sample homework problems similar to those John assigned

Quizzes. Quizzes came from the test bank, which was an accumulation of questions from John's many years of teaching mathematics. Most of the items in the test bank came from various textbook publishers' tests. All quizzes were given during class time, and in most cases, they took about 15 minutes to complete. All the quizzes were

graded, and the grades were recorded. Sample quiz questions (taken from John's test bank) appear in Figure 3. Before giving back the graded quizzes to the students, John would normally tell them the questions that most of them had missed. He would then go over those questions.

- Find an equation for the ellipse with center (1, 3), one focus at (1, 1), and one vertex at (1, 2)

A. $\frac{(x+1)^2}{9} + \frac{(y-3)^2}{25} = 1$

B. $\frac{(x-1)^2}{9} + \frac{(y+3)^2}{25} = 1$

C. $\frac{(x-1)^2}{25} + \frac{(y+3)^2}{9} = 1$

D. $\frac{(x+1)^2}{25} + \frac{(y-3)^2}{9} = 1$

- Graph $\frac{(x+2)^2}{25} + \frac{y^2}{9} = 1$

Center _____

Vertices _____

Foci _____

Figure 3. Sample problems from the ellipse quiz taken from John's test bank.

Tests and examinations. Most of the test questions came from the test bank. The test bank questions came from various textbook publishers' tests. John indicated that over the years he had written a few test questions. The reasons for constructing some of the items himself depended on the level of mathematics in the course. John saw his dependence on the test bank as a time-saving endeavor.

J: Well, you know when we are solving quadratics—especially when we get past the complex numbers where, you know, solutions could be, you know, radicals or complex numbers . . . —I make up all [the] stuff. But when we are doing rational equations or rational expressions, I tend to use what is in the bank just so that the numbers, you know, the answers come out reasonable—[something] that these kids, something they would recognize. But with the Algebra II, Algebra I, you know, those previous classes, I made up a lot of my own. And you know—

I: OK. So you have your own bank that you use?

J: Well, yes. It is on a disc that I have been doing, you know. All of them are on a disc, and I cut, copy, and paste from years and years. Yeah, I have a flash drive with that on it. And there are also the authors—. That all the textbooks we've bought the last 8 years, 10 years, come with a test generator CD. So you can mix and match and pull those in or not, you know, just use one of their questions. Or if you like the data, you know, if it is a data set, you can pull the data set and put it on your test and ask questions—your own questions—about that data.

I: Do you really find those helpful?

J: Oh, absolutely. Time, yeah.

I: In terms of time?

J: Yeah, because . . . just trying, you know, depending on the level of your course, you know. Like, when I'm teaching Analysis, I don't care if the answer is like $1 \pm i\sqrt{12}$. I mean, kids can handle that really easy when they are solving quadratics. But in the other classes, like Algebra I, you would like all the numbers to come out to be integers or reasonable fractions, you know, something like that. So using the test bank, you can see the question and the answer. And so it is easier to just pick one. . . . Obviously, when you solve this by factoring, the answers are going to be 3 halves and 5 halves; so that's pretty good. So pick that one. (John interview; November 24, 2008)

In general, the tests were composed of two sections. In the first section, students were not allowed to use any graphing utility. In the second section, graphing utilities were allowed. Sample test questions (taken from John's test bank) are shown in Figure 4.

<ul style="list-style-type: none"> Form a polynomial of degree 4 whose zeros are -2, 0, 2 and 3. (Do not leave answer in factored form.) Graph $f(x) = x^2 - 2x - 3$ by finding the vertex, x intercept(s) and y intercept. Identify on each on the graph. <p>Vertex _____</p> <p>X intercept(s) _____</p> <p>Y intercept _____</p>
--

Figure 4. Sample test problems for John's Analysis course taken from John's test bank.

The entire 90 minutes of the lesson were used for the tests, which were normally taken a day after a review of the entire chapter. Sometimes, practice sample tests from the students' textbook were used during the review for the test. All tests were closed notes and closed book. John graded all the tests and recorded the grades.

John indicated, however, that as far as the final examinations were concerned, all the teachers teaching the same course, say, Algebra II, wrote the test together by simply pulling questions from a common test bank. The choice of the questions depended on the common material that all the teachers had covered.

Analysis of Oral and Written Assessment Items Used in Mathematics

Table 2 presents results of my coding of the oral and written assessment items used by John in his Precalculus class. The results show that most of the oral and written assessment items were in Group A (see chapter 2 for explanation of Groups A, B, and C).

Table 2
Percent of John's Oral and Written Assessment Items Assessing Thinking Skills in Mathematics

Group	Description	Oral questions	Test and quiz items
Group A	Factual knowledge	27	0
	Comprehension	43	23
	Routine procedures	27	71
Group B	Information transfer	0	2
	Application in new situations	0	0
Group C	Justifying and interpreting	0	4
	Implications, conjectures and comparisons	3	0
	Evaluation	0	0

It is important to note, however, that the written assessment items did not assess any recall of factual knowledge. For example, students were never asked to state a theorem, whereas in oral assessment, students were at times asked to do that. Other thinking skills—that is, those in Groups B and C—were seldom assessed either in oral or written assessment items.

Further analysis of the written assessment items showed that none of the items were open ended. The majority of the questions, over 90%, were open-middled. According to Bush and Greer (1999), open-ended items are those that have multiple solutions and whose solutions can be found in several ways. Open-middled items are those that have one correct answer, but there are multiple ways of getting to that answer. There was no clear rationale for the use of these items. However, I noted that John used multiple-choice items in quizzes and not in tests. The small percentage of multiple-choice items could be attributed to the fact that John did not like such items, as he indicated in the assessment interview. But it is also possible that John just used the format in the textbook publisher's test book. The same could be said of the absence of open-ended items in the quizzes and tests.

John's Grading Process in Mathematics

John used a grading process that was uniform for all teachers teaching the same course, as mandated by the school administration. Three types of assessment instruments—namely, homework, quizzes, and tests (chapter tests, midterm, and final)—were used to determine the students' grades. The breakdown of the percentages for each of these instruments is displayed in Table 3. Chapter tests were the main determinant of the students' grades, contributing 50% to the grades. It is important to note here that only written assessments were used to determine the students' grades.

Table 3
Grading Schedule for Mathematics

Instrument	Test	Quiz	Homework	Midterm	Final
Percent	50	10	10	10	20

John's Instructional Practices in AP Statistics

John usually started his lesson with a review question from the previous day's topic. He would then ask the students if they had questions from the previous homework problems. John seemed to enjoy himself in the statistics class. His approach to teaching statistics involved the use of real-life examples that students could identify with to provoke discussion. He had earlier indicated in the conceptions interview that he liked statistics more than mathematics. He said it was very easy to demonstrate to the students the important role that statistics plays in their lives. In the class on February 12, John had an interesting discussion with his students. He told them that the accident response time by the medical people in the county was $6\frac{3}{4}$ minutes. That comment triggered a debate, and the students asked a number of questions such as:

- How did they measure that?
- How do you know that is the time when you are involved in the accident?

John did not answer those questions immediately but went on to finish the rest of the example:

J: Insurance people were not really convinced that that claim was true. A sample of 400 from the population of the county was taken to find out whether that claim was false. How do we find out that?

Some of the students had seen emergency situations in the area, and the response time had been greater than $6\frac{3}{4}$ minutes. The discussion went on to include the idea of mean response time—that is, some response times will be greater than $6\frac{3}{4}$ minutes, and others will be less. The students participated very well when this example was used. John used this scenario to introduce the concept of significance test to demonstrate how such claims can be tested statistically. This example was in line with John's belief about

the role of a teacher. In the conceptions interview, he had described his role as a statistics teacher. Apart from preparing his students for the AP Statistics examination, he saw that he had another role:

It's [to] make them more aware statistically so that when they see a report or they see a graph or some kind of picture in the newspaper, they can make sense of it. And now even analyze it to find out if it is true or if it is not true. Or that, you know [someone] is trying to pull the wool over their eyes. Those kinds of things, you know, just make them aware. (John interview; November 24, 2008)

During the interview on assessment, John mentioned multiple representations. He said that one way of making sure that students demonstrate understanding of a particular concept or problem is to ask them to express their solutions in a number of ways. It was the same approach that he demonstrated in class: to make sure that his students appreciated the notion of multiple representations. He also mentioned that whenever he demonstrated how to solve problems, he would use technology wherever applicable and also other methods without technology. For example, he showed his students that problems can be solved in multiple ways. In the lesson, the students mentioned that they did not understand Part b of the homework question from the textbook (Yates, Moore, & Starnes, 2008) John used. A problem similar to the homework problem is shown in Figure 5.

Motor Vehicles: Ford Motors make up 14 % of all motor vehicles registered in the United States. You plan to interview an SRS of 500 motor vehicle owners.

- a) What is the approximate distribution of your sample that own Ford cars?
- b) How likely is your sample to contain 20 % or more who own Ford cars? Do a Normal probability calculation to answer this question.
- c) How likely is your sample to contain at least 15 % who own Ford cars? Do a Normal probability calculation to answer this question.

Figure 5. Homework problem similar to the one that John assigned.

John asked the students to calculate the z -value. The students did that, and he drew the picture on the board similar to the one shown in Figure 6. John asked the students a number of questions.

J: Can you tell me where on the curve I can mark off $z = 3.87$?

S1: The point will be far off to the right.

J: What will be the probability that $z > 3.87$, approximately?

S2: It will be very small, approximately 0.

T: Can you check the exact answer using your TI [calculator]?

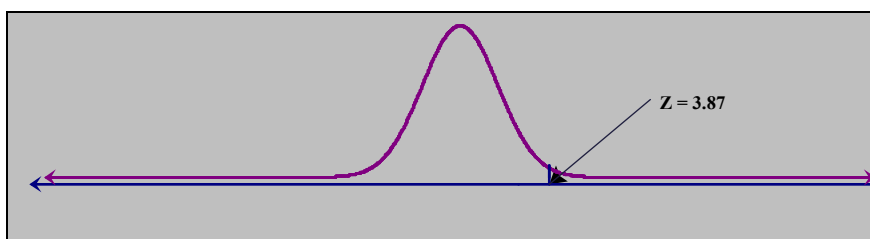


Figure 6. John's drawing of a normal curve.

All the students used their calculators to find the correct probability. Though the question did not ask the students to specifically use a diagram, John encouraged them to draw the curve and make a rough estimate of the probability that they were looking for. It was clear from the students' reaction that the calculation made sense once John introduced the diagram of the curve. Though simply using a graphing utility could have solved the problem, John believed that students should be equipped with multiple ways of solving a problem. If students were able to solve problems in multiple ways or different ways, that showed that they understood the concept.

John also wanted the students to see how different statistical concepts were connected. Whenever he was discussing a concept, he would link it to other concepts. Though John depended very much on examples from the textbook, he easily extended those examples to cover other areas related to a particular question. He was willing to go

beyond the book. He seemed to have a firm grasp of the content. He was very creative in his instructional approach and got his students motivated.

John used his experience in teaching AP Statistics to highlight some of the misconceptions students have in statistics, though at times he started by mentioning the misconception before giving the students the opportunity to see if they could make the same mistake. For example, John discussed the difference between finding the probability for an individual value of x , say $P(x > 68)$ and $P(\bar{x} > 68)$, given that one is sampling from a normal population, $N(64.5, 2.5)$. He told the students: “The difference is that in the two situations the mean is the same, but the standard deviations will be different because the sampling standard deviation of \bar{x} is narrower.” He cautioned them, “A lot of students get mixed up here, so you have to be careful.” To illustrate the point John drew two sketches on a transparency.

John’s Assessment Practices in AP Statistics

Informal Assessment

John used a lot of questioning in class. He always emphasized the full interpretation of numerical answers in context and asked for answers in full sentences. If a student said, “The mean is 12,” John would correct him or her and say, “You’ve got to say it fully, like the mean age for girls is 12.” He wanted his students to be precise in the use of statistical language. Often in class, he would tell the students that statisticians are paid for interpreting statistical information rather than just understanding the mechanics of operating a calculator. This comment was also consistent with John’s belief that though technology was important in statistics, it was the interpretations that were more useful.

When John gave his students work to do in class, he would normally walk around to see how they were doing the problems. He encouraged the students to work together in class; however, he never made it mandatory for students to do that. He used no activities that were specifically designed as group work. All the students did the same activities or problems at the same time.

Formal Assessment

John used four main assessment instruments for his formal assessment in AP Statistics. These were homework, quizzes, tests, and projects.

Homework. All the homework problems that John assigned in AP Statistics came from the textbook that the students used. Sample homework problems similar to those assigned from the textbook (Yates, Moore, & Starnes, 2008) appear in Figure 7. John always gave the students an agenda sheet detailing what would be covered in a given lesson, including the homework problems and the dates for tests and quizzes. There was an average of 4 or 5 problems on each homework assignment not including reading assignments and practice tests, which were sometimes included as homework. Students were expected to bring their solved homework problems to class. Before covering the day's lesson, John would ask the students if they had any questions from the homework. If there were questions, he would discuss them with the class. During this time, the students were expected to make corrections. Normally, John would go over several problems, especially those that the majority of students said they had difficulties with. All the problems discussed in class would later be collected and graded for accuracy. John called the set of solved problems a "homework quiz." The rest of the problems would be graded for completion only. John recorded all the homework grades.

- (1) State whether the boldface number is a parameter or a statistic, and (2) use the appropriate notation to describe each number.

How tall? A random sample of male high school students has a mean height of **64.5** inches, which is greater than the **63**-inch mean height of all adult American men.

- Based on the information from the University of Hope, the heights of male students closely follow a normal distribution with mean 69.2 inches and standard deviation 3.1. inches. The heights of female students follow a normal distribution with mean 63.5 and standard deviation 2.7 inches. A particular group of University of Hope students has 12 males and 8 females.
 - a) Treat the students in this class as random samples of male and female students. Use random variables to calculate the probability that the mean height of male students is greater than the mean height of female students. Show your work.
 - b) Design a calculator simulation to estimate the probability described in (a). Describe your simulation method clearly enough so that a fellow student could carry it out without further explanation from you. Then carry out the simulation. Interpret your results.

Figure 7. Sample homework problems similar to those John assigned for his AP Statistics class.

Quizzes. All the quizzes came from the quiz book (Yates & Starnes, 2008) that accompanied the textbook. John would make copies of them without any changes. Normally, a quiz would cover a given section of the textbook that had been discussed in class. At times, however, John would combine two sections, and the students would take a single quiz on the combined sections. At least a day before a quiz, John would tell the students what sections the quiz would cover and also the number of problems that would be included. At times, he would conduct a quick review to remind the students of the topics they were supposed to know in preparation for the quiz. During the review, he would also answer any student's questions.

The quizzes were usually taken a few minutes after the beginning of class. All the quizzes were closed notes and closed book. There were no restrictions on the use of technology (graphing calculators). The quizzes were normally to be completed within 15

minutes, but John was flexible in allowing more time if the students needed it. He graded each quiz, recorded the grades, and returned the papers to the students the next class period. Figure 8 shows a sample of quiz problems similar to those John used for his AP Statistics class.

<ul style="list-style-type: none"> • HRM County data shows that 22% of county children under the age of 6 live in households with incomes less than the official poverty level. A study chooses an SRS of 500 children. What is the probability that more than 20 % of the sample are from poverty households? (Remember to check that you can use the Normal approximation) • A random sample of 1500 teenagers (ages 16 to 19) was asked whether they watched football on television; 1019 said they did. Construct and interpret a 99% confidence interval for the population proportion p. Follow the Inference toolbox (a box containing the steps used in constructing confidence intervals).

Figure 8. Sample quiz problems similar to those John used for his AP Statistics class.

Tests. Most of the tests were sample tests (Yates & Starnes, 2008) from the publishers of the teachers' textbook. At times, John would draw from some other sources, like past examination questions, and put them together to form a test. He spoke about the construction of test items in the assessment interview.

- I: OK, do you get to form your own questions sometimes?
- J: I do. Oftentimes, I form my own questions, but it is so time consuming that, you know, over the years, you know, I write maybe one or two new questions per chapter per year. So I imagine in another 10 years, then I could actually put together a test bank or someone else. But I don't—[pause]. The majority of the test questions are not my own.
- I: OK. So, in AP Statistics you are saying you have a test generator. So do you have the resources for all those tests that you talked—?
- J: Yes, yes. And I could pick the different multiple-choice questions out of the test bank. And that's what I use, you know. I use other people's multiple-choice [questions] because it is so hard to come up with good distracters and other answers, you know. So—. And then the free-response question is easy to write; except some of it, it is so data driven that I've got to take, I've got to get data from somewhere,

you know, doing all that hard work. And if I did that now, all [I've] got to do is to search the bank to find the questions that I need for a particular topic. (December 8, 2008)

The tests were meant to mimic the AP Statistics examination, so they had two parts. Part 1 contained multiple-choice items, and Part 2 contained free-response items. The free-response items included some short-answer questions and an investigative task. Sample test items similar to those John took from the quiz book (Yates & Starnes, 2008) accompanying the textbook are shown in Figure 9. All the tests were closed notes and closed book. When I asked John why, he said, “I do not [make the test open notes or open book], because the students spend most of the time flipping through the pages instead of concentrating and understanding the questions.”

- The sampling distribution of a statistic is
 - a) The probability that we obtain a statistic in repeated random samples
 - b) The mechanism that determines whether randomization was effective
 - c) The distribution of values taken by a statistic in all possible samples of the same sample size from the same population
 - d) The extent to which the sample results differ systematically from the truth
 - e) None of the above. The answer is _____.
- The State of Alabama is considering additional restrictions on the number of vehicles allowed to enter Birmingham Botanical Gardens. To assess public reaction, the state asks a random sample of 200 visitors if they favor the proposal. Of these, 110 say “Yes.”

Construct and interpret a 99% confidence interval for the proportion of visitors to Birmingham Botanical Gardens who favor the restrictions.

Figure 9. Sample test problems similar to those John used for his AP Statistics class.

The day before a test, John would devote the entire lesson to review. He would usually give the students a practice test. Though in the AP Statistics examination, students are given a formula sheet, John did not give them one for his tests. He

explained: “We compartmentalize [the test] within one chapter, but when they take the AP exam, they have 14 chapters.” An entire lesson (90 minutes) was allowed for each test. All tests were graded, the grades were recorded, and the test papers were returned in the next class period.

Analysis of Oral and Written Assessment Items Used in AP Statistics

Table 4 presents results of the coding of oral and written assessment items used by John in his AP Statistics class. The results show that most of the assessment items,

Table 4

Percent of John’s Oral and Written Assessment Items Assessing Thinking Skills in AP Statistics

Group	Description	Oral questions	Test and quiz items
Group A	Factual knowledge	27	25
	Comprehension	14	24
	Routine procedures	43	29
Group B	Information transfer	5	3
	Application in new situations	0	0
Group C	Justifying and interpreting	11	15
	Implications, conjectures and comparisons	0	7
	Evaluation	0	0

both oral and written were in Group A. In the written assessments, the percentage of the items requiring factual knowledge, comprehension of factual information, and routine use of procedures were not that much different, whereas for the oral assessment items, the majority of the items in Group A required comprehension of factual knowledge. Items requiring thinking skills in Group B were hardly used. The percentage for written assessment items in Group C was twice that of the oral assessment items. It was evident,

however, that posers of both oral and assessment items ignored the application in new situation and evaluation thinking skills.

Further analysis of the written assessment items showed that less than 2% of the items were open ended. In general, there were no multiple-choice or true-false items in the quizzes. In the tests, however, almost 50 % of the items were multiple-choice. John used such items so as to conform to the structure of the AP Statistics Examination.

Projects. Projects in John's classes usually involved an extended investigative task. All the projects came from the teacher's resource book. John gave only one project a semester. At the time of the study, he had not yet assigned a project. He told me that he was still thinking about what kind of project he would give.

John's Grading Process in AP Statistics

John used four types of instruments to determine the students' grades: homework, quizzes, tests, and projects. Chapter tests contributed a higher percentage to the students' overall grade, with homework and the midterm contributing the least. Only written assessment instruments were used to determine the students' grades. The breakdown of the percentages for these instruments is displayed in Table 5.

Table 5

Percent Distribution of Grades for John's AP Statistics Course

Instrument	Test	Quiz	Homework	Midterm	Project
Percent	50	10	10	10	20

Factors Influencing Assessment in Mathematics

I identified three factors that affected John's assessment practices in mathematics: the caliber of the students, the administration policy, and standardized tests.

Caliber of Students

John's perception of the type of students in his mathematics courses—mainly Algebra II and Mathematics Support I—affected the way he assessed the students. He found it difficult to ask the students in those courses questions involving applications. He said they could do routine problems only.

And for the kids outside of Statistics, it is a lot harder. We do—. Like today, we were doing direct variation with a square. So Y is varying directly as X^2 . And when it is all equation driven, they don't have a problem. But when you put it in context, like the intensity of a light [as] you move away from it, then all of a sudden they don't have X and Y anymore. They've got D and I . And that the relationship is the same, but they get lost. So they—[pause]. The assessment for the math, it's, you know, it seems to be . . . reduced to a single number. You know, with the direct—an equation where they can substitute it in without difficulty. But when they have to take the information from a problem like the stem of the problem and then put it into the formula, get a number, and when they bring it back, they can be able to express what that means, you know, outside of the classroom walls. That's very difficult. (John interview, December 8, 2008)

John explicitly discussed the issue of his students' expectations. He thought that the students in Algebra II seemed not to want to stretch themselves:

I guess it's really a shame, but the expectation enters into it, too. Because the kids in Statistics, they want to be able to do what I, what we as educators, want. And especially with the new [Georgia Performance] Standards. We want them to be able to see how [mathematics] works in real life and take a word problem and model it mathematically. And . . . come up with some solution, and then, you know, put it back to a workplace. And [say], "This is how we would do it." And with the kids that I have in Algebra II, most of them have failed at least one or two math classes in their high school career. And they don't want that challenge. What they want is the bare bones and minimum. . . . So I try to give them more than that, but it is hard, you know. It's like the old saying, "You can lead a horse to water, but you can't make him drink." So in the class, you model for them something about mathematics. You model the, how beautiful it can be, and how it improves their daily lives. . . . But for a lot of them, it is just a very complicated system of rules that don't apply to them outside of their classroom. That's unfortunate, but that's what we see a lot. (John interview; December 8, 2008)

Although John said the students in Algebra II did not want to exert themselves fully, he saw the situation as different in Precalculus. The students in the Precalculus course, according to John, were motivated to learn. He said:

The difference . . . is that Precalculus is a math elective. It is an academic elective. So the kids that go that far in math, it is because they want the challenge. And so it is easier, you know, without having to work with the motivation, the education becomes a lot easier because those kids are willing to do whatever you ask of them to keep them challenged. And in Algebra II, that's the third course of our mandatory four-course sequence for graduation, and a lot of these kids, well, I mean, they are good at the other things. I mean, I have kids that are really good in literature or drama, or maybe even music or science. And they just don't like math, you know. So when you are in the required course load, you know, that content—. [In] the courses that deal with just the requirements, the variation in student ability and motivation [is] very, very wide. But when you talk about teaching the elective courses, the academic electives in math, these kids, the variation in their motivation and their ability shrinks. And so then you can take that whole group and move them a lot further and a lot faster than you can with kids that perceive the course as a necessary evil to graduation instead of an opportunity to be challenged. (John interview; December 8, 2008)

John did not do any projects with his algebra students, because of the lack of commitment that he saw on the part of the students:

J: In mathematics, I just gave an extra credit project; it is called "career poster." OK, so they were supposed to interview someone that uses math, and they were supposed to, you know, list a problem, OK, and then [write] a paragraph to explain how this person would solve that problem in their job. And then they [were] supposed to turn it in. And then [I gave] very specific instructions on what they [were] supposed to do. And [not] one . . . of them of [their posters] was worth a dime.

I: Really?

J: You know, after you have been teaching for 20 years, you can see [the difference between] a poster that was put together the night before it was due in 15 minutes, and the poster that the kid started on, took his time, did the interview, wrote it, proofed it, put it together nicely, and then, you know, . . . turned it in. And so it has been my experience in mandatory math classes that the posters just don't—. . . . They take so long to grade, and they require so much more effort that, you know, it is hardly worth it. (John interview, December 8, 2008)

He also differentiated between teaching advanced level mathematics and remedial mathematics.

And I like math. When I teach an Advanced Algebra II or I am teaching Analysis, which is our Precalculus class, I really like the math. I really do, so I like it all. But when I have to teach a remedial [mathematics course], that is more like a job than, you know, in other times. Like, most of the time I'm having fun. So it is not really like a job. But when I'm having remedial classes, then it is definitely a job. (John interview; December 8, 2008)

Influence of Standardized Tests

John indicated that he had never liked multiple-choice tests. He had to prepare his students for standardized tests, however, so he was left with no choice.

OK, and so then multiple choice; I don't like multiple choice too much, because I always call it multiple guess. And you know, they can eliminate a few things. But standardized tests are all multiple choice, and so [the students] need to be able to see some of those. And so, you know, you want that to be a component, you know. (John interview; December 8, 2008)

School Administration Policy

The school administration had a policy that dictated the grading scheme for all the mathematics teachers teaching the same course. This policy was mainly due to the administration yielding to pressure from parents.

We were told years ago that we all have to have the same kind of grading scheme because at the school there are people saying, "I don't want teacher Y, because they are too hard; their tests count too much; they don't give us credit for homework. And this teacher over here does something [different]." So for each class that we teach, we have [requirements] for how their grade is broken down. (John interview; December 8, 2008)

The school policy was the following: "Grades to be 70 and above are passing, and credit is awarded at the end of each semester for each course successfully completed." The cutoff points for letter grades were determined at the school level. At John's school, the ranges were as follows: A = 90–100, B = 80–89, C = 70–79, and F = 0–69. This

information was posted on the school's Web site and also appeared in the school handbook.

Factors Influencing Assessment in AP Statistics

After going through the information John provided in the interviews and the information collected through class observations, I identified a number of factors that influenced John's assessment practices in AP Statistics. The factors were time, school administration policy, the external examination, technology, the textbook publishers' tests, the caliber of the students, and technology.

Time

John indicated that the county school district had decided that the number of hours for teaching AP Statistics should be reduced, a development he was not happy about. John noted that over several years the performance of the students on the AP Statistics examination had declined because of the changes in the number of contact hours with the students. John thought that teaching AP Statistics every other day had affected the students most in how they did homework assignments.

Time was reduced from 180 days to 90, and the [AP Statistics examination] pass rate dropped immediately from 83% to 51%. Changes are underway to change the schedule so that the teacher sees the students every day for 45 minutes. This will be helpful . . . because when I don't see them but every other day; they tend not to think statistically that whole day. . . . And then they go, "Oh, shoot, I had homework. I didn't do it today, you know. Tuesday when I got it assigned, I waited. Then I forgot how to do it. Can you help me?" (John interview; December 8, 2008)

To emphasize the effect of time on the students' performance, John also indicated that the pass rate for AP Statistics had been high even at the time when the school had a large number of students taking the course. In other words, the overall drop in students'

performance was clearly a result of the reduction in contact hours and not because of large classes.

Our numbers were really bulging before the new school opened. So we had a large number of kids here, but we had [an] 83% pass rate, with half of those passers getting 5s [on a scale of 1 to 5, where 5 is the highest score]. (John interview; December 8, 2008)

Because there was not enough time to cover all the course material, John indicated that he had to change the number of projects he gave his students in a semester: “I used to give [the students] two [projects] a semester, but now we are down to 90 days. They are going to, they get one each semester.”

Apart from the fact that it was difficult for John to write his own assessment items in AP Statistics because most of the questions were data driven, he also indicated that it was time consuming to write test items in AP Statistics.

School Administration Policy

AP Statistics was opened up to more students in John’s district to give them the opportunity to experience a more rigorous curriculum. This decision was made by the school district administration. John was not against the decision, but he said that the decrease in the contact hours had affected most the students with borderline abilities. John suggested that those students needed extra help because they could not do most of the work on their own. He said:

And we have opened the door more to not being so selective, you know, so that we can offer [the course] to more people. Which means a benefit, I believe, to some of these students getting a chance to see a rigorous curriculum before they leave high school. . . . We have been doing that when we had them for all year, and we still weren’t successful. And then scaling back to half—to 90 days from 180—it affects the borderline kids more than it does the top 5%, because we are still going to get our 5s and 4s because of the gifted students that are taking the course. But the ones that are, you know, traditionally, you know, A or B students, they are

getting 2s and 3s—somewhere in that range—because they don’t have the strong, I guess, they are just not as intelligent, you know. They can’t get it all on their own. They don’t get as much on their own as some of the other kids can. (John interview; November 24, 2008)

John’s AP Statistics class had fewer students because the original class that had 31 students was split into two to reduce the class size.

External Examination

John had indicated to me that part of his responsibility was to prepare the students for the AP Statistics examination. Although he was quick to say that he did not teach to the examination, it was apparent that his assessment was mainly geared toward it. In most of the lessons I observed, he talked about the examination. Some of the things he said or did included the following:

- The review of homework, quizzes, and tests was centered on identifying misconceptions and common mistakes found in the AP Statistics exam.
- When John said in class, “Big points come from interpreting the meaning of the confidence interval in the context of the problem,” he was emphasizing the importance of interpretation in the AP Statistics exam.
- Most of the time, the students were given specific ways as to how they could tackle certain problems. For example, John said, “In the exam, you need to give a clear explanation of what you are doing in the simulation.”

The AP Statistics examination included multiple-choice items, so John gave his students some practice every time he gave them a chapter test. Otherwise, he said that he would not have used multiple-choice items in his assessment.

Technology

The main technology used in the AP Statistics examination was the TI-83 graphing utility. The availability of technology affected John's approach to assessment, such as putting more emphasis on interpretations than computations and explaining solutions in the context of the problem. The following is his explanation:

[Technology] plays a big role [in the course] because when, I mean, the TI will do everything for you. It will give you—. It will do a test; it will draw the picture; it will, you know, it is marvelous. It frees up all of your mental energy to talk about the final result, you know, the conclusion in the context of the problem. Because you can spend all your energy there and not on the fact that you are finding Z -scores, and you are doing all that stuff and having a big long trail of manipulations to get to your p -value. You know, the calculator will draw the picture. You sketch the picture, you shade it in, and it will help you explain why this part is shaded and what this p -value means in [the] context of the problem, and you are good to go. The same thing with random variables: You find the mean of a random variable and talk about, you know, just it; the technology, it just makes the computations faster. So we get to spend time on interpretation. And that's, I think, the most important part [of statistics]. (John interview; November 24, 2008)

John indicated that even though computer technology was necessary in statistics, he did not dwell on it much because the students were not going to use computers in taking the AP examination.

If I use MINITAB, or DATA-DISC, or FATHOM, you know, they have good applets, and they can let you—. [A computer] has more memory; you can observe more data on that stuff. But I focus so heavily on the TI-83 or 4. (John interview; November 24, 2008)

In fact, during the time of the study, John took his AP Statistics students to the computer lab only once. Graphing calculators, however, were used on daily basis.

Textbook Publisher's Tests

Almost all the homework problems came from the textbook (Yates et al., 2008), and the quizzes and tests came from those supplied by the textbook publishers (Yates &

Starnes, 2008). As already noted, John did not make any change in the test items; if anything, what he did was just to rearrange questions. Evidence of the source could be seen on the test copies, which bore page numbers, test numbers, or quiz numbers from the test book.

Caliber of Students

Although John had already indicated that the decision to allow more students to take AP Statistics meant that even relatively weak students were enrolled in the course, he still described AP Statistics students as bright kids who would not have any trouble reading. One of the homework assignments John used to give his AP Statistics students was a reading assignment. John appeared motivated to do that partly because he expected his students to be able to read the material and at least understand it with minimum assistance from him. He also described these students as self-confident and outspoken.

Like I said, it is an AP class. So you are getting the best of the brightest students, the ones that read and will write well, the ones that have more self-confidence, the ones that traditionally speak out; they do service projects. They know they are not shy or introverted during that kind of stuff. (John interview; November 24, 2008)

Summary of John

The analysis presented in this chapter indicates that John was able to anticipate students' problems, give real-world examples and also demonstrate multiple ways of solving problems in AP Statistics. His content knowledge was equally good in both AP Statistics and Precalculus. However, his pedagogical content knowledge in AP Statistics was superior to his pedagogical content knowledge in Precalculus. Perhaps this difference could be attributed to the fact John had indicated that he liked teaching AP Statistics more than Precalculus. His preparation in AP Statistics seems to have played a

role as well. In fact, John had more experience in teaching AP Statistics, and better pedagogical content knowledge in AP Statistics, than Mary, who we meet in the next chapter.

CHAPTER 5

THE CASE OF MARY

General Background

Mary was a White female who at the time of the first interview was in her mid 50s. She grew up in a small town in a Southeastern state. Mary developed an interest in sciences at a very young age, mainly because of her father. She said this about him:

He had a great influence on me mostly on science. When I was little, I never played with dolls like the other kids did. Dad used to buy me a chemistry set, and I would play with it to make different combinations of substances. I remember my mother would get scared that maybe I would make something dangerous. (Mary interview; December 4, 2008)

As a young person, Mary wanted to be a doctor. Apart from science, her dad also got her interested in mathematics:

My dad used to ask me to solve some mathematical problems mentally. For example, when we were traveling in our car, say at 50 miles per hour, he would ask me to calculate how long it would take us to get to our destination if we were moving at that speed. (Mary interview; December 4, 2008)

Mary described herself as having always been good in mathematics and science and as having enjoyed the two subjects. She had always tutored her peers in both her high school and college mathematics classes. It was not until a college professor pulled her aside and asked her whether she had considered a career in teaching that she decided to go into teaching. Because she had already started helping her fellow students with mathematics, she thought she could just go on to become a mathematics teacher.

Mary attended a university in a Southeastern state. She obtained bachelor's degrees in mathematics and mathematics education and also obtained a specialist degree in mathematics education. She was a National Board Certified teacher. She started teaching in 1993. In addition to a statistics course she took as an undergraduate, she attended a 1-week seminar offered by the same university where she obtained the degrees stated above, to prepare her to teach AP Statistics. At the time of this study, she was the chairperson of the mathematics department at her school.

Mary did not like her experiences as an undergraduate student. She described some of her professors as very difficult to approach if she had a problem, and she thought that she had learned most of the mathematics on her own. However, she considered work that she did in graduate school as very helpful because it was tailored to meet the needs of her classroom teaching. She was also full of praise for the applied project in her specialist degree program:

Really and truly if I were to say what has helped me reflect on teaching and my grading practices, [it was] going through that process, because it was huge. I mean the process itself was just monumental, and I mean putting everything together and having to look at yourself and the ways that you do things and trying to tell somebody how you do things that you don't talk [about] or share. You just have to give it to them and try to explain it in words, and [your understanding is different]. That was probably one of the biggest moments I think for me. (Mary interview; December 4, 2008)

Mary had initially trained as mathematics teacher and not a statistics teacher. She could not recall whether or not she had liked statistics as an undergraduate student. Her first experience in teaching statistics came as a result of the state department of education changing the state curriculum for the Algebra-Trigonometry course. That change meant that a third of the course had to be statistics. She had started teaching AP Statistics when

she moved to a new school, where she was teaching at the time of this study. She thought that her experience with regular statistics was a good preparation for teaching AP Statistics. She conceded that teaching AP Statistics was a big challenge:

I started teaching the statistics, the actual statistics course, not the AP Statistics, because I knew I was going to be teaching the AP Statistics. And I said, “Let me kind of get my feet wet first”; so I turned to regular statistics before I ever taught the AP [Statistics]. Just to kind of like, you know, get a good feel of what exactly the course was going to be all about, knowing that AP [Statistics] was going to kind of supercharge it a little bit. . . . And then at the same time, I was on the AP [Statistics] Learning Community at the same time, kind of hearing what they were going through, and it scared me to death (laughs). (Mary interview; December 4, 2008)

Though she knew that teaching AP Statistics would be a challenge, Mary indicated that she enjoyed teaching statistics more than mathematics because it had interesting tasks for students to do.

Mary's Conceptions

Conceptions Related to Mathematics

When I asked Mary to give me her definition of *mathematics*, she hesitated a bit. After a few tries, it emerged that her view of mathematics was mainly based on relationships. Her interest in looking at relationships started at a young age.

I: Now what is your view of *mathematics*? How would you define mathematics?

M: My role?

I: Your view of *mathematics*.

M: My view of mathematics? Keep talking, like “view” as in the world or

. . .

I: What would you say is *mathematics*?

M: I would say I don't know that I would have to. I am not sure I understand what you are asking, but I don't think, you know, as a high school teacher [that] I would have to say what is mathematics, because they have already had 8 years of it. When they come here then they may say, “What is trigonometry?” You know, and something like that and trying to show them the relationships that, you know, we usually

talk about triangles and then shapes, and relationships that things have with one another. But I don't know that I would have to define *mathematics* to someone. Do you think you would have to define *mathematics* to someone?

I: Oh, I just thought that maybe you would have any definition. Say this is the way I look at mathematics. Maybe some people say mathematics is about formulas, or mathematics is about—.

M: Oh, Lord! No, you see, you know how I feel about formulas (laughs). Maybe if you would define, just tell me about the relationship between anything, objects, people, you know—I always look at numbers and kind of look at relationships between numbers. Like—. I don't play the lottery, but I look at the numbers as they come out, and kind of look at the relationship, and “Oh, look at the [relationship between the numbers].” I can always make some sort of relationship.

I: Relationship, yeah.

M: With the numbers that are picked. And my mom used to tell me, “Why do you do that? Or how do you do that?” But mathematics, I guess is the relationship between many things: shapes objects, people, relationships between numbers, and numbers themselves. (Mary interview, December 4, 2008)

Mary also looked at mathematics in terms of its usefulness in other subject areas. She looked at mathematics as a tool. She said that most of the “applications of [mathematics] are scientific.” She indicated that she used a lot of science examples in her mathematics class.

It was difficult to fit Mary's conceptions into the Perry (1970) scheme from the responses she gave about mathematics. However, when Mary explained the difference between mathematics and statistics, her conception of mathematics emerged more clearly. She said:

When I talk with other teachers of [mathematics], I tell them that these [statistics] problems are not skills and drills. You know these skills in math. In statistics the problems are more like, “Here is a scenario; now how do you solve it?” Because it is no more of, you know, “Simplify the square root of 50.” You know, there is no more of that. Those are gone! Okay. So it is just a different way of looking at the [problems]. When they [students] first check in the book, they go, “Gosh, where are the problems?” (Laughs.) And I go, “See that number right there? That

paragraph is the problem.” (Laughs.) (Mary interview; December 4, 2008)

Comparison between mathematics and statistics revealed Mary’s conceptions of mathematics as dualistic. The phrase “problems are not skills and drills” implies that the only way to learn mathematics is to master the skills by doing them over and over again. This approach seems to bring in the element of learning by “committing to memory,” which is typical of a dualistic perspective of mathematical knowledge according to Perry. The example “Simplify the square root of 50” seems implicitly to suggest that mathematics is just about carrying out procedures and getting the answer, which can either be wrong or right.

In another example, Mary presented a different view of her conception of mathematics. She indicated that she wanted her students to be able to use multiple approaches to solve problems. She specifically gave an example of solving quadratic equations:

Most [students] understand teaming⁶ with it or, yeah, most of them already know and can sing the little song. But that’s, you know they, that’s one [way]. It’s, and I tell them, I say that’s just one way; that is one way of solving a quadratic. And they, they are funny because I wanted to show them so many ways. And a lot of them say, “Why do you, why do you show us all these different ways?” And I say, “I want you to understand that you don’t have to use one method.” And perhaps one method doesn’t always work. Like factoring. Factoring may not always work. And so, you know, but you know, and they will find out. “Well, what, what always works?” (Laughs.) “The quadratic formula.” (Laughs). (Mary interview; December 4, 2008)

In this example, it is clear that Mary is communicating multiplicity in terms of the methods that can be used to solve mathematical problems, specifically quadratic equations. This evidence is seen in the following phrases: “It’s, and I tell them, I say

⁶ A song that people use as a mnemonic to remember the quadratic formula.

that's just one way" "I want you to understand that you don't have to use one method." One can also see Mary moving towards relativism, however, if one considers comment "But you know, and they will find out. . . Well, what, what always works?" This statement is an indication that Mary wants the students to make a choice among different methods of solving quadratic equations. The evidence presented here is not very clear as to whether Mary helped the students to see that the quadratic equation works all the time or she let the students discover that on their own.

In conclusion, one can say that through what Mary said in the conceptions interview, her view of mathematical knowledge has elements of both dualism and multiplism. Dualism is seen in her definition of mathematics problems as "drills" and "skill," and multiplism, in the multiple ways of solving mathematical problems.

Conceptions Related to Statistics

Mary's definition of statistics was a little difficult to understand. In the conceptions interview, she talked about the use of variables, and also the concept of causation in terms of variables.

I: How would you define *statistics* to someone who has never done statistics?

M: I can say, well, we can talk about the variables—. There's different parts of statistics. I say we started looking at relationships between [variables]—. Has [a] thing happened because another thing happened? (Mary interview, December 4, 2008)

Later in the same interview, however, she described statistics in terms of what it is used for. For example, she said, "Teachers use statistics to calculate the mean and the standard deviation of the students' scores in a test." She went on to say that calculating the standard deviation was a mathematical process.

According to Mary, statistics was made up of two parts. The first part involved the use of mathematics, and the second involved writing and explanations. She saw statistics as a subject with its own language that had to be used precisely. She explained:

The beginning of statistics—. We are looking at the probability—. That part of the statistics course is using your math skills. When you start writing it, when you start doing hypothesis testing where you have to watch for semantics and the way that you say things, like you, um, “You fail to reject.” It is always interesting to teach the kids that you cannot accept (laughs) a hypothesis. You choose to be free, you know, reject it or fail to reject. (Mary interview; December 5, 2008)

On page 84 above, Mary described statistics problems as not “skills and drills.” In the same interview, she went further to describe the nature of solutions and also the type of thinking involved in statistics:

It’s a different way of thinking that, you know, is not completely right or wrong. . . . I mean [for] some problems we wrote, there’s no right or wrong answer [as long as] they can defend it. Because it depends on “What do they think? How close is close?” For example, if you are looking at the percentage. You know, if a surgery or a new medicine is going to be 60% [successful], perhaps it worked on 60% of the people that tried it. Is that one good enough to be released by [the] FDA [Food and Drug Administration]? What if it is only 40% [successful]? Do you still want to release it because it has [worked on] 40% of [the] people that take it? So it is kind of interesting hearing their thought processes of—. And the kids have their opinions. Answers depend on whether somebody cares if it [the success rate] is better than 50% or less than 50%. So it is kind of neat hearing the kids. And that part of it has a little bit; it is a little bit [different from] the mathematics. (Mary interview, December 5, 2008)

By saying that statistics problems are not skills and drills, Mary was communicating a view of statistical knowledge that went beyond dualism. Statistics problems demand more than just performing skills and learning by committing to memory. We get a good picture of Mary’s position in the Perry (1970) scheme when she said, about statistics problems, “There’s no right or wrong answer [as long as] they can defend it.” That statement implied that statistics problems can have multiple solutions; therefore we can

say that Mary acknowledged multiplicity in the number of solutions. More evidence can be seen when Mary said, “And the kids have their opinions.” This statement is characteristic of Perry’s (1970) Position 4, in which there is acceptance of a diversity of opinions. From what Mary said about statistics, it appears that she held a view of statistical knowledge that was multiplistic.

Conceptions Related to the Difference Between Mathematics and Statistics

Mary communicated that she expected her students to have some basic mathematical skills because some of the problems they did in statistics required some knowledge of mathematics. Just as problem-solving skills were crucial in mathematics, the same skills could be utilized in statistics. She did not expect her students to be learning any mathematics in statistics. As already noted, Mary differentiated the two subject areas by looking at the problems and what it takes to get them right. Statistics problems were seen as demanding that one have a lot of information and requiring the students to come up with their own ways of solving the problems.

Conceptions Related to Assessment in AP Statistics

When I asked Mary what assessment in statistics meant to her, her first reaction was to talk about informal assessment. It was fairly obvious that her view of assessment involved a number of things, such as monitoring students’ progress while teaching and focusing on addressing students’ questions. Addressing students’ questions was probably the part of informal assessment that Mary took most seriously. Evidence can be found in this exchange:

I: What does classroom *assessment* mean to you?

M: It is an ongoing process. Literally, during, while you are teaching, [if] you are constantly monitoring the learning, you should be all right. Student teachers, for instance, that’s always a weakness because they

are so scared, I think. They are nervous about being out there. Their focus is, let me just teach this, you know. And, you know, they are not quite ready (laughs) to kind of watch the kids and monitor for that learning. And for, you know, that I know I am using that weird analogy, maybe it's because [our students] just had a student teacher. But watching [the student teachers] get bothered by being asked a question, you know, whereas I welcome a question, unless it is just the same question for the third time. "What page are we on?" (Mary interview, December 5, 2008)

Mary believed that assessment should be something that was useful to students and not just to keep them occupied or just for them to get a grade. Her experiences with assessment were partly shaped by what she saw some professors do in her college days. She described her experience as follows:

M: I would get very frustrated. If I felt like the professor or teacher would ask us to do something; do it like a project, [then] I would look at [it] and say, "There is no way I'm using this project!" So then I began thinking, "This is busy work." And to this day, I hate busy work. And so I think my teaching also shows [that] I hate busy work, because you know, I feel like if this is something that's going to take [the students] where they need to go, that's fine. But if it is something, "Oh, here is a project I want you to do." It's like, you know, there are so many projects that I did in undergrad, that I was, like, I did it for the grade.

I: They were not meaningful—in a nutshell.

M: No, no. They didn't have any meaning whatsoever for, other than I'm going to make that A. I mean that was, that was it. And [I] don't have [the projects] today, didn't use [them in my] first 10 years [of] teaching. I, you know, probably then will never use them. (Mary interview, December 5, 2008)

Mary believed in equipping the students with the ability to ask good questions in class. She said, "You know, as teacher you always look and see how good do the students do. Do they ask the right types of questions?" To achieve that goal, she indicated that she assigned students the odd-numbered exercises (those for which there were answers at the back of the textbook). She explained how she guided her students to ask good questions:

If you only give them [students] even [numbered] problems, some of them will change, you know. [They] don't have a clue if they got the answers right or not. So if they come in, and they have checked an odd-number problem, and they got it different, they will come in and say, "I didn't get what they had." And I'll say, "Well, tell me what you didn't get." We can kind of look and see where, if there was a mistake, is it something that they didn't understand? Or was it just some manipulation of numbers? (Mary interview; December 5, 2008)

Mary saw assessment as crucial in helping her to find out after teaching if the students "got it or not" before embarking on homework. She illustrated her point by using what she called a "ticket out the door" method. She was mainly concerned in making sure that her students had all the necessities to be able to do their homework.

And ticket out the door is simply, "Oh, here are two points. Find the distance, and find the midpoint." And then they hand it in. And that way, I know before they walk out and go home to do the homework; do they understand what we did in class? And that kind of lets me know back to back; who's got it and who doesn't before they go home and try the homework. (Mary interview; December 5, 2008)

Mary believed in giving her students the opportunity to work together collaboratively. She had a strong belief in the use of group quizzes. Citing one of the advantages of doing that, she said:

The conversations that go on as students are discussing the problems are helpful. Usually [I] like to get those heads together to be thinking and to talk more about the mathematics. And it is kind of neat to kind of walk around when they start doing it and listen to the conversations. You know, even if I'm not walking around, you can still hear the conversations that are going on between them. (Mary interview; December 5, 2008)

She also viewed the use of group quizzes as important because that allowed students to catch each other's errors and correct each other's mistakes and helped the students look at multiple approaches to solving problems.

Mary also saw group work as important because it helped those students who were struggling. To Mary, the process of working through the problem was more

important than the end of the problem, the answer. She clearly wanted to communicate that message to her students.

And so that was kind of neat: how they take on their own different roles within their little groups, of working the problems. And I do have some kids that are at the lower end, but I feel like they are getting something out of it, too. They are watching the people work the problems. They might not themselves be working them, but they are seeing the process going through; what the students are doing to work it. (Mary interview; December 5, 2008)

I was also interested in finding out whether Mary had a preference in assessment methods. She preferred those that incorporated fun, where students played games like Jeopardy. Playing Jeopardy allowed the students to work together in groups and got all of them involved. It was also important to help students' review before an exam. She said:

And so, and it is kind of neat because I have the kids get in groups, small groups, and let them pick their [group] name—. And the neat thing about that is that we actually get the whiteboards out, and they work on the whiteboards. Some of them [who] don't like whiteboards work on paper, but most of them pick whiteboards because they can write on it and erase real quick . . . if they mess up. So the kids are all working in their little groups of three or four. And you get, it is almost like everybody gets involved. And it is like a review. I do it right before a test, and so it is like a review session. And the kids, it is fun, the kids are competitive. They are competitive already, and it is kind of neat; it gets them into it. (Mary interview, December 5, 2008)

Among assessment items, Mary preferred what she referred to as "open-ended" items, because it was easier for her to go through the student's thinking process. She said:

I want to walk through [the solution], and I want them to show their work. And that's they have gotten, you know, they have gotten used to me. And I don't know who has taught them before; but they have gotten used to the fact that [they need to] work the problem, you know. And if you've worked the problem, and you got the mathematics there, I'm not going to take everything off. (Mary interview; December 5, 2008)

Mary also shared her opinion about the use of multiple-choice assessment items:

I used to never give multiple-choice tests. I can't stand making them out (laughs) because, you know, it is kind of like I want to work the problem and be done with it. That's my philosophy. But trying to make out a multiple-choice [test] and make out a good multiple-choice question is not easy. (Mary interview; December 5, 2008)

Conceptions of Assessment in Mathematics

In Mary's response to the question on assessment in statistics, I noticed that she did not distinguish between assessment in statistics and assessment in mathematics. She spoke in general terms, so I wanted to find out whether there was any difference between the way she viewed assessment in Honors Algebra II and AP Statistics. This was her response:

Assessment, well, probably not so much. I think [that] I'm probably more alert in [statistics], I'm probably more heavy with them [statistics students]; I don't do quite as much quizzing on the statistics as I do with the regular math class, simply because, number one, I don't have the time. We only have 90 days to teach this course—. This course used to be 180 days. It is shorter, it's not—; we've been cut in half, our time. It used to be taught in 180 days. And [in the] past 3 years, we've had to teach it in 90. (Mary interview; December 5, 2008)

Mary's Instructional Practices in Mathematics

Mary's Honors Algebra II classes followed the same routine each day. Normally the class would start with a review of the homework problems assigned in the previous lesson, a brief introduction of the day's topic, examples, and then class exercises. The mathematics class was highly interactive; the students felt free to interact with the teacher. She could at times joke with them. In most cases, Mary encouraged students to work in groups whenever she assigned some class work. The students seemed used to that, and every time she asked them to form groups, they did that with great ease. The idea of working in groups was consistent with Mary's belief that students benefit a lot

when they work in groups. Her teaching involved students in a number of ways, such as asking them to come to the board to solve problems or demonstrate solutions to the rest of the class. Students also used whiteboards to do some of the class exercises. Her approach to teaching involved guiding students through problems and constantly asking questions as they solved the problems. Mary appeared to be conversant and comfortable with the material she taught.

Mary always reminded her students to dwell on the understanding of the mathematical concepts instead of just applying procedures blindly. She made a few comments to that effect. For example, when the class discussed absolute values, Mary noticed that her students were not able to tell which of the problems had solutions and which did not. She gave the students two examples, $|x + 2| < -4$ and $|x + 2| = -4$, and asked, “What is happening to those problems that do not have solutions?” Students were not able to give a reason, so she asked, “Do you understand what the absolute value means? Think about the distance. Is it possible for distance to be negative?” She then commented, “We are so much hung up on the procedure, and we do not think about what the question is asking us about.” She noticed that in this case students were just applying the algorithm for solving linear equations.

Mary also seemed to believe that the best way to teach certain concepts was to highlight the students’ misconceptions. For example, in one instance she asked the students to simplify $-(3x - 4)^2$. Some students had $(-3x - 4)^2$ as an answer, whereas others said the answer was $(-3x + 4)^2$. Mary asked the students, “How did you get that?” The students replied that they had distributed the minus sign. Mary told them that it was better to square the parenthetical expression first before using the distributive property.

Mary valued students' contributions whenever the class discussed mathematical problems. Figure 10 shows a problem similar to the one that was discussed on January 14. Students had difficulties with this problem, which was assigned as a homework problem. Mary explained to the students that since this was to be a function in t , they had to replace z with t in the function containing C . The students were able to do that and came up with a function: $C(t) = 360t + 850$. After that, the students were able to finish the rest of the problem. Mary then extended the problem and asked the students this question: "What if you were asked to find how many candles you could make for \$2000?" One student volunteered to answer the question and explained that he wrote the expression as: $2000 = 360(t) + 850$. Together with the class, Mary found the value of t (i.e., $t = 3.19$), which was then used to calculate the number of candles by using the function $z(t) = 90t$ (i.e., $z(3.19) = 90(3.19) = 287.5$, which was rounded down to 287). One student indicated that he had solved the problem differently. Mary asked him to explain how he did it. The student said, "I just used $2000 = 4z + 850$, and solved for z , and I found 287.5." Mary was visibly happy and told the class not to just stick to one way of solving problems and that they could always choose the one that worked well for them.

Though in most cases Mary made sure that her students understood the logic behind most of the procedures, there was at least one instance where she introduced a procedure to students without offering any explanation. That happened when she introduced the concept of finding the inverse of a function. The class then looked at an example of how to find an inverse of a function: $t = 50 + 3v$. To find the inverse, Mary told the students that they needed to interchange v and t and solve for t . That is, they needed to solve $v = 50 + 3t$ for t . Mary asked the students, "What do you have to do

first?” One student replied, “Take 50 to other side of the equation.” Mary said, “Okay, go ahead and solve for t .” There was no further discussion after that.

The function $C(z) = 4z + 850$ closely approximates the cost of a daily production of a run of z candle stands. The number of candle stands produced is represented by the function $z(t) = 90t$, where t is the time in hours since the beginning of the production run.

- Give the cost of a daily production run, C , as a function of time, t .
- Find the cost of the production run that lasts 5 hours.
- How many candle stands are produced in 5 hours?

Figure 10. Example problem on functions similar to the one discussed on January 14

Mary's Assessment Practices in Mathematics

Informal Assessment

Mary mainly used questioning as part of her informal assessment. She also used class monitoring, which included observations like listening to students' conversations while they worked on problems

One of the things that Mary said assessment should achieve is to help her find out if the students had understood what she had taught them. One way of doing that, according to her, was by asking students a “what if” question. In that way students would be extended to think a little bit more. In the interview on assessment, she said the following:

I love the “what if” questions that you can get from a kid. . . . But I do like the “what if” questions [in which] the kids seem to have understood something that you've said or they studied and want us [to] apply it to maybe a different situation. Or what if you changed a little incident? Just one little incident in a situation or the equation or the problem or whatever it is. . . . And I love that because I say, “Well, let's see.” So it kind of opens the door to kind of, to explore more. (Mary interview, December 5, 2008)

She used that approach on a number of occasions in her Honors Algebra II class. An example is in the extension to the problem in Figure 10. She handled the extension extremely well, and she appeared to be at ease extending mathematical problems or giving counterexamples.

In most cases, Mary answered a lot of students' questions by asking them a series of questions that led them to come up with the answer themselves. Consider the following example: The students were given a question in a warm-up exercise that asked them to simplify $|3x - 3| > 2$. Some students wrote the answer as $x > \frac{5}{3}$ and $x < \frac{1}{3}$. Mary told them that they had to put *or* instead of *and*. The students asked why the solution had to be written as $x > \frac{5}{3}$ or $x < \frac{1}{3}$, and not as $x > \frac{5}{3}$ and $x < \frac{1}{3}$. Mary responded by asking them the following question: "Can you give me a number that is both $< \frac{1}{3}$ and $> \frac{5}{3}$ at the same time?" One student responded that it was not possible to have such a number. She then said that was the reason she put the *or*, because it was possible for x to take values less than $\frac{1}{3}$ or to take values greater than $\frac{5}{3}$, but not both at the same time.

Mary asked her students a lot of questions during the lesson, and there were many times when she could not wait for them to think about the question before she would ask another question or give a hint. Consider an episode in which Mary introduced the greatest integer function (GIF). She first discussed examples of the real-life applications of the GIF. She then asked the students a question: "Who else uses a step function?" The students did not answer the question immediately. Mary then asked the students

another question: “Anybody mailed anything lately? If you are shipping something, they use a range for the shipping costs.” Mary then said, “You would want me to use the greatest integer function because it means that I would always be rounding down.” (She did not ask the students whether in shipping they would round up or down.) She then said, “We wish the post office people would use the greatest integer function.”

Mary walked around the classroom to listen to students’ conversations, monitoring their work and also allowing them to answer each other’s questions. In her class, it was common to hear her say to the students: “I like it when y’all answer each other’s questions because maybe you will understand it better than me explaining.” Whenever students did the explaining, Mary listened carefully. She would then comment if the student doing the explaining got it right. She did not grade any class exercise in her Honors Algebra II class.

Formal Assessment

Mary used four main instruments for her formal assessment in mathematics. These were homework, quizzes, tests, and exams.

Homework. The textbook (Shultz et al., 2001) was the main source of most of the homework problems in Honors Algebra II. At times, however, Mary would get some problems from the Internet and also worksheets from a practice workbook accompanying the textbook. Sample homework problems similar to those Mary assigned from the student’s textbook are shown in Figure 11. Mary usually assigned homework problems at the end of the lesson. Sometimes when students had not finished their class work, those problems became part of the homework. Students were given on average 7 problems a day. They were expected to bring their completed homework to the next class

meeting. Mary would normally ask them whether they had done their homework. I did not see her look at the students' notebooks to confirm who had completed the homework. If there were problems that the students could not do, she would discuss them with the students. Sometimes she put the solutions on the board or a transparency for the students to make their corrections. She did not grade any homework. When I asked her why, she said, "I do not grade those. I want them to take personal responsibility for their learning."

- Use substitution to solve the system of equations. Check your solution
 $y = x - 5$
 $y = 3x + 4$
- INCOME Jane works 30 hours or fewer per week selling magazines and tutoring. She earns \$30 per hour selling magazines and \$20 per hour tutoring. Jane needs to earn at least \$600 per week.
 - a) Write a system of linear inequalities that represents the possible combinations of hours spent selling magazines and hours spent tutoring that will meet Jane's needs.
 - b) Graph the system of linear inequalities. Is the solution a polygon?
 - c) Find a point that is a solution to the system of linear inequalities. What are the coordinates of this point and what do the coordinates of this point represent?
 - d) Which solution in the solution region represents the best way for Jane to spend her time? Explain you think this is the best solution.

Figure 11. Sample homework problems similar to those Mary assigned in Honors' Algebra II.

Quizzes. Quizzes came from a number of sources including the teachers' resource book (Shultz et al., 2001) accompanying the textbook and also from the Internet. Figure 12 shows a sample of quiz problems similar to those Mary used in Honors Algebra II.

All the quizzes were given during class time and usually took 15 minutes to complete. All the quizzes were closed notes and closed book. Mary would normally

spend some time on the day of the quiz to answer students' questions before giving them the quiz. In one class observation while the students were reviewing for a quiz, Mary made the following comment to the class: "I do not normally put questions on the quiz we haven't gone over in the homework. Maybe I could if I wanted to check whether you have done homework." This comment suggested that Mary just wanted to put questions on quizzes that students had seen before or were familiar with. This practice brings in an element of memorization because the students could just memorize the solutions to these problems and do the quizzes without much or any understanding of the concepts. There was no formula sheets used for the quizzes. All the quizzes were graded, and the grades were recorded.

- Which of the following is an irrational number?

[A] $-\sqrt{25}$

[B] $\sqrt{28}$

[C] $-\frac{1}{\sqrt{4}}$

[D] $\sqrt{16}$
- Evaluate $3x + 4y + z$ for $x = 5$, $y = 12$, and $z = 7$.
- Find the maximum and minimum values, if they exist, of $C = 3x + 4y$ for each of the constraints.

$$3 \leq x \leq 8$$

$$2 \leq y \leq 6$$

$$2x + y \geq 7$$

Figure 12. Sample of quiz questions similar to those Mary used in Honors Algebra II.

Tests. Mary got most of the test questions from the textbook publisher's test book (Schultz et al. 2001), but sometimes she would write her own questions. Usually a day before the test, Mary would spend some time going over the topics that would be included in the test. She would normally go through practice problems from the textbook to prepare students for the test. On one occasion, she gave her students a warm-up

exercise with 8 questions in preparation for a chapter test. All the tests were closed notes and closed book. I never saw Mary give students formula sheets, but she indicated that normally if there was a formula that students needed, she would give it to them. The entire lesson was usually used for the tests. All the tests were graded, and the grades were recorded. Test questions similar to those Mary used in Honors Algebra II appear in Figure 13.

<ul style="list-style-type: none"> • Write each of parametric equations as a single equation in x and y. $x(t) = 3t + 8$ $y(t) = 5t - 1$ • An airplane is 5000 feet horizontally and 6500 feet vertically from where it must land. Every second the plane travels 230 feet horizontally and descends 30 feet. Write parametric equations that represent the plane's flight path as it approaches its landing point. How long will it take to touch the ground?
--

Figure 13. Sample test questions similar to those Mary used in Honors Algebra II.

Exams. The exams were mainly composed of the work covered in the term. The time I conducted the study was at the beginning of the semester, so there was no copy of the final exam.

Analysis of Oral and Written Assessment Items Used in Mathematics

Table 6 presents results of the coding of oral and written assessment items used by Mary in her Honors Algebra II class. Most of the assessment items for both oral and written assessments were in Group A (as described in chapter 2). Almost all the written assessment items required students to use routine procedures. Eleven percent of the oral assessment items were in Group C, but there were no written assessment items assessing thinking skills in either Group B or Group C. Clearly, the results show that written assessment items were dominated by items that required the use of routine procedures.

This dominance was to be expected because most of the items used in the quizzes or tests were similar to those used in homework assignments or discussed in class.

Table 6

Percent of Mary's Oral and Written Assessment Items Assessing Thinking Skills in Mathematics

Group	Description	Oral questions	Test and quiz items
Group A	Factual knowledge	27	0
	Comprehension	14	6
	Routine procedures	43	94
Group B	Information transfer	5	0
	Application in new situations	0	0
Group C	Justifying and interpreting	11	0
	Implications, conjectures, and comparisons	0	0
	Evaluation	0	0

A further look at the written assessment items revealed that none of the items was open ended even though Mary had indicated in the assessment interview that she liked open-ended items. It would appear that Mary might have viewed any open-middled questions as open ended. All the items in the tests were open middled. In the quizzes, there was a mixture of multiple-choice, true-false, and open-middled items. Mary did not seem to have criteria to determine the number of items of a certain format in the quiz; for example, in one quiz 85 % of the questions were multiple choice, whereas in another quiz only 23% of the questions were multiple choice.

Mary's Grading Process in Mathematics

Mary used four assessment instruments to determine the grades in the Honors Algebra II class. The distribution of percentages by assessment instrument is displayed

in Table 7. One interesting observation is that homework did not contribute anything towards students' grades. Mary did not explain the school requirement of the grade distribution if there was one. She did, however, explain that it was the school's policy that homework should not count for students' grades in Honors Algebra II and AP Statistics.

Table 7
Percent Distribution of Grades for Honors Algebra II

Instrument	Quiz	Test	Midterm	Final
Percent	20	60	10	10

Mary's Instructional Practices in AP Statistics

Mary's classes usually started with her going over questions from the previous lesson as well as assigned homework. Mary's approach in statistics mainly involved showing the students how different problems could be solved. At times, she would encourage the students to copy down some notes. She very much taught from the book and depended on it greatly. After seeing some examples, the students were given similar questions to try in class. Usually they worked in groups and would check their solutions with Mary. She would then discuss those problems that students had difficulties with.

Mary communicated to me that much as she wanted her students to understand some statistical procedures or application of formulas, she still wanted them to understand what the question was asking instead of blindly applying formulas. Though computations were important, she wanted the students to develop the ability to check their own work to see if their answers made sense. An example of Mary's approach can be seen in an episode involving the discussion of the problem shown in Figure 14.

A coin is tossed three times. If X is the number of heads you can get, find the distribution of X . What is the mean of X

Figure 14. Example of a problem on finding probability distribution.

The probabilities for the possible number of heads in three tosses of a coin were calculated in the previous lesson, so together with the class Mary put up a table on the whiteboard showing the distribution of X . The table is shown in Figure 15. In the previous lesson, the students had learned that the mean of a random variable could be calculated by using the formula $\mu_x = \sum x_i p_i$. The students applied the formula and found the value of μ to be 1.5.

Number of heads X :	0	1	2	3
Probability:	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Figure 15. Probability distribution of number of heads in three tosses of a coin.

After this example, Mary stressed the importance of understanding what was going on. She said:

When you [students] are doing this, I do not want it to be computational. I want you to understand what is going on. Use common sense. For example, if there were [three heads], then on average you would expect to have 1.5 heads and 1.5 tails. Is having a head independent of having a tail? These are independent and equally likely. This can also help you to check whether what you have in the calculator makes sense.

She then gave the students a similar question to try on their own using the approach she had just demonstrated.

At times, Mary would ask students to read certain sections of the book on their own during class time. On certain occasions, however, students would spend only a few minutes doing that. After that, there would be no real discussion. Mary would just give the students examples. For example, in one class, Mary had introduced the rules for

finding the means and variances of the sum and difference of two random variables. She wrote the rules on a transparency and asked students to copy that down. Immediately afterward, she asked the students to read the information in the book on the rules. There was no further elaboration of the rules; she just cautioned the students about them:

You should not apply the rules for means to the rules of variances, because they are not the same. This is the one [the variance of the difference] that a lot of students have missed in the past because they look at this as the difference. So they say $\sigma^2_{X-Y} = \sigma^2_X - \sigma^2_Y$. This rule of differences only applies to the means and not the variances.

The fact that Mary commented on the misconception regarding the use of the rule shows that she realized how important it was, yet it seemed she just glossed over it.

Mary's instructional approach depended on the topic she was teaching. There were situations in which she would go deep to the concept. At other times, she would just lecture to the students, simply telling them what they were supposed to do or not to do. Some examples of that approach are discussed below.

Mary's Assessment Practices in AP Statistics

Informal Assessment

Most of Mary's informal assessment was done by monitoring students while they solved assigned problems in class. As noted above, Mary said in the assessment interview that she always liked her students to work in groups. After giving the students the problems, she would then walk around to various groups to listen to their conversations and answer any questions related to the assigned work. Most of the questions she asked were linked to the problems or examples given in class, and some followed up on the activities done in class. In most cases, she just paraphrased the questions from the book.

Though Mary asked a lot of questions and also answered students' questions while going over problems, that task was made a bit difficult because of the way Mary managed the class. She had told me about some of her experience with college professors:

There were some [professors], I have to [say], and there were some professors that were easier to approach than others. So if you had difficulty with the problem, you could go. Or I felt like I had, they were more approachable than others. (Mary interview; December 5, 2008)

Mary created an atmosphere in her class where there was a good interaction with the students. I concluded that Mary's wish was to have a free atmosphere where students felt "at home" and not be intimidated to participate. She wanted to make herself approachable and not be like some of her college professors. However, some students might have taken that approachability a bit far. Sometimes when Mary was teaching, the students would just walk up to the front of the room to ask her a question while she was busy addressing another student's question. Sometimes the way students answered the questions was not orderly. It made it difficult for Mary to provide feedback to the students and sometimes to address their questions. A number of times some students were at Mary's desk, wanting to ask her a question, but she apparently did not see them, so she just addressed the questions of those who shouted. At times, however, she would tell a student, "You have answered or asked a lot of questions. Give others a chance."

Formal Assessment

Mary used four main instruments for her formal assessment in AP Statistics. These were homework, quizzes, tests, and projects.

Homework. All the homework problems came from the textbook (Yates et al., 2008). Mary always gave her students an agenda detailing what would be covered in a

given lesson and the homework problems. Also included were tests and quiz dates.

Sample homework problems similar to those Mary assigned in AP Statistics are displayed in Figure 16.

There was an average of five or six problems in each homework assignment. Students were expected to bring their solved homework problems to class. I never saw Mary check an individual student's work. Homework review was mainly dominated by Mary, with very little input from the students. The review was mostly Mary telling the students how they were supposed to respond to the problems. The students would then copy down the correct solutions. This procedure was not surprising, because in the assessment interview, Mary had communicated to me that she did not have enough time with her AP Statistics class. She said:

I used to do so many fun, cool projects, and I mean—. And you hope they get all the mathematics and all the statistics in it, but you know that they might not. So you kind of go back to the—. “Here are some problems in class to do, and here are some problems for homework to do.” And [you] just talk about how you [students] were supposed to have done it. And [you ask them], did you do it that way? And then maybe give them a little quiz to make sure that they've done it after they have done several different sections or whatever, to do an assessment that way. (Mary interview; December 5, 2008)

Though I have indicated that Mary dominated the review of homework, there were also some occasions when she would find out from the class if anybody had answered a particular question, and she would ask that student to come to the board and show the rest of the class how he or she solved the problem. Mary never collected any homework for grading. Here is one exchange she had with the students:

M: I have noticed that some of you are not doing your homework. That will drop your grade.

S1: Are you going to drop our grades?

M: No. You know what happens when you do not do your work.

After listening to that conversation, I asked Mary why she did not grade homework. She replied as follows:

I go through the problems with them in class, and they check their work. I do not give any homework grade. To me, homework is not punitive. I do not grade this class [AP Statistics] for effort. But I do that with my Math Support class. Whatever they do, I give them a grade because that is a different [class]. (Mary interview; January 15, 2009)

- A coin is tossed three times. There are 8 possible arrangements of heads and tails. For example, HHT means the first two tosses result in heads and the third toss results in a tail. All the 8 arrangements are (approximately) equally likely.

Write down all 8 arrangements of the results of the tosses. What is the probability of any one of these arrangements?

- Let X be the number of heads. What is the probability that $X = 2$?
- Find the distribution of X . That is, what values can X take, and what are the probabilities for each value?

- The table below gives the distribution of the number of basketball games young adults watch on TV every day. The probability distribution of the number X of games watched is as follows:

Number of games x_i	0	1	2	3	4	5
Probability p_i	0.03	0.16	0.30	0.23	0.17	0.11

- Calculate the mean μ_x and the standard deviation σ_x .
- Describe the details of a simulation you could carry out to approximate the mean number of toys μ_x and the standard deviation σ_x . Then carry out the simulation. Are the mean and standard deviation produced in your simulation close to the values calculated in (a)?

Figure 16. Sample homework problems similar to those from Mary's AP Statistics class.

Quizzes. All the quizzes Mary gave came from the quiz book (Yates & Starnes, 2008) accompanying the textbook. From the samples that were made available to me, it was clear that Mary just made copies of them without making any changes. At times, she

would combine two quizzes (i.e., covering two adjacent sections) and let students take that as a single quiz. Quiz problems similar to those Mary used in the AP Statistics class are shown in Figure 17.

1. A certain probability density function is made up of two straight-line segments. The first segment begins at the point $(2.5, 0)$ and goes to the point $(2.5, .5)$. The second segment goes from $(2.5, .5)$ to the point $(x, 2.5)$.

(a) Sketch the distribution function, and determine what x has to be in order to be a legitimate density curve.

(b) Find $P(2.5 < X < 3)$.

(c) Find $P(X = 2.5)$.

(e) Circle the correct option: X is an example of a (discrete) (continuous) random variable

2. A report says that the weights of 25-year old men have a mean of 75.5 kilograms (kg) and standard deviation 3.05 kg. There are 2.20 pounds in a kilogram. What are the mean and standard deviation in pounds?

Figure 17. Quiz problems similar to those Mary used in AP Statistics.

All the quizzes were given at the beginning of class. On the day of the quiz, Mary would make a quick review of the materials for the quiz. Though the quizzes were normally completed within 15 minutes, usually a half a lesson was devoted to the review and the quiz. All the quizzes were closed notes, closed book. Mary had two kinds of quizzes: an individual quiz and a group quiz that students took in pairs.

Before recording the grades for the quizzes, Mary would give the quiz papers back to the students to go over and also to check whether the grades on the papers were accurate. She would then collect the papers again and record the grades.

Tests. Sometimes Mary just used tests straight from the sample test book (Yates & Starnes, 2008). At times, she used questions from online resources and also from past examination papers. She never wrote her own test items. All her tests had two parts. Part 1 contained multiple-choice items, and Part 2 contained free-response questions. Sample test problems similar to those Mary used in AP Statistics are shown in Figure 18.

- IQ scores of 7th graders are normally distributed with mean = 100 and standard deviation = 15. What is the probability that a randomly selected will have an IQ score higher than 120?
 - a) 0.908
 - b) 0.815
 - c) 0.184
 - d) 0.092
- Suppose the Athens-Times asks as a sample of 200 Athenians their opinions on the quality of life in Athens.
 - a) Is this study an experiment? Explain why or why not.
 - b) Identify the sample and the population in this study.

Figure 18. Sample test problems from Mary's AP Statistics class.

All the tests were closed notes, closed book. A day before the test, Mary would use the whole period to review the material that would be on the test. She would mainly go over practice problems. Though I never saw Mary give the students a formula sheet, as is the case in the AP Statistics examination, she indicated to me that she would give such a sheet any time she felt it was necessary. Tests took the entire lesson (90 minutes). All tests were graded, and the grades were recorded.

Analysis of Oral and Written Assessment Items Used in AP Statistics

Table 8 presents results of the coding of oral and written assessment items used by Mary in her AP Statistics class. An analysis of the assessment items showed that the majority of them required the use of routine procedures. The percentage for written assessment items, however, was higher than that of the oral assessment items. Items in

Groups B and C were almost nonexistent in the written assessments, with only 4% in each category. For oral assessments, there were no items in Group B, whereas the percentage of items in Group C was 14%. Though small, this percentage was more than three times that of the written assessment items.

Table 8

Percent of Mary's Oral and Written Assessment Items Assessing Thinking Skills in AP Statistics

Group	Description	Oral questions	Test and quiz items
Group A	Factual knowledge	17	2
	Comprehension	26	11
	Routine procedures	43	78
Group B	Information transfer	0	4
	Application in new situations	0	0
Group C	Justifying and interpreting	9	4
	Implications, conjectures and comparisons	5	0
	Evaluation	0	0

I analyzed the written assessment items further to determine whether they were open ended or not. The results indicated most the items (over 95%) were not open ended. The only items that could be described as open ended involved the designing of simulations. Half of the items in the tests were multiple choice. However, there were no multiple-choice items in the quizzes.

Projects. The projects were mainly extended investigative tasks. On average, Mary gave 1 project per semester.

Mary's Grading Process in AP Statistics

Three assessment instruments were used to determine a grade in Mary's Honors Algebra II. The distribution of grades was decided by the school administration. Tests were the greatest contributors of the student grade, contributing 60%. The distribution of percentages per assessment instrument is displayed in Table 9.

Table 9
Percent Distribution of Grades for AP Statistics

Instrument	Quiz	Test	Project
Percent	20	60	20

Factors Influencing Mary's Assessment in Mathematics

A review of Mary's documents, responses from interviews, and classroom observations revealed that there were four factors that influenced her assessment practices in statistics. The factors were standardized tests, teaching gifted students, technology, and administration policy.

Standardized Tests

As already mentioned on page 92, Mary did not like multiple-choice items. Because her students were to take standardized tests, however, she was obliged to prepare the students for that.

Because they [administration] say, these kids need to practice what kind of test they are going to take. Okay. So fine, I will do it, you know. They [the students] take the SATs [SAT Reasoning Tests] or EOCTs [End of Course Tests]; and that's a multiple-choice [test]; so most of these kids are going to be taking one or the other if not both. (Mary interview, December 4, 2008)

Mary was not really sure, however, whether that testing practice had changed her teaching. She said, "Well, it is kind of interesting whether I change my teaching."

Classroom observations did not indicate any great emphasis on multiple-choice questions.

Teaching Gifted Students

Mary said that throughout her teaching career she had mostly taught gifted students (honors students). That experience in a way had influenced the way she looked at the purpose of assessment. According to Mary, the purpose of assessment was to find out “who has got that mathematical insight.” She indicated that it had made her happy over the years to learn that the students think differently. She explained:

I’ve been teaching honors forever, since like, probably over half of my teaching career I have taught honors students—. And that’s part of teaching the gifted because you really want to see who’s got that mathematical insight. [That] just makes you smile—. I have some that are just phenomenal students that just, you just look at them, and you think, “Gosh! You know, what a great mind!” We want to talk to them so that you can see how, what; how did your mind get to that answer? You know, because they just think differently. They just, there are some of the truly gifted mathematical students. Just, I mean, it is hard to explain to somebody. (Mary interview; December 4, 2008)

The type of tasks Mary would give to students would depend on whether they were the “bright students” or the “low-level students.” For example, tasks that involved a lot of reading would not be given to low-level students. She compared the tasks that she would give her AP Statistics students and those she would give her Algebra 1 or Mathematics Support I students:

I can’t imagine giving that task because my low-level students would have, number one, it had too many words. And then, yes, I [would] wind up having to kind of read it to them and tell them what the task is about. And then you probably would have only the ones that would say, “Hey, that kind of sounds like it would be interesting to do. I’ll get to play with the calculator.” But Lord knows what they might play on the calculator because they are just not directed. (Mary interview; December 5, 2008)

Technology

Technology had changed the way Mary taught and the areas of mathematics she concentrated on. For example, it changed the type of questions she asked students. Manipulations were no longer a way of assessing whether students had understood a particular concept. Technology provided the opportunity to ask deeper questions that required students to offer explanations. She said:

I want to get past the manipulations. They've done the manipulations. They were, where they are now, in that junior course, because they could manipulate. I said, "Now I want them to learn and say, what have you got? Okay, what is an asymptote? Why is this shape? Why does this graph look like it does? Well, why is this? You know, why is this [graph] so? Converging at this point but never touch—." And so I say, "It changes completely the way that you teach." (Mary interview; December 5, 2008)

School Administration Policy

Mary communicated to me that the school's administration had decided to ease up a little bit on the requirements for those students wishing to enroll in Honors Algebra II. Because of that decision, Mary stated that she had watered down the material that she taught to accommodate those students. It had also greatly affected the pace at which she covered the topics, and therefore, led to a decrease in the number of topics covered each semester. She argued that she had to adjust the type of questions she had used when she had the "really gifted students."

It was also the school's policy that homework assignments grades should not count towards students' final grades in Honors Algebra II. In this case, we can say that the administration dictated which assessment instruments the teachers could use to assign students' grades. However, it is also important to note here that Mary was also not a big proponent of using homework grades.

Factors Influencing Mary's Assessment in AP Statistics

There were a number of factors that influenced Mary's assessment practices in statistics. The factors were time, the external examination, the textbooks publishers' tests, students not doing homework, caliber of students, technology, and pedagogical content knowledge.

Time

Mary indicated that the school district had dictated that the duration of the AP Statistics course be shortened. Because she wanted more time with her AP statistics students, Mary had suggested that instead of following the A-B-A-B-A⁷ schedule, it would be better for the school to change that so that she could be teaching AP Statistics on Monday, Wednesday, and Friday, as is done in most colleges. However, that suggestion was not taken into consideration. Mary indicated that lack of time affected her assessment practices greatly.

Lack of time affected the way Mary conducted assessment. Because of lack of time, she would just give students problems to try by simply telling them what to do, instead of covering the material thoroughly and letting the students practice on their own. She was even contemplating giving her students take-home quizzes to make up for the lost time. She wondered:

Would it be nice to let them do them [the quizzes] at home and bring them in? Oh, definitely (laughs) it would be. But, and I still, there's time when I want them to do group quizzes, which would be, you know, a take-home quiz because those kids have study groups. (Mary interview; December 5, 2008)

⁷ A-B-A-B-A is a schedule where a course is taught every other day. For example, in one week the course is taught on Monday, Wednesday, and Friday. The following week the course is taught on Tuesday and Thursday. Then the cycle is repeated.

I also noticed on one occasion that Mary gave quizzes in a sort of a haphazard fashion. For example, Mary gave her students group quiz immediately after giving them an individual quiz. The conduct of the group quiz showed that Mary was feeling the pressure of trying to cover as much ground as possible. Going through her schedule, I noticed that she had not put a date for the group quiz, but she ended up giving it. In the schedule, she had indicated that she would give the students a quiz for Section 7.1, and the other quiz would be a quiz for section 7.2. But apparently, that was not done, so Mary might have thought that, to cover ground, she just had to give them both quizzes the same day. The second quiz (group quiz) did not go as planned, because instead of the students doing the quiz in their pairs, they spent most of the time asking Mary what to do. She ended up explaining most of it, and the rest of the questions were given as homework

External Examination

Most of the classes that I observed Mary teach, she spent a lot of time talking about the AP Statistics examination and giving students instructions on what they should do. She seemed to continually draw the students' attention to the importance of the AP Statistics examination. Examples of some of the things she said about the examinations are as follows:

- After the exam when I see you I will ask you, "How many multiple-choice questions did you attempt? How many did you get right? You should do Number 6 first." I know one student once came to me crying, "I did not do Number 6." Her friend said, "You should do Number 6 first because it a simulation."
- The last question in [the] AP [Statistics examination] is a simulation. They allow you 25 minutes. That is why I am going through it slowly. In simulations, make sure that your explanation is very clear to an ordinary person.

- Be careful the way you write. They grade at a table. They grade a lot of papers. They may not have time to look at your small writings.

Mary depended on the previous AP Statistics examination questions to form her quizzes and tests. The format of her tests mimicked that of the AP Statistics examination.

Mary indicated that the way points are awarded in the AP Statistics examination influenced the way she graded her students in AP Statistics class tests. Most of the time she did not use the raw scores to assign the students' grades. She said:

I want to be fair to my students and also to encourage them. For example, if you look at these grades then it means that any student below 69 has failed. This implies that a lot of students will fail. So I try to use something close to the AP score conversion. For example, a score of 69 under the AP conversion would not be a fail. (Mary interview, January 26, 2009)

Textbook Publishers' Tests

Remember that Mary indicated that she never wrote any questions for AP statistics. She very much depended on the teachers' resource book for all her tests. I noted above that Mary did not modify any of these questions. It is fair then to say that all in all, the publishers' tests had a great influence on her informal assessment practices.

Students Not Doing Homework

Mary communicated that she was having problems assessing students' abilities in statistics because some students were not doing their homework. Mary expressed her dilemma as follows:

Right now, I'm just having trouble with them. They are just kids, and they don't want to do anybody's homework, much less some statistics homework to do. And some of these kids will take, like I told you yesterday, some of these kids are taking three AP courses. And so that means they are all on A-B-[A-B] schedules, which means [that] if they are taking three AP courses, they could easily be taking three, six; they could be taking seven classes [at] one time. (Mary interview, December 5, 2008)

Mary often reminded the students that AP Statistics was a college-level class; therefore the students needed to put in more effort.

Caliber of Students

There was a difference in the way Mary viewed teaching with regard to teaching the AP Statistics “bright students” and the Algebra II and Mathematics Support 1 “low-level” students. This is what she said about AP Statistics students when they were given an exercise to complete:

You know, they [students] figured it out. I mean that was kind of—. I feel like in some instances I am more of a facilitator. “And here’s something I want you to try and see what you get.” And making sure they [students] get to that point, you know, just kind of heading them down, herding them, or heading them down that path. I’m not sure which one sometimes. But, um, I am a little bit more also—. I think, I feel like I am a little bit more relaxed teaching my Statistics class than I would be in [an] Algebra II class; definitely more than I am [in] Math 1’s poor class. So it is just a different chemistry in the classroom as well as the material. The material may be harder, but yes, you’ll be more relaxed with the kids. You can try to get them to understand, just understand the mathematics and not have to remember it or memorize it. (Mary interview; December 5, 2008)

Technology

Mary thought that there was no need to spend time teaching students to work on the computer, because they did not need one in the AP Statistics exam. Furthermore, she was not sure whether the school system could afford some of the statistical packages like MINITAB or FATHOM. However, she stated that applets from computers were very useful. She believed that although technology was available, it was important for her students to be able to solve problems manually and not become overly dependent on the technology, without understanding the basics. Mary indicated that the main benefit in terms of assessment was that technology allowed the students to do more problems. She said:

For AP Statistics, it [technology] allows them to look, to maybe [be] able to look at [a] problem, [to] see if I got it right. It is a tool. It is a

completely—. I hope [it] is their tool and not a crutch. I definitely want it to be a tool. But I do—. It also gives me opportunities for us to do additional problems. You know, I am always—maybe because I am so much older than most of the teachers—looking at it and saying, you know, “Will they get it in one [problem]?” Because [some] students [can’t] get it in one problem, you know. That’s—. I still don’t—. I haven’t answered that question for myself. If I just do one problem, will they really get it? (Mary interview; December 4, 2008)

Mary did not put any restrictions on the use of technology (e.g., graphing calculators) on the quizzes or tests that the students took in AP Statistics.

Teachers’ Pedagogical Content Knowledge

Mary had already indicated that teaching AP Statistics was a challenge and that she was still learning a few things. This learning was reflected in the way that she conducted her assessment in class when she depended so much on the textbook and would not deviate much from it. There were occasions in which she did not address some of the students’ questions fully, because she had to go back to her original solution to check or confirm the right answer. For example, on one occasion, the class was discussing a question from a test the class had taken. The question was on calculating probability of the union of two events, given the probabilities of two independent events. This was a multiple-choice question with 5 possible answers. A lot of students had missed that question. During the review, some students wanted to know why the solution was wrong. Mary just indicated that she would go back to her original answer to see if she had marked the wrong one on the answer sheet. She seemed to be extra cautious and wanted to make sure that her facts were right. Students usually struggle with the concepts of independent events and mutually exclusive events.

On some occasions, Mary’s comments were reduced to emphasizing procedures and not conceptual understanding. She seemed to struggle to make connections between

what the students had learned previously and the current topic. An example can be seen in the following episode: Mary had given the question similar to the one shown in Figure 19 to her students. This was an extra problem after she had noticed that her students were

The design of an electric circuit calls for a green resistor and a red resistor connected in series so that their resistances add. The components used are not perfectly uniform, so that the actual resistances vary independently according to Normal distributions. The resistance of green resistors has mean 115 ohms and a standard deviation of 3.0 ohms, while that of red resistors has mean 275 ohms and standard deviation 3.3 ohms.

- (a) What is the distribution of the total resistance of the two components in series?
- (b) What is the probability that the total resistance that the total resistance lies between 340 and 360 ohms? Show your work.

Figure 19. Example on the use of the rule of means and variances.

struggling in the group quiz with the concept of random variables. The next day, Mary started by asking the question, “How do we find the mean?” One student responded, “Add them together and divide by two.” Mary indicated that the student had to use the rule of adding means, which had been learned in the previous lesson. That is, “If X and Y are random variables, then $\mu_{X+Y} = \mu_X + \mu_Y$.” The students had looked at a similar question in the group quiz in which they were asked to find the mean of Y , where $Y = \frac{X_1 + X_2}{2}$, and X_1 and X_2 were individual observations. It is possible that the student might have thought that any time one has two variables, then to find the mean one simply adds the means of the variables and divides by 2. However, this misconception was not clearly dealt with. Mary did not discuss the difference between the mean of a set of observations and the mean of a random variable. At the end of the lesson, Mary indicated to me that the students had already covered the topic on finding means on a set of observations. In this case, Mary was limited to the rules the students were dealing with at that time and

could not link that knowledge to what the students had done previously to make a comparison between the two concepts. This type of scenario could explain why Mary was limited in ability to extend her questions to what was beyond the book.

Summary of Mary

Mary was very comfortable with both her mathematical content knowledge and her pedagogical content knowledge. It was very easy for her to offer explanations to students on the concepts they had problems with. She addressed students' misconceptions very well and could anticipate the problems students were to have with some of the concepts. She was able to give them numerous examples and also demonstrated multiple ways of solving mathematical problems. The situation was very different in AP Statistics. Though her content knowledge would seem sufficient, she did not appear as comfortable with her pedagogical content knowledge in AP Statistics as she did in mathematics, which seems to have had a very big effect on the way she taught AP Statistics. Mary was still building her confidence in teaching AP Statistics. Remember that at the time of this study, Mary had only taught AP Statistics for only 2 years. However, she was still enthusiastic about teaching the course.

CHAPTER 6

CROSS-CASE ANALYSIS

In this chapter, I present findings related to the research questions. The main focus of the chapter is to present a comparative analysis of the two case studies in relation to the research questions. The study investigated the assessment practices of high school teachers of mathematics who also teach AP statistics. To understand these practices, I addressed the following three questions:

1. What are teachers' conceptions of assessment in mathematics and AP Statistics?
2. How do teachers assess students in mathematics and AP Statistics?
3. What factors influence teachers' assessment practices in mathematics and AP Statistics?

Conceptions of Assessment

I identified two forms of assessment—namely, formal and informal—according to the activities done or the manner in which they were done. Both teachers, John and Mary, viewed informal assessment as less structured than formal assessment. It was something that could be done while teaching and could even be done outside the classroom, at least according to John. Formal assessment was seen as more structured. It was done at specified times. Mary and John viewed assessment in both mathematics and AP Statistics as a continuous process. It was therefore important for them to monitor the

students' learning as the lesson progressed. Assessment was not viewed as separate from instruction but rather was integrated into instruction.

Both teachers indicated that the main purpose of classroom assessment was to find out whether the students understood what they had learned. They both were interested in making sure that students did not just reproduce what they had learned but demonstrated the ability to apply their knowledge in new situations. To achieve that, Mary used what she called "what if" questions, whereas John believed that it was important to ask students to explain things in a number of ways.

The teachers viewed assessment as a shared responsibility between the teacher and the students. Both Mary and John believed that it was important to involve students in the assessment process. They believed that their students benefited more if they heard explanations from their fellow students than from other sources because the students spoke the "same language." In particular, John believed that such an approach was beneficial because it improved the students' communication skills, and also it helped him as a teacher to identify misconceptions. According to Mary, giving students group quizzes allowed them to work together. All in all, both Mary and John considered listening to students' conversations or explanations to be a vital way of assessing whether the students had understood the concepts.

Assessment, in their view, was not limited to the teacher asking questions. It also meant giving the students the opportunity to ask questions on whatever they did not understand. Mary viewed giving the students the opportunity to ask questions as one way of assessing their understanding of the material. Mary's explanation quoted on pages 88–89 concerned the importance of students' questions. John also hinted at the importance

of anticipating students' questions during instruction. He said, "So when you get in class, and somebody asks you some difficult question, or you get all off track, or it takes you in another direction, you know then your lesson plan is not good." Both teachers, therefore, believed that questioning was a two-way process in which both the teacher and the students could ask questions.

Assessment instruments should have certain characteristics to be considered useful. Both Mary and John believed that assessment instruments should be able to provide students with the opportunity to demonstrate their understanding of the concepts or the material learned. But at the same time, the instruments should enable the teacher to diagnose students' problem areas. Consequently, both teachers had strong reservations about the use of multiple-choice questions. According to John, multiple-choice questions allowed students to guess answers at times. Mary believed that multiple-choice questions were not fair to students, because when grading such questions she would look only at the answer and not the process used in getting that answer. Mary preferred "open-ended questions" because they allowed the students to work through a problem and show their work. John believed that alternative assessments such as projects or portfolios were very important because they assessed students' multiple skills, unlike paper-and-pencil tests. However, he used projects only in AP Statistics.

Both teachers indicated that they liked questions about real-life applications. For example, Mary said:

I like the kids trying to ask questions. And I like the problems that are interesting enough to make them think about their lives. Look at it not just to be a test, you know. Something about, you know, some medical thing that, that—. Well, medical problems are great. They seem to make [the students] think more of what their lives are now as well as them looking

on what their lives are going to be, especially in this time when we have all this energy [crisis]. (Mary interview; December 4, 2008)

As John said in the quotation on page 51, he preferred the GPS mathematical tasks because of their real-life applications.

The two teachers indicated that, for them, the purpose of assessment went beyond just assigning students grades. As noted earlier, Mary indicated that she did not want to do any task when she was in college just for a grade. She wanted any assessment to be meaningful. Although both teachers indicated that their role as a teacher was to prepare students for examinations, they saw the purpose of classroom assessment as not limited to achieving that goal. Their other goal was to make sure the students were able to transfer what they learned in class to the real world. For example, in AP Statistics, the teachers wanted to make sure that the students became “statistically literate” so that they would be not just consumers of information they saw in newspapers but also critical of the reports they read.

Although both teachers thought assessment should be used for holding students accountable for their learning, they did not believe in the use of assessment to punish students. The teachers had different rationales for their beliefs about the use of assessment grades. Mary did not believe in assigning homework grades. She viewed it as punishment if students did not get a grade when they did not do their homework. John, in contrast, believed in assigning homework grades. To him, what was important was to be reasonable and flexible in carrying out the process. The purpose of grading homework, in his view, was to hold students accountable but also to make sure that they were motivated to do their work. Both teachers wanted their students to be successful and did not want to do anything to jeopardize that.

Mary and John believed that assessment methods are always subject to review. Both of them were National Board Certified teachers. They indicated how their specialist degree course had really helped them to be reflective about their teaching and grading practices. They both believed that it was important for teachers to spend time looking at themselves critically in terms of their instructional and assessment practices in order to be able to make adjustments if needed.

Neither teacher made any distinction between his or her conception of assessment in mathematics and AP Statistics. They both spoke in general terms. Therefore, it was not possible to isolate their conceptions of assessment in mathematics and statistics. When asked specifically whether their views on assessment were different, they said no. But each indicated that there were factors that influenced their assessment practices in the two subject areas. These factors are discussed below in the response to Question 3.

In summary, it would appear that the teachers' conceptions of assessment were similar. Students were an important part of the assessment process. For all purposes of assessment, making sure that students understood what they had learned was probably its most important feature.

Assessment Practices

Informal Assessment. For both teachers, informal assessment in Mathematics and AP Statistics involved the following practices:

- Asking students questions in class that they could respond to orally.
- Allowing students to respond to fellow students' questions, with the teacher listening to their explanations.
- Responding to students' questions and providing verbal feedback on their progress.

- Walking around the classroom to monitor students as they worked on assigned problems.
- Answering students' questions by asking them follow-up questions, so that the students themselves would answer the questions they had asked.
- Allowing students to work problems on the board.

Questioning was identified by both teachers as the main way of carrying out informal assessment in both mathematics and AP Statistics.

Oral Questions

Comparisons of the oral questions the teachers asked revealed that both teachers relied on assessment items that required recall of factual information, comprehension, and use of routine procedures in mathematics and AP Statistics. For both teachers, over 80% of the oral items in mathematics and AP Statistics assessed Group A thinking skills (items that require recall of factual information, comprehension, and the use of routine procedures; Smith et al., 1996). In terms of subject matter, there appeared to be no difference between the two teachers in the use of items assessing Group A thinking skills. There was a heavy emphasis on the use of these items in both subjects. Items assessing Group B thinking skills (items that require information transfer and application in new situations) were largely absent in mathematics and AP Statistics for both teachers; the percentages of such items ranged between 0 and 5. And in the case of the items in Group C (items require justifying, interpreting, comparing and evaluating), subject matter did not make any difference. For Mary, percentages were small and almost identical in both subjects. In contrast, John did not use items requiring the skills in Group C in mathematics, but he did in AP Statistics, which indicated that subject matter determined whether he used questions measuring thinking skills in Group C.

Formal Assessment

For these teachers, formal assessment involved the use of homework, quizzes, tests, and projects (the last in AP Statistics only). For both, homework was the main assessment instrument they used daily in the mathematics and AP Statistics courses. The emphasis on homework was evident in both courses, where a lot of class time was spent going over homework problems. The textbook was the main source of the homework problems for both teachers and in both courses. The two teachers used the same book in AP Statistics. At times in their mathematics courses, the teachers used worksheets from the teachers' resource book, and sometimes Mary downloaded worksheets from the Internet. John gave his AP Statistics students reading assignments before the beginning of a new section or chapter. He did not do that in mathematics. There was no evidence that Mary assigned any reading homework assignment in AP Statistics or mathematics, though at times she asked students to read certain sections of the book in class. Review of homework was done daily. In Mary's classes, though the students participated in the review of homework in both mathematics and AP Statistics, Mary seemed to dominate most of the review in AP Statistics. Although John graded homework assignments in both mathematics and AP Statistics courses, Mary did not do that in either course.

Quizzes were second to homework in terms of the frequency of usage as an instrument of formal assessment. In AP Statistics, both teachers depended on the sample quizzes from the textbook's publishers; therefore, their use of quizzes was similar. In mathematics, both teachers wrote some questions and used others from the textbook publisher. It was difficult, however, to tell from the quizzes which questions were written by the teachers. They all seemed similar to those taken from textbooks. Only Mary gave

group quizzes. She did that in her AP Statistics class, in which she allowed students to use their notes and textbooks.

Both teachers used chapter tests, midterms, and final exams in mathematics; there were no final exams in AP Statistics. The AP Statistics tests were similar for the two teachers because they both used similar resources. Tests in mathematics were similar in that neither teacher used any multiple-choice questions. The teachers gave no clear rationale for the not using multiple-choice questions. One explanation could be that they did not like such questions, as they indicated. But also the resources they used (textbook publisher's tests) did not have multiple-choice questions, so they just used the type of questions that were available to them.

In terms of technology, specific instructions were given in situations where technology use was allowed or prohibited. For both teachers, all the tests in mathematics and AP Statistics were closed notes and closed book. Both said, however, that they could allow the use of formula sheets in both AP Statistics tests and mathematics tests. The teachers also expected students to be able to memorize some formulas in mathematics, such as the quadratic formula. The memorization of formula or statistical procedures was never mentioned by either teacher. It would appear that this decision was motivated by the fact that in the AP Statistics final examination, students are given formula sheets.

Projects were used only in AP Statistics. Both teachers acknowledged that they were only able to give one project in a semester because there was not enough time to do more than one. There were no projects for the mathematics classes. John said he wished he could do alternative assessment like projects, especially poster boards. That was not possible, however, for the reasons discussed below in the response to Question 3.

Written Assessment Items

For both teachers, the items used in the tests and quizzes in mathematics and AP Statistics were dominated by Group A thinking skills. The percentage of the items in this category ranged between 78 and 100%. There appeared to be no difference in terms of subject matter in the use of questions assessing Group A thinking skills. Items assessing Group B thinking skills were used minimally in both mathematics and AP Statistics by both teachers, with the percentages 4% or below. In John's case, the analysis indicated that he was more likely to use items requiring Group C thinking skills in AP Statistics than in mathematics (22% of the written AP Statistics items were in Group C compared with only 4% in mathematics). The subject matter determined the way he used items requiring Group C thinking skills.

Tests and Quizzes

The numbers of tests and quizzes given in both mathematics and AP are displayed in Table 9. On average, both Mary and John gave more quizzes in AP Statistics than they did in mathematics. On average, however, the number of tests John gave in AP Statistics did not differ much from that in mathematics. This similarity was consistent with what John said, quoted on page 52, about the number of tests he gave in the two courses he taught. On average, Mary gave more tests in AP Statistics than mathematics. The results presented here are inconclusive, however, when one takes into account the number of lessons considered. There are a number of variables that influenced the number of quizzes or tests given: For example, the length of chapters varied a lot, which affected the number of quizzes given.

Table 10

Frequency of Quizzes and Tests Used by Mary and John in Mathematics and AP Statistics

	Mary		John	
	Mathematics	AP Statistics	Mathematics	AP Statistics
Lessons	18	6	18	15
Chapters	2 ¼	1	2½	2 ½
Quizzes	2	2	2	4
Tests	1	1	2	2

Grading and Reporting Performance

Both Mary and John graded all the quizzes, tests, and projects. As discussed above in connection with Question 1, however, Mary did not grade any homework assignments in either her Honors Algebra II or her AP Statistics classes. Both teachers wanted to be fair with the grades of the students. They had a standard procedure of giving the students their graded papers before recording the grades. The teachers would then go over the tests and quizzes together with the students. While the students made corrections on their papers, the teachers would also encourage them to check the accuracy of the points awarded to them. Students were also given the opportunity to explain the reasons behind some of their responses in the tests and quizzes.

The teachers had different grading schedules. And for each teacher the grading schedules between mathematics and AP Statistics were also slightly different. I presented the grading schedules in chapters 4 and 5. In both subjects, performance on the quizzes and chapter tests was the main determinant of the students' grades. These accounted for 60 to 80% of the grade. John provided the rationale for his grading schedule in mathematics, which is given on page 61. Mary indicated that the school administration told the teachers that homework grades could count only in lower

mathematics classes such as Mathematics Support I. Both teachers used letter grades to report on students' performances in mathematics and AP Statistics. They differed, however, in the way they determined the cutoff points for the numerical averages. John used the cutoff determined by the school's administration. Mary's school did not post any cutoff points for the students' scores, so she determined her own. Mary would at times arrange all the students' averages in order, such as 54, 55, 57, 77, 79, 88, and 90. She would then consider what she called "natural" cutoff points; for example, 57 and 77, would be the cutoff points so that the grades 54, 55, and 57, would be grouped together. Similarly, she would put 77 and 79 in another group. These separate groups would then be assigned different letter grades. At times, Mary used the normal curve to determine the students' letter grades. I asked her whether she thought using the curve was helpful in determining the students' grades. Mary said:

I like to do the standard normal curve. I like to see where [the grades] fall within a certain standard deviation from the mean. And I want to see what the average of the class was. I would like to see and say, "What does this tell me about the kid?" (Mary interview; November, 2008)

She said that she normally looked at how many students were below and above the mean to decide how to convert averages to letter grades.

Overall Analysis of Assessment Practices

For this analysis, I used the *Assessment Standards for School Mathematics* of the National Council of Teachers of Mathematics (NCTM, 1995). In using the Assessment Standards, I wanted to critique the way the teachers carried out classroom assessment. In chapter 3, I presented an overview of the six assessment standards and the questions that I would focus on each standard.

Mathematics Standard

What mathematics is reflected in the assessment?

The Mathematics Standard advocates activities that encourage students to communicate about mathematics. John indicated that he wanted his students to communicate mathematically:

In my math class they have to explain their answer in context anyways so, you know. If they give me a rate and there is no unit, then I hand it back to them and say it doesn't make sense. It is not a ratio, it is a rate, you know. So we are trying to go the way of the [state] standards now. In Math 1 they are supposed to be able to communicate, and use multiple representations. So we are forcing them now to give us things. (John interview; November 24, 2008)

In the assessment interview, John mentioned the importance of mathematics beyond the classroom environment. For John, getting the right answer was not enough; he expected the students to be able to think about the applications of mathematics outside the classroom. Further evidence of the mathematics that John wanted his students to learn could be seen when he talked about the state standards: He hailed the standards because of their emphasis on real world applications. He said:

And especially with the new standards, is we want [the students] to be able to see how it works in real life and take a word problem and model it mathematically, and they come up with some solution and then, you know, put it back to a work place and say this is how we would do it. (John interview; November 24, 2008)

Mary indicated the type of mathematics she wanted her students to learn when she discussed why she liked free-response questions: "I like those kids to show work, you know. And I really like application a lot more than drill and practice." She talked about two kinds of students: those who can apply their knowledge and those who can only do routine problems. She said:

He is the kind of kid that I could put a little bit different scenario, and he would figure it out; whereas some of the other students will look at—. “Oh no, I don’t know how to work this problem. It doesn’t look like the one[s] I’ve already worked. (Mary interview, December 4, 2008)

This comment shows that doing routine problems was not enough for Mary; she wanted her students to move beyond that. It is clear that Mary wanted her students to be able to relate the learned material to new or unfamiliar situations.

Both teachers saw the importance of communication in mathematics, solving real-world problems, and applying it to new situations. Sometimes there was a disparity between what the teachers did or said in class during instruction and what they did during assessment. For example, both teachers talked about the importance of applications, but their assessment items did not assess application that much. Analysis of the oral and written assessment items in both mathematics and AP Statistics indicated that the majority of the assessment items assessed recall of factual information, comprehension, and use of routine of procedures. The *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989) requires that all aspects of mathematical knowledge and its applications be assessed. The standards include mathematics as problem solving, mathematics as communication, mathematics as reasoning, and mathematical connections. The two teachers in this study assessed mathematics only as communication and mathematical connections. They did not assess the other aspects of mathematical or statistical knowledge such as problem solving or reasoning.

How does the assessment engage the students in realistic and worthwhile mathematical activities?

Though both teachers wanted their students to learn significant mathematics or AP Statistics and move away from just applying procedures to solve problems, that was

not the case, as noted above. The students were not given problems that would present them with the opportunity to reason mathematically or formulate problems. The problems that students worked on or did during assessments would not be described as “worthwhile mathematical activities.” NCTM’s definition of “worthwhile” mathematical tasks (activities) is given on page 8. The problems that students solved in the quizzes or tests were similar to those covered in class or homework and did not involve any real problem solving. The *Curriculum and Evaluation Standards* (NCTM, 1989) indicates that “problem solving is much more than applying specific techniques to the solution of the classes of word problems” (p. 137). Overall, there were more activities set in realistic context in the AP Statistics classes than in the mathematics classes. The teachers had very little to do with this difference, which arose primarily from the nature of the AP Statistics content.

Learning Standard

How does the assessment relate to instruction?

The Learning Standard indicates that “assessment does not simply mark the end of a learning cycle. Rather it is an integral part of instruction that encourages and supports further learning” (NCTM, 1995, p. 13). John saw the need to involve students in the assessment process (see pp. 49–50). He believed that assessment and instruction could go together. Therefore, while assessing students, he could also teach them the things they had problems with to “support further learning.” In class, he used students’ responses to questions or their explanations to fellow students to drive instruction by focusing on their misconceptions.

Mary's definition of *assessment* presented in chapter 5 showed that she understood that there was connection between instruction and assessment. When Mary discussed the use of the Jeopardy game as a means of assessment, she talked about the use of assessment as a teaching tool.

There are pretty good [sides] to it [Jeopardy Game]—so they [students] can share with the rest of the class, “This is what I did to solve the problem.” And it is kind of neat because it becomes not a graded assessment—which not all assessments are—but it becomes a very good teaching tool. You know, like a review tool, [but it's also] a teaching tool for them. (Mary interview; December 5, 2008)

The individual case studies indicated that both teachers integrated their assessment with instruction. Evidence from classroom observations showed that the teachers used informal assessments such as oral questions, observations, and listening. The NCTM (1995) Assessment Standards describe these as “the primary source of evidence for assessment that is continual, recursive, and integrated with instruction” (p. 46).

One way of addressing the Learning Standard is to make sure that “mathematics activities are consistent with, and sometimes the same as, activities used in instruction” (NCTM, 1995, p. 13). When I discussed with them the role of technology in mathematics and AP Statistics, both Mary and John indicated that they focused on the use of the graphing calculator because that was what the students used in the AP Statistics examination. As already noted, in both the mathematics and the AP Statistics courses, calculators were used during instruction and were also available during assessment. Both Mary and John indicated that they were a bit reluctant to spend time on computers, as they would not be available during examinations.

How does the assessment allow the students to demonstrate what they know and can do in novel situations?

Both teachers' assessment allowed students to demonstrate what they knew as far as applying routine procedures or recall of factual knowledge is concerned, but the students were given little opportunity to demonstrate such knowledge in "novel situations." The overwhelming use of problems in the quizzes or tests that were very similar to the ones the students had seen or done in class defeated that purpose. This practice was, however, contrary to what both teachers said about the types of mathematical and statistical problems they preferred. They wanted their students to be able to apply the mathematics or statistics they had learned to new situations.

Equity Standard

What opportunities has each student had to learn the mathematics being assessed?

One thing that I observed in both schools was the availability of learning materials; for example, books and calculators. I discussed the details in chapter 3. Therefore, in terms of learning materials, all the students had equal access, and the materials were adequate to enable all the students to learn mathematics and AP Statistics.

In chapter 5, I described the way Mary conducted her assessment in mathematics, where students just raised hands but were not recognized. In a situation like that, the students that are not recognized might have felt left out and might not participate fully in future classroom activities. This practice, in turn, could affect the mathematics they learn. But as discussed earlier, there were a few occasions when Mary let the students know that they had to give other students a chance to answer questions. This comment was an indication that Mary may have discovered that her assessment practice was not

equitable. John made sure that all the students were given equal opportunity to participate in class. He specifically targeted those students he considered passive participants.

In some situations, the teachers did not give the students enough time to think about the questions they asked. This practice can have an adverse effect on the students' learning, most especially the slow learners. It could also have set a bad precedent, as sometimes students could just sit back when asked questions, knowing that the teacher would answer anyway if they did not. The *Professional Standards for Teaching Mathematics* (NCTM, 1991) discusses the role of the teacher in discourse: Teachers have to decide “when to provide information, when to clarify an issue, when to model, when to lead, and when to let students struggle with a difficulty” (p. 35). As noted previously, the teachers described students in AP Statistics, Precalculus, and Honors Algebra II as “bright students.” One would expect the teachers to give those students more time to grapple with the problems before any intervention.

The absence of group work in John's class could have been a problem to those students who benefit more when they work in groups. As I indicated earlier, John left it open for students to work in groups but did not really enforce that. For students to see real value of group work and benefit from it, the teacher would be better advised to make sure that the students are indeed working together in their groups. Mary did that in her classes.

The fact that the assessments were mainly dominated by timed paper-and-pencil tests could have been a problem to those students who do not do well when they take tests under timed conditions. I did not, however, see any student asking for more time during the quizzes or tests. In fact, it was not uncommon to see some students finish

taking the quizzes or tests before the allotted time was up. For practical purposes, it might not be possible to give students tests or quizzes in class that are untimed. I must also indicate that I never interviewed any students to learn about their preferred methods of assessment. The teachers indicated, however, that their students did not like multiple-choice questions in AP Statistics, because they found them difficult.

How does the assessment provide alternative activities or modes of response that invite each student to engage in the mathematics being assessed?

The Equity Standard says: “Because different students show what they know and can do in different ways, assessments should allow for multiple approaches” (NCTM, 1995, p. 15). One way of addressing that standard is to give students the opportunity “to do major pieces of work that are more elaborate and time-consuming than just short exercises, sets of word problems and chapter tests” (p. 36). The major pieces of work might include performance tasks, projects, and portfolios. As noted earlier, neither teacher gave projects in mathematics although they were both cognizant of the importance of projects. Both teachers gave projects in AP Statistics. Despite the reasons that the teachers gave for not doing projects in mathematics, it can be argued that they were compelled to do projects in AP Statistics because they were preparing the AP Statistics students for an external examination. The AP Statistics examination has an investigative task, so the projects were done to prepare the students for that task. In the mathematics courses, the students took an internal final examination, so the teachers had control over what could be asked in it. In short, projects did not play a role in the teachers’ classroom assessment practices in mathematics. Formal assessments were dominated by short quizzes and chapter tests. It is fair to point out here that both teachers

did allow students to use multiple ways to solve problems, which is indicative of an assessment that is equitable. John used this practice often in AP Statistics, and Mary did it in mathematics. My conclusion is that the assessment practices of the teachers in mathematics did not provide the students with many “alternative activities.”

Openness Standard

How do students become familiar with the assessment process and with its purposes, performance criteria, and the consequences of the assessment?

The Openness Standard indicates that students must be informed of “how they will be expected to demonstrate that knowledge and what the consequences of assessment will be” (NCTM, 1995, p. 17). John outlined his expectations in terms of students’ performance in statistics, for example, when he discussed the difference between mathematics and statistics and how students were selected to take the AP Statistics course. So it is clear that even before teaching the students, John made sure that students were well aware of what was expected of them:

There is reading, the answers to the questions aren’t one word, one number answers; they are always in context; and stuff that they are not used to that I make them aware that they will have to [get used to]. And some of them stick with it, and some go, “OK, never mind.” (John interview; November 24, 2008)

In class, John often mentioned how students were supposed to respond to questions in the quizzes, tests, or AP Statistics final examinations to obtain maximum points. Similarly, Mary spent a great amount of time in her AP Statistics class talking about what the students were expected to do in the AP Statistics examination. When Mary discussed why she did not like multiple-choice questions but liked open-ended questions, she outlined what she expected of her students:

I want to walk through [the problem], and I want them [the students] to show their work, and that's they have gotten, you know, they have gotten used to me. And I don't know who has taught them before; but they have gotten used to the fact that work the problem, you know, and if you've worked the problem, and you got the mathematics there, I'm not going to take everything off. (Mary interview, December 5, 2008)

In this case, Mary clearly outlined what she wanted her students to do and also what the consequences of doing that were.

Classroom observations, as discussed in chapters 4 and 5, showed that the teachers spent a lot of time doing review work before giving the students any formal assessment. During that time, both teachers notified the students about the material that would be covered in the quiz or test. Students were also presented with sample test materials from the textbook in preparation for the quizzes or tests. Both teachers presented their students with grading schedules, enabling the students to know the consequences of each assessment (see chapters 4 and 5). In brief, both teachers did a good job in preparing their students for any formal assessment.

How are the teachers and students involved in choosing tasks, setting or creating criteria, and interpreting results?

There was no evidence that the issue of involving students was addressed by either of the two teachers. All the tasks were chosen by the teachers. There was no apparent role played by the students in setting or creating criteria and interpreting results. This phenomenon could be due to the fact that the teachers did not use many alternative assessments, such as performance tasks. The only thing I noticed was that the teachers would encourage the students to show their work when solving problems. The teachers seemed to suggest that they were concerned with how the students solved the problems and not just the answer. This suggestion appears to be the “grading criteria” they said

they “discussed” with the students. I did not see the teachers give the students any rubrics on how they were going to grade their homework, quizzes, or tests. On one occasion, however, I saw Mary give her AP Students a rubric that she was going to use to grade their projects. The discussion that ensued after that was to help the students understand the rubric. The students did not have any input in terms of creating the criteria.

Inferences Standard

What evidence about learning does the assessment provide?

Both teachers were able to use assessment to determine what the students had learned. On a number of occasions, I asked the teachers about their observations after they had taught a lesson or given a quiz or test. For example, on January 22, Mary taught a lesson in AP Statistics on rules of means and variances. I asked her to give me a reflection of that lesson. She said, “They [students] still have difficulty with the meaning of variance. After going on a couple of examples about expected value, I still had a student ask me what *mu* (μ) meant.” Similarly when she gave a quiz on the rules of means and variances, which I discussed in chapter 5, after grading the quiz, she noticed that her students had not learned the rules the way she wanted them to, so she gave them another opportunity to do the problems again. This practice demonstrates that Mary used the evidence from the quiz to determine whether the students had learned or not. As another example from the mathematics course, on page 93, I discussed how Mary used certain types of questions to address students’ misconceptions. Such questions gave her evidence as to whether the students had grasped the required concepts fully.

On January 26, John had a lesson on complex numbers. He asked his students a lot of questions on the rules of exponents, which he was going to use to derive a formula for describing the pattern for powers of “ i .” The students’ responses to his questions led him to conclude that some of them still had problems with the properties of exponents and division with remainders. This practice implies that John’s assessment (questioning) provided him with the evidence that the students had not learned the concepts he wanted. Consequentially, he switched gears and started talking about rules of exponents.

Both Mary and John used item analysis to determine the items that most students had difficulties with in the quizzes or tests. I did not have access to students’ work, however, to see whether the teachers put any comments on their graded papers. I also could not establish whether, when the teachers looked at the students’ papers, they documented the specific errors students were making. I can only assume they did, based on classroom observations, where both teachers would mention the items on which students made errors.

What multiple sources of evidence are used for making inferences, and how is the evidence used?

From my discussion with John on his preferred instruments of assessment, he mentioned instruments other than tests such as projects and portfolios. Though John never explicitly said that he was looking for multiple sources of evidence to make valid inferences about students’ performance, I conclude that he understood that just using one form of assessment was not enough to measure the students’ performances.

Mary was also aware that there could be multiple ways of assessing students’ understanding of the material taught. For example, she lamented the lack of time to do

“cool projects.” She also discussed the use of group quizzes as one way of assessing students. She was also keen on listening to students:

And it's, it's, you know, having like, you know, looking [at] them [students] and having conversations with the students that—. Really and truly having them to talk to you about what they did. Most of the kids, you know what they know by hearing them say things and knowing they are not just, you know, telling you something back that you already said to them. (Mary interview; December 4, 2008)

As discussed already, both teachers said they believed that there were several ways of assessing students' learning. Both believed in the use of multiple assessment methods as a way of finding out whether the students had learned the concepts. As discussed above, however, only written assessments (quizzes, tests, portfolios) were used to determine students' grades. I was not able to determine whether the teachers had any documentation of the evidence they got from classroom observations. I did not see them record anything when they listened to students in class. There was evidence, however, that the teachers knew their students' abilities very well—the teachers knew the students who were struggling and would at times ask those students to respond to questions in class. This was more common in John's classes than Mary's classes.

Both teachers used the evidence they got from assessment results for a number of purposes. In some cases, the teachers had to reteach some concepts when they saw that the students did not do well in a quiz or did not answer some oral questions satisfactorily. Both teachers used both formal and informal assessment to gather evidence about students' learning and take appropriate action.

Coherence Standard

How is professional judgment used to ensure that the various parts of the assessment process form a coherent whole?

For this question, I considered the four phases of mathematics assessment as indicated in the *Assessment Standards for School Mathematics* (NCTM, 1995): (1) plan the assessment, (2) gather evidence, (3) interpret evidence, and (4) use the results. I did not have access to the teachers' lesson plans to see whether they planned their oral questions in advance. For AP Statistics, the teachers used publisher's tests and quizzes, so if there was any planning it was just picking a particular version of a quiz they liked. For John, in the mathematics course, it was just a question of pulling questions from the test bank. Though the teachers wrote some items in mathematics, it is difficult to tell how much planning went into that since the items were not noticeably different from those supplied by the textbook's publishers. In the analysis of oral and written assessment items, I found very little evidence of professional judgment in the selection of tasks. The teachers used items that assessed the same thinking skills. I have discussed above under the Inference Standard the evidence teachers gathered, interpreted, and used.

How does the assessment match its purposes with its uses?

Assessment Standards for School Mathematics (NCTM, 1995) puts the purposes of mathematics assessment into four categories: (1) monitoring students' progress, (2) making instructional decisions, (3) evaluating the students' achievement, and (4) evaluating programs. For this study, I was able to observe evidence of only the first two purposes.

I have discussed above how the teachers monitored student progress. On page 54, I discussed how John conducted questioning in class. One thing that I mentioned was how he acknowledged students' responses. In other words, John gave oral feedback to the students. As I noted, even if a student gave a correct answer, John would at times ask the student to explain. This practice is consistent with what the Coherence Standard stipulates: "Feedback must be part of an assessment system that gives students consistent messages about what mathematics is valued and legitimate ways to demonstrate that knowledge" (p. 34). Similarly, Mary gave students a lot of feedback when she walked around class to listen to their conversations while they worked in groups. She also wanted students to use appropriate methods to solve problems. Her questioning practices made it difficult, however, for her to provide timely feedback for the reasons I explained in chapter 5. Because I did not have access to students' work, it was not possible to establish whether students got any written feedback. Sometimes the teachers just made general comments about the things students were struggling with: for example, John said: "You guys are still having problems interpreting a confidence interval in context." And Mary said: "Most of you did not do well in this quiz."

I presented above information on how assessment was used by both teachers to make instructional decisions. Both teachers matched the assessment purpose with its use. In this study I could not establish, however, whether the use of assessment to monitor students' learning promoted the students' growth in mathematics or AP Statistics.

The analysis presented here, on the one hand, shows that there were a number of areas where the assessment practices of the teachers addressed the assessment standards; on the other hand, it shows that they were still some areas that needed improvement. In

some cases, the teachers seemed aware of shortcomings in the way they conducted assessment; they made the necessary adjustments. The teachers seemed unaware, however, of some of the issues with their assessment practices. For example, I do not think that the teachers realized that almost all their assessment items were assessing the same low-level thinking skills.

One can look at and critique many aspects of assessment practices using the assessment standards. In this chapter, I looked at only a few of those aspects. I think the analysis presented here, however, sheds light on how the teachers carried out assessment in mathematics and AP Statistics.

Factors Influencing Assessment

Standardized Tests

Both teachers acknowledged the influence of standardized tests and AP examinations on their assessment practices. Neither teacher liked the use of multiple-choice questions for reasons already discussed. Both John and Mary spent a lot of class time in AP Statistics class talking about examinations. In both Honors Algebra II and Precalculus, students took internal exams as final exams. Both teachers seldom talked about these exams in their lessons.

Textbook publisher's tests

The influence of the textbook publishers' tests was more apparent in AP Statistics for both Mary and John than it was in their mathematics courses. They both depended on extracting questions from the textbook publishers' tests. They did not modify the questions. The situation was slightly different in mathematics (Honors Algebra II and Precalculus). Though the teachers relied on the questions from the textbook publishers'

tests, they wrote some of the questions themselves or modified some of them. Both teachers seemed to have some flexibility with what they did in mathematics but seemed restricted in what they did in AP Statistics, where they followed the curriculum strictly.

Time

Time was identified by both teachers as one of the main factors that influenced their assessment practices. Lack of time forced them to adjust some of their practices. Insufficient time seemed to be an issue for both teachers in AP Statistics. Neither Mary nor John complained about a lack of time in their mathematics courses. To a certain extent, there was some flexibility in the amount of material they were expected to cover in mathematics. Insufficient time also meant that it was difficult for the teachers to design their own assessment items, especially in AP Statistics. Apart from quizzes and tests done in class, no other work done in class was graded. They claimed to have no time to grade the work done in either AP Statistics or mathematics classes. Most of the time during the lesson was spent giving students examples or going over homework problems.

Teacher's Expectations

The teachers' expectations as to what the students were capable of doing influenced the way the teachers carried out instruction and assessment. The teachers had higher expectations for the students taking AP Statistics and upper level mathematics like Honors Algebra II and Precalculus than for students taking other courses. Mary and John referred to the AP Statistics students as "bright kids" or "the brightest of the brightest" or "the gifted." The AP Statistics students were expected to be able to relate solutions to problems in context of the problem. Mary described the AP Statistics students as

independent learners who would figure out things on their own, whereas students in lower-level classes (Algebra I remedial) needed more guidance. John said the assessment for mathematics was sometimes reduced to a single number. Students in lower-level mathematics courses were not able to apply what they learned to newer situations. He described them as having no work ethic and not wanting any challenges. As noted in chapter 4, John indicated that he would not do any projects with the mathematics students, because most of them would not devote any time to producing a meaningful project. Both teachers regarded teaching Algebra I remedial students as hard work and said they needed to motivate those students more than the others.

Technology

The use of technology was crucial to assessment in both mathematics and AP Statistics classes. The main technology tool used in each class was a graphing calculator. The teachers indicated that the use of technology was beneficial in that it allowed the students to do many activities. Mary indicated that the fact that calculators made calculations quicker meant that the students would have additional practice with problems. John that indicated the use of technology in statistics was really important because time that would have been spent doing laborious statistical computations could then be used in the interpretation of statistical results. The teachers also indicated that that in general the availability of technology changed the type of questions that they would ask students. They would move away from questions that just involved computations in both the mathematics and the AP Statistics courses. Technology was used almost daily in AP Statistics, whereas in mathematics it was not used as much.

Though technology was seen as very important, both teachers indicated that they did not want the students to depend on the calculator; the students still had to understand certain things or be able to see the logic behind the computations so that they would be able to check their work if something went wrong. In both mathematics and AP Statistics, the teachers would demonstrate various ways of solving problems with and without the use of technology.

Pedagogical Content Knowledge

There was a remarkable difference between John and Mary in the way they carried out instruction and assessment, most especially in AP Statistics. The individual case studies showed the effects of the difference in the number of years that the two teachers had taught AP Statistics: John was more experienced than Mary. He was more likely than Mary to expand on the information he got from the textbook, use numerous examples or extend questions, and ask probing questions. Mary seemed to be very much restricted to what the textbook suggested. As a result, she did not expand much beyond the textbook questions. The fact that she did not connect very well the different concepts in statistics meant that her questions asked isolated facts. It is important to notice, however, that her formal assessments were not affected since both teachers used the same resources. Both teachers appeared comfortable with mathematics.

Subject Matter

Both Mary and John indicated that there was a difference between mathematics and AP Statistics when it came to assessment. Although both teachers focused on misconceptions, Mary discussed more misconceptions in mathematics than she did in AP Statistics. John did the opposite; he discussed more misconceptions in AP Statistics than

mathematics. Both John and Mary acknowledged that they liked teaching AP Statistics more than mathematics because AP Statistics had more interesting tasks.

Other Factors

There were other factors that influenced the teachers' assessment practices. They included aspects such as school or district policies. Because the schools discussed in this study were in the same county, the decision by the county to change the duration of AP Statistics course from 180 days to 90 days affected both schools.

At both schools, it was the policy that grades be assigned as A, B, C, and F. No student could be assigned a grade of D. At John's school there was a common test for all the algebra classes. Both teachers discussed how the decision by the administration to increase the numbers of students taking upper level mathematics meant that even those students who did not qualify to take the courses were still taking them. That meant diluting the content of the courses.

CHAPTER 7

SUMMARY AND CONCLUSIONS

The purpose of this study was to investigate the assessment practices of mathematics teachers who also teach AP Statistics. The main focus was to explore the teachers' conceptions of assessment in mathematics and AP Statistics, the way that the teachers carried out assessment in the two courses, and the factors that influenced the teachers' classroom assessment practices in the two courses. Two teachers who at the time of this study were teaching mathematics and AP Statistics participated in the study from November 2008 to February 2009. The two teachers taught in high schools in the same county. I collected data using three different methods: (1) interviews, (2) classroom observations, and (3) artifacts. Analysis of the data occurred throughout the data collection period. The first part of the analysis was done using inductive analysis (Patton, 1998). I used a number of frameworks to analyze the data. I used the Perry (1970) theory of intellectual development to understand the teachers' conceptions of mathematics and AP Statistics, the mathematics taxonomy framework by Smith et al. (1996) to analyze the tasks the teachers used for assessment, and the six assessment standards by NCTM (1995) to gain further understanding of the teachers' assessment practices.

The teachers' conceptions of mathematics and AP Statistics ranged from dualistic to relativistic. There was no difference between the teachers' conceptions of assessment in mathematics and AP Statistics. The main aim of assessment was to promote students'

learning. I identified two forms of classroom assessment: formal and informal. Formal assessment involved the use of homework, quizzes, and tests. Projects were used only in AP Statistics. Informal assessment involved oral questions, monitoring, and listening to students' conversations. Only written assessment items were used to assign students grades. The majority of the assessment questions (oral and written) in the mathematics and AP Statistics classes assessed recall of factual information, comprehension of factual information, and routine use of procedures. A number of factors influenced the teacher's assessment practices such as time, textbook publishers' tests, teachers' beliefs about assessment, and teachers' pedagogical content knowledge.

Research Related to This Study

The study focused on three research questions (see p. 10). Though there have been numerous studies on assessment practices of mathematics teachers, no study has looked at the assessment practices of mathematics teachers who also teach AP Statistics. Some of the findings in this study are similar to those from previous studies on assessment, whereas others are different. Below, I present the findings of study as they relate to the previous studies.

Teachers' Conceptions of Mathematics and AP Statistics

Thompson (1984) reported that teachers' conceptions about mathematics in one way or another play a significant role in shaping their instructional practices. This study confirmed that finding. John's conception of mathematics played a role in his instructional practices. For example, his conceptions of mathematics were mainly dualistic. His instructional approach did not involve multiple ways of solving mathematical problems. Similarly, Mary's conceptions of mathematics were multiplistic

and relativistic. In class, Mary discussed multiple ways of solving mathematics problems and also encouraged students to choose the methods they liked. Though Mary's conceptions of statistics were mainly multiplistic, her instructional approach involved a lot of "telling students" what to do. That practice is indicative of dualism. All these findings are consistent with what Kesler (1986) found. Kesler reported that sometimes teachers' conceptions of mathematics are not always congruent with their instructional behavior. Similarly, Thompson (1984) reported of the complexities of studying the relationship between the teachers' conceptions of mathematics and teaching.

Teachers' Conceptions of Assessment

A number of studies (Brown, 2004; Sanchez, 2002; Susuwele-Banda, 2005) have looked at teachers' conceptions of assessment. In his study, Susuwele-Banda found that teachers viewed assessment as tests. In contrast, the teachers in the present study viewed assessment as an ongoing process that has to be integrated with the lesson. Assessment occurred throughout the lesson and not just at the end of a topic or chapter. Unlike the teachers in Susuwele-Banda's study, the teachers in the present study had multiple ways of assessing students. In addition to using quizzes and chapter tests, the teachers in this study used informal assessment, such as questioning and monitoring students' progress by looking at how they worked on problems the teachers assigned as class work. They also used assessment to diagnose students' problems. I indicated already that the teachers did item analysis. Both teachers discussed the "teaching moments" that arise as a result of assessment. These teachers did not use assessment just to assign students grades; they used assessment for other purposes as well.

Though I have indicated above that both teachers looked at assessment in terms of students' accountability (see p. 124), a result consistent with Brown's (2004) finding, Mary presented a picture that was slightly confusing. Mary wanted her students to be responsible for their learning, which is why she never graded their homework. But at the same time she wanted the students to be held accountable for their learning. It would appear that Mary may have looked at grading homework as bad, if it meant that her students were to lose some points for not doing their homework. According to Brown (2004), this view might have made Mary look at assessment as irrelevant if it was just meant to hold students accountable. Mary did not say explicitly, however, that she viewed assessment as irrelevant if it was used to hold the students accountable. On the whole, the teachers in this study seem to agree with Brown's (2004) conception of assessment as improvement of teaching and learning. The conception of school accountability was not mentioned directly by either teacher. Both said, "We live in a test-driven society." That statement might indicate an awareness of the school accountability conception of assessment.

Characteristics of Assessment Items

Previous studies (Senk et al., 1997; Stiggins et al., 1989) have indicated that teachers ask mostly low-level questions. The present study yielded the same conclusion. The teachers in this study used oral and written assessment items that required a low level of thinking; namely, recall of factual information, comprehension, and use of routine procedures. As in Stiggins et al.'s study, the teachers in the present study ignored other thinking skills such as applications in new situations, justifying, and interpreting. Stiggins et al suggested that it might be difficult for teachers to guide students and

evaluate answers in evaluation questions. I did not establish in this study that teachers had the same problems with evaluation questions as these suggested by Stiggins et al. (1989). There were situations, however, when the teachers showed the ability to evaluate students' answers that were different from what they expected. For, example, John did that in his AP Statistics class when he discussed simulations. The absence of items assessing higher levels of thinking could indicate that the teachers never thought about classifying the assessment items the way I did and not necessarily that they did not know how to handle such questions.

Although Stiggins et al. (1989) found that there was a difference between oral questions and written questions in mathematics in terms of the thinking skills assessed, I found that there was no difference in skills assessment between the oral items and the written items the teachers used in their mathematics classes. Both oral and written items assessed the same low-level thinking skills: recall of factual information, comprehension, and use of routine procedures. The same was true with AP Statistics. John used Group C (items requiring justifying, evaluating, etc; (see Smith et al., 1996) thinking skills more in AP Statistics than he did in mathematics, but Mary did not. Unlike in the study by Senk et al. (1997), where they reported that only geometry teachers used questions that required justification, I could not determine from the data in this study the influence of subject matter on the thinking skills assessed.

The use of open-ended assessment items was almost nonexistent. Neither John nor Mary used open-ended items in mathematics. These results are consistent with those of Senk et al. (1997).

Assessment Instruments and Grading

The teachers in this study indicated the types of assessments they liked. They wanted assessments that gave students the opportunity to show their work (not multiple-choice questions). McMillan and Nash (2000) reported similar results. Although the teachers in McMillan and Nash's study indicated that objective items allowed students to memorize instead of understanding concepts, in the present study John indicated that such items allowed students to guess answers. I also found that the teachers did not like objective items for other reasons: They found constructing multiple-choice questions difficult, most especially coming up with appropriate distracters.

Studies by Ohlsen (2007), Sanchez (2002), and Senk et al. (1997) indicated that teachers depend on major examinations and quizzes as the main source of assessment of students' mathematics learning. These studies found that mathematics teachers tend not to use other forms of assessment such as oral presentations or team projects. Assessment of students' performance is mainly determined by written instruments (quizzes and tests). I found the same thing in the present study. All the students' grades in the mathematics courses and the AP Statistics course were determined by written assessments (tests, quizzes, and homework) only. Though the teachers in this study used informal assessments, they did not use any of those to determine the students' grades.

Factors Influencing Assessment and Grading Practices

A number of studies have reported factors that influence mathematics teachers' assessment and grading practices. Senk et al. (1997) reported about the influence textbook publisher's tests have on mathematics teachers' assessment practices. According to Senk et al., mathematics teachers depended on the textbook publisher's

tests as a source of their test questions. The findings of the present study are consistent with those findings. The teachers in this study depended on the textbook publisher's tests. As indicated already, this dependence was more evident in AP Statistics than in mathematics. As in the study by Senk et al. (1997) the teachers in the present study did not modify those tests. It is possible that the teachers did not want the questions in the AP Statistics class to look any different from those students would be asked in the AP Statistics examination. It could also be that the teachers did not have confidence in forming their own items in AP Statistics. Whereas McMillan and Nash (2000) and Ohlsen (2007) reported that teachers preferred to write their own tests instead of using textbook publisher's tests, the teachers in the present study did not explicitly state their preference. I got the impression that the teachers wanted to write some of the items on their own, especially in mathematics, and use some items from the textbook publisher's tests. John had said that writing AP Statistics test questions was very difficult because most of such questions were data driven.

Teachers' pedagogical content knowledge has an influence on the way they carry out assessment (Sanchez, 2002; Enderson, 1995). Findings in the present study support that conclusion.

Senk et al. (1997) reported that newer forms of assessment such as written reports or projects took time to prepare and grade and that discouraged teachers from using them. The same was true with the teachers in this study, most especially in AP Statistics, where they indicated that they needed more time to cover the material, so they resorted to just giving one project per semester. This finding is similar to that of Sanchez (2002), where teachers faced the dilemma between devoting their time to finishing the curriculum or to

using open-ended assessments. Time was not a factor for the teachers' failure to use projects in mathematics. Senk et al. (1997) reported how the students' attitudes made it difficult to implement some assessment strategies. The teachers in the present study had to deal with similar issues: John could not do projects with his Algebra I students because, he said, the students were not committed.

Limitations of the Study

The participants in this study were chosen by using purposeful sampling. They were chosen on the basis that they were teaching both mathematics and AP Statistics. The two participants were also members of an AP Statistics learning community. I have explained in chapters 4 and 5 how the two teachers started teaching AP Statistics. It might be difficult to extend the findings of this study to other mathematics teachers who also teach statistics as they may possess different characteristics. The findings of this study, however, can still enlighten us on the issues related to assessment practices of mathematics teachers who also teach AP Statistics.

The participants in this study came from the same school district. The schools might have had similar characteristics that are different from those of schools in other districts. I had initially decided to include a third teacher from a different school district, but for reasons beyond my control, it was not possible.

The study was conducted for 10 weeks, but I observed each teacher only six times in each course. This number of observations might not be enough to get a very clear picture of the teacher's assessment practices. A longer period would be beneficial—probably for the duration of the entire course (whole semester)—when the teachers would have covered a lot more topics. In a short period it is possible that the teacher might just

teach a few topics, and that can make it difficult on the part of the researcher to get a good picture of the way the teacher carries out assessment as he or she teaches different topics.

I reported on pages 35–36 about the teachers' busy schedules. Time was the limiting factor on what I would discuss with the teachers after observation. The informal interviews I had with the teachers were very brief, and I did obtain much information. It would have been better if the teachers had more time to really explain some of their practices after classroom observations.

Though I was able to get much information from interviews, observations, and field notes, I think it would have been better to audiotape the lessons I observed. I must indicate, however, that handling that much data would have demanded a lot more time and expense.

Implications for Teacher Education Programs

It is currently the case that only about a dozen states explicitly requires competence in assessment as a condition to be licensed to teach. Moreover, there is no licensing examination in place at the state or federal level in the U.S. that verifies competence in assessment. Thus teacher preparation programs have taken little note of competence in assessment, and the vast majority of programs fail to provide the assessment literacy required to enable teachers to engage in assessment for learning. (Stiggins, 2001, p. 5)

Despite the fact that studies on assessment have received much attention in the mathematics education community in recent years, results from the present study indicate that there are still a lot of areas of assessment that teacher education programs and professional development programs need to work on. Though there have been some changes from the time that Stiggins (2001) talked about assessment literacy, educators still have some way to go to understand the assessment practices of teachers.

There was a great difference between the preparations of the two teachers in the study. Though both teachers had attended the College Board training session, the teacher who had taken graduate level statistics courses seemed to have a better grasp of the AP Statistics content and was very confident. Considering that more statistics topics are finding their way into the secondary school curriculum, teacher education programs need to pay attention to statistics education. Most mathematics education programs for teachers offer a number of mathematics education courses in addition to mathematics courses. If mathematics teachers are expected to teach both statistics courses and mathematics courses, then statistics education deserves the same treatment that mathematics education gets. Education programs can help by requiring preservice teachers to take more courses that will help them with both statistical content knowledge and also pedagogical content knowledge in statistics.

Closely related to the training of statistics teachers is the issue of certification. Most states require mathematics teachers to get certification before they can teach mathematics. I have not found any evidence, however, of whether any states require teachers to get certification before they can teach statistics. If this is not happening already, then it might be helpful to require teachers to get certification in the teaching of statistics knowing that a lot of high school teachers of mathematics do not know how to teach statistics (Franklin et al., 2007).

Although the *Assessment Standards for School Mathematics* (NCTM, 1995) advocates the use of alternative assessments such as oral presentations, the present study shows that such assessments are still lacking. Apart from the fact that the teachers in this study indicated that they did not have enough time to use alternative assessments, it is

also possible that they did not see the real value of these newer methods of assessment. Teachers normally use what they value and think will be helpful to their students. Mary's episode (reported on p. 89) on why she did not like some of the assessment tasks she did in college clearly sends a message on the way teachers feel about what they learned in college. Perhaps teacher education programs ought to spend more time discussing assessment issues—to make sure that teachers are convinced that “new methods” of assessment are worth their time. If teacher education programs are modeling assessment techniques, then it is important to make sure that they present teachers with specifics that they can take with them to the classroom. Teachers need to see the uses or the values of assessments in practice. Just covering theory will not be helpful to them.

The teachers in the study indicated that they found making multiple-choice questions difficult. It is possible they did not have the necessary skills to write their own items. This raises the question, how much time are teacher education programs devoting to assessment literacy? It might be helpful to have such programs require all preservice teachers to take measurement courses. Collaboration between teacher education programs in mathematics and measurement programs could be helpful in making sure that what teachers learn in measurement is related to what they do in mathematics.

Closely related is the issue of assessment tasks. In this study, I found that the items teachers used for assessment were “low level.” As I have indicated, the teachers in the study did not seem to think of the items in terms of the thinking skills required to do them—considering that the teachers used items that measured the same thinking skills. According to the recommendations of *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), mathematics is more than just applying procedures to

routine problems: Mathematics also involves making connections, problem solving, communication, and reasoning. The fact that a number of researchers (Sanchez, 2002; Senk et al., 1997; Stiggins et al., 1989; Susuwele-Banda, 2005) have examined the nature of the items that mathematics teachers use and reached the same conclusion—teachers tend to use low-level items—could mean that teacher education programs are not spending time in addressing the issues involving the characteristics of the items teachers use in their assessments.

Professional developers also ought to spend a lot of time with practicing teachers to help them mainly in the development of tasks that teachers can use in the classrooms. Teachers need the necessary knowledge to enable them to develop or identify tasks that can measure all the thinking skills, instead of just concentrating on the recall skills. Abraugh and Brown (2006) found that engaging high school mathematics teachers to critically examine the mathematics tasks in terms of levels of cognitive demand influenced the way the teachers thought about the mathematics tasks they used. The study further found that some of the teachers changed the patterns of how they selected the tasks to use in the classroom. Sanchez (2002) also reported the effect of involving teachers in assessment projects. In-service teachers ought to be involved in assessment projects. That way they can acquire the necessary skills to develop their own tasks.

Finally, those who conduct teacher education programs ought to redouble their efforts in reaching out to practicing teachers to make sure that those teachers take courses on the new assessment techniques. The two teachers in the present study were members of an AP Statistics teachers learning community. They clearly already had an interest in upgrading their skills. Teacher development programs could take advantage of such

enthusiasm. In institutions where statistics, mathematics, statistics education, and mathematics education courses are taught in different departments, there is great need for collaboration to make sure that the teachers are equally well prepared in both the content and the education courses.

Recommendations for Future Research

The findings of this study have shown that there are still a number of things that need to be investigated concerning classroom assessment. Some of them are as follows:

1. Both teachers in the study indicated that they were not satisfied with some of the things they learned in college. Because these teachers were trained more than 20 years ago, research should be done on whether teachers who have been trained in recent years feel the same way. A study could be done to investigate how new teachers feel about the training they got in college. It would be helpful to track such teachers, to establish whether they think there is any alignment between what they were taught in college about assessment and what they find useful and practical in the field. Such a study could help teacher education programs in designing their courses.
2. The teachers in this study depended on publishers' tests. One factor was the issue of convenience and time. It is important to investigate, however, whether teachers have enough training to enable them to write their own tasks—tasks that measure various thinking skills and not just recall of factual information or low-level tasks. Are teachers required to take any measurement courses? Do these measurement courses discuss the characteristics of assessment items in terms of thinking skills? Does it make any difference if they do?

3. A study similar to the current study could be conducted on a large scale. The study could include different counties or different school districts. Several other variables could be considered such as using teachers who are initially trained to teach mathematics and statistics. The issue of the difference between mathematics and statistics has been discussed for a long time. Teachers in the present study were able to articulate the differences between mathematics and statistics. It is important to investigate, however, whether that knowledge of the difference translates into the identification of tasks. For example, are teachers able to distinguish tasks that assess statistical reasoning from those that assess mathematical reasoning? Can teachers develop their own tasks that will show the difference between items assessing mathematical reasoning and items assessing statistical reasoning?
4. The two teachers indicated that their students did not like multiple-choice questions in AP Statistics because they found them difficult. A study could be conducted to find out if in general AP Statistics students find multiple-choice questions difficult. Is it because of the emphasis of context in statistics? Is it because the students have difficulties comprehending the questions? Or could it be that students have difficulties in understanding statistical language? The two teachers did not indicate that students in mathematics complained about multiple-choice questions. Are multiple-choice questions in mathematics easier than those in AP Statistics?
5. The two teachers in this study had different backgrounds in terms of their preparation to teach AP Statistics. It is likely that many high school teachers who

are currently teaching mathematics and statistics are mathematics teachers first who are being asked to teach statistics. A study could be conducted to look into the preparation of teachers in teacher training colleges—to establish whether enough is being done to train them to teach both mathematics and AP Statistics. Furthermore, it is important to investigate how the statistics component is taught. Rossman, Chance, and Medina (2006) observed that though some institutions might already be doing some work related to assessment, the teacher preparation programs were only concentrating on the mathematical aspects of statistics.

6. The teachers in the study complained about lack of time in AP Statistics. This complaint came as a result of a school district policy. The argument was that 90 days were enough to cover the course (the teachers wanted 180 days). These teachers were clearly frustrated with the lack of time. Future studies should explore how school districts come up with such policies and how much input the teachers have who are teaching the courses.
7. The study raises a number of issues about homework that require further exploration. For example, should homework be graded? The two teachers differed in their beliefs about that. There is a need to investigate teachers' views about homework and grading. Should homework be graded for accuracy or completion?
8. This study looked at the assessment practices of mathematics teachers who also taught AP Statistics. Another study could be done to investigate the assessment practices of mathematics teachers who also teach ordinary statistics. Such a study could help us to understand how these teachers carry out assessment in the two

kinds of courses. In the present study, external examination had a big influence on the teachers' assessment practices in AP Statistics.

9. Since many statistics topics are finding their way in the high school curriculum, a study could be done to investigate whether there are any differences between the way teachers carry out assessment when they are teaching statistics topics and mathematics topics respectively.
10. In the present study I looked at how assessment is carried out in mathematics and AP Statistics. I concluded that the AP Statistics examination had an influence on the teachers' assessment practices. Another study could be done to compare the assessment practices of the teachers in two AP courses, such as AP Statistics and AP Calculus. Such a study could help us to understand whether there are any differences in teachers' assessment practices due to the difference in subject matter.

Closing Remarks

The study has shown that a number of factors influenced the teachers' assessment practices in mathematics and AP Statistics. Mary and John indicated that AP Statistics students were bright. Therefore, they could give the AP Statistics certain tasks (for example, tasks that required reading). They would not give such tasks to students in low-level mathematics, for example, Mathematics Support I. The difference in the type of tasks could therefore be attributed to the caliber of the students and not necessarily to the difference between mathematics and AP Statistics. Furthermore, the teachers said that they enjoyed teaching AP Statistics because it had interesting tasks. Probably if the mathematics courses they taught had interesting tasks, there would be no difference

between the ways they taught the two courses. Mary explained why she started by teaching ordinary statistics before teaching AP Statistics (see p. 83). It is therefore possible that the tasks that Mary gave to her AP Statistics students would not have been given to the students taking ordinary statistics even though both courses are statistics.

In AP Statistics, students took an external examination. As noted already, the examination had a strong influence on what the teachers did. Therefore, the way they assessed the students in AP Statistics had much to do with the examination and not necessarily the caliber of students. Mary taught Honors Algebra II, and John taught Precalculus. These two courses could be considered as upper level courses. Therefore, one would expect the teachers to give these students challenging tasks (at least at par with the AP Statistics tasks). The fact that the teachers did not do that clearly shows the strong influence of the external examination. There was no external examination for the mathematics courses. It is possible that if the two teachers were teaching two AP classes, say AP Statistics and AP Calculus, then teachers would have used similar tasks in the two courses.

In this study, I used a number of frameworks to help me understand the different aspects of the teachers' assessment practices. I have indicated already in chapters 2 and 3 why I used each framework for a particular purpose. Though that is the case, there is a connection among all the frameworks used. For example, on the conception of assessment, both teachers indicated that it was important to involve students in the assessment process—allowing students to answer each other's questions. This practice of involving students allows them to view the teacher in a different role. The teacher is not the only source of correct answers or knowledge. According to Perry (1970), under

multiplicity, learners learn to think independently because the teacher “the authority” does not just provide the answers to the learners. The two teachers also indicated in the assessment interview that they did not want to give students routine problems but problems that involved application. The teachers said that if assessment only involved the use of procedures then that would encourage the students to memorize these procedures without really understanding the mathematics or AP Statistics problems. Memorizing procedures is typical of dualism. One can therefore argue that the teachers’ conception of assessment reflected the conception of mathematical or statistical knowledge beyond dualism.

I used the Assessment Standards (NCTM, 1995) to analyze the teachers’ assessment practices and Smith et al. (1996) taxonomy to analyze the assessment items. I also discussed on page 8 why good tasks are very important. The type of mathematics or statistics tasks a teacher uses can reflect the conceptions of mathematical or statistical knowledge—the Perry’s (1970) framework can be used to analyze these tasks. I have reported already that items that assessed recall of information, comprehension of information and routine use of procedures dominated the assessment items in both mathematics and AP Statistics. The emphasis on these items was on using the right procedure and getting the right answers— characteristics typical of dualistic conception of knowledge as stated above. Therefore, the assessment items that the teachers used reflected the conception of mathematical or statistical knowledge that was dualistic. Lack of open ended assessment items meant that the students were not given the opportunity to approach mathematics or AP Statistics problems in multiple ways —the students did not have the opportunity to choose to or compare different methods of solving mathematics

or AP Statistics problems. Because of that, the assessment tasks did not reflect the other conceptions of mathematical knowledge such as multiplicity or relativism.

As noted already, Mary liked group work and listened to students conversations. Such an approach to assessment encourages students to come up with multiple ways of approaching mathematics or AP Statistics problems. John discussed multiple ways of solving AP Statistics problems. I can say that some of the teachers' assessment practices portrayed mathematical knowledge and statistical knowledge as multiplistic.

The purpose of this study was to investigate the assessment practices of mathematics teachers who also taught AP Statistics. Though a number of issues have been discussed in this report, it is obvious that one study alone cannot address many issues related to assessment. The findings of this study serve as a stepping stone to future research studies. Studies involving statistics have always involved college students. Now may be the right time to start thinking seriously about conducting more studies in middle schools and high schools. As noted on page 1, the GAISE project (Franklin et al., 2007) has been in the forefront advocating the improvement of the teaching of statistics in Grades PreK–12. The launching of the CCSS (see pp. 2–3) is a welcome development insofar as recognizing the need for more statistics topics in the high school curriculum. The NGO and CCSSO report shows a significant difference in the way statistics is viewed in U.S. state curricula. For example, in the Georgia Performance Standards (GPS), statistics appears together with mathematics such as Accelerated Mathematics I: Geometry/Algebra II/Statistics. In the current NGO and CCSSO document, statistics appears on its own as Statistics and Probability. This organization might signal that statistics is going to be treated as a course separate from mathematics that will be taught

independently. The organization is not clear from the document, however, because it still appears under mathematics standards. It also appears that the common core standards will cover more statistics and probability topics than the ones currently been offered by the GPS. The statement of support by ASA presented on pages 3 and 4 seems to suggest that there will be more statistics in the high school curriculum. The popularity of AP Statistics implies that there is a high demand for teachers who can teach AP Statistics.

The continual use of paper-and-pencil tests while ignoring other forms of assessment is worrisome. The fact that teachers are able to articulate the benefits of using open-ended assessments but are unable to support their claims with classroom assessment practice is something the educational community needs to take seriously and address. In line with the recommendations of the *Curriculum and Evaluation Standards* (NCTM, 1989), teachers need to make sure that their assessments are addressing all aspects of mathematical and statistical knowledge.

Assessment is dynamic. As statistics and mathematics content keep on developing, and as also technology is becoming increasingly important in the two subject areas, there is need to make sure that assessment keeps up with those developments. Studying assessment practices of teachers should be at the center of our education. More importantly, studies involving assessment in statistics are critical, as I have indicated earlier that statistics is a relatively new subject in the secondary school.

REFERENCES

- Arbaugh, F., & Brown, C. (2006). Analyzing mathematical tasks: A catalyst for change. *Journal of Mathematics Teacher Education*, 8, 499–536.
- American Statistical Association. (2005). *Endorsement*. Retrieved on May 20, 2008, from <http://www.amstat.org/education/gaise/ASAEndorse.htm>.
- American Statistical Association. (2010). *Statement of support*. Retrieved on November 8, 2010, from <http://www.corestandards.org/assets/CommonCoreStateStandardsforStatisticsProbability.pdf>.
- Brown, G. T. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education*, 11(3), 300–318.
- Bush, W. S., & Greer, A. S. (Eds.). (1999). *Mathematics assessment: A practical handbook for Grades 9–12*. Reston, VA: National Council of Teachers of Mathematics.
- Champion, R. (Winter, 2002). Sampling can produce solid results. *Journal of Staff Development*, 3(1). Retrieved on May 22, 2008 from <http://www.learningforward.org/news/jsd/champion231.cfm>
- Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, 10(3). Retrieved on July 10, 2008 from www.amstat.org/publications/jse/v10n3/chance.html
- College Board. (2007). *AP Statistics course description*. Retrieved February 13, 2008, from <http://www.collegeboard.com>.
- Copes, L. (1979). *The Perry development scheme and the teaching of mathematics*. Paper presented at the annual meeting of the International Group for the Psychology of Mathematics Education, Warwick, England.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. Thousand Oaks, CA: Sage.
- Dougherty, B. J. (1990). Influences of teacher cognitive/conceptual levels on problem-solving instruction. In G. Booker, P. Cobb, & T. N. Mendicuti (Eds.), *Proceedings of the Fourteenth International Conference for the Psychology of Mathematics Education* (pp. 119–126). Oaxtepec, Mexico: International Group

for the Psychology of Mathematics Education.

- Enderson, M. C. (1995). Assessment practices of preservice secondary mathematics teachers (Doctoral dissertation, University of Georgia, 1995). *Dissertation Abstracts International*, 56, 3040A.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Schaeffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK–12 curriculum framework*. Alexandria, VA: American Statistical Association. (Also available at <http://www.amstat.org/>) (First published in 2005)
- Gal, I., & Garfield, J. (1997). Curricular goals and assessment challenges in statistics education. In I. Gal. & J. B. Garfield (Eds.), *The assessment challenges in statistics education* (pp. 1–13). Amsterdam, The Netherlands: IOS Press.
- Garfield, J. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education* [Online], 2(1). Retrieved July 8, 2010, from <http://www.amstat.org/publications/jse/v2n1/garfield.html>
- Garfield, J. (1995, April). How students learn statistics. *International Statistical Review*, 63(1), 25–34.
- Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Needham Heights, MA: Pearson Allyn & Bacon.
- Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, 38, 427–437.
- Helms, J. M. (1989). Preservice secondary mathematics teachers' beliefs about mathematics and teaching of mathematics: Two case studies (Doctoral dissertation, University of Georgia, 1990). *Dissertation Abstracts International*, 51, 440A.
- Kesler, R., Jr. (1986). *Teachers' instructional behavior related to their conceptions of teaching and mathematics and their level of dogmatism: Four case studies* (Unpublished doctoral dissertation, University of Georgia, 1985). *Dissertation Abstracts International*, 46, 2606A.
- McGalliard, W. A., Jr. (1983). Selected factors in the conceptual systems of geometry teachers: Four case studies. (Doctoral dissertation, University of Georgia, 1982). *Dissertation Abstracts International*, 44, 1364A.
- McMillan, J. H., & Nash, S. (2000). *Teacher classroom assessment and grading practices in decision making*. Paper presented at the 2000 annual meeting of the National Council on Measurement in Education, New Orleans, LA.

- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.
- Merriam, S. B., & Associates. (2002). *Qualitative research in practice*. San Francisco, CA: Jossey-Bass.
- Meyerson, L. N. (1978). Conception of knowledge in mathematics: Interaction with applications to a teaching methods course (Doctoral dissertation, State University of New York, Buffalo, 1977). *Dissertation Abstracts International*, 38, 733A.
- Moore, D. S. (1988, January). Should mathematicians teach statistics? *College Mathematics Journal*, 19, 3–7.
- Mulhall, A., & Taylor, P. (1998). *Using participatory research methods to explore the learning environment of rural primary school pupils* (PLA Notes 33). London: International Institute for Environment and Development.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (1991). *Professional standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM), National Council of Supervisors of Mathematics (NCSM), Association of State Supervisors of Mathematics (ASSM) and Association of Mathematics Teacher Educators (AMTE). (2010). *Statement of support*. Retrieved on November 8, 2010, from http://www.corestandards.org/assets/k12_statements/NCTM-NCSM-ASSM-AMTE-joint-statement.pdf.
- Ohlsen, M. (2007, Fall). Classroom practices of secondary school members of NCTM. *American Secondary Education*, 36(1), 4–14.
- Owens, J. E. (1987). *A study of four preservice secondary mathematics teachers' constructs of mathematics and mathematics teaching* (Doctoral dissertation, University of Georgia, 1987). *Dissertation Abstracts International*, 48, 588A.
- Patton, M. Q. (1988). *Qualitative evaluatative methods*. Thousand Oaks, CA: Sage.

- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Perry, W. G. (1970). *Forms of ethical and intellectual development in the college years: A scheme*. New York, NY: Holt, Rinehart, & Winston.
- Perry, W. G. (1999). *Forms of ethical and intellectual development in the college years: A scheme*. San Francisco, CA: Jossey-Bass. (Original work published 1970)
- Punch, K. F. (1998). *Introduction to social research: Quantitative and qualitative approaches*. Thousand Oaks, CA: Sage.
- Quellmalz, E. S. (1985). Developing reasoning skills. In J. R. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 86–105). New York, NY: Freeman.
- Rapaport, W. J. (2006). *William Perry's scheme of intellectual and ethical development: A journey along the 9 "Perry" positions (as modified by Belenky et al. 1986)*. Retrieved April 13, 2008, from <http://www.cse.buffalo.edu/~rapaport/perry.positions.html>
- Romberg, T. A. (1992). Perspectives on scholarship and research methods. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 49–64). New York, NY: Macmillan.
- Rossman, A., Chance, B., & Medina, E. (2006). Some important comparisons between statistics and mathematics and why teachers should care. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance* (Sixty-eighth annual yearbook of the National Council of Teachers of Mathematics, pp. 323–333). Reston, VA: NCTM.
- Sanchez, W. M. B. (2002). Conceptualizing mathematics teachers' use of open-ended assessment items (Doctoral dissertation, University of Georgia, 2001). *Dissertation Abstracts International*, 63A.
- Senk, S. L., Beckmann, C. E., & Thompson, D. R. (1997). Assessment and grading in high school mathematics classrooms. *Journal for Research in Mathematics Education*, 28, 187–215.
- Schultz, J. E., Ellis, W. Jr., Hollowell, K. A., & Kennedy, P. A. (2001). *Annotated Teacher's Edition: Algebra 2*. AustinTX: Holt, Rinehart & Winston.
- Smith, G., Wood, L., Coupland, M., & Stephenson, B. (1996). Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal for Mathematics Education, Science and Technology*, 27(1), 65–77.

- Stiggins, R. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement*, 22(4), 271–286.
- Stiggins, R. J. (1998, January). Revitalizing classroom assessment: The highest instructional priority. *Phi Delta Kappan*, 69(5), 363–368.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5–15.
- Stiggins, R. J., Griswold, M. M., & Wikelund, K. R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement*, 26(3), 233–246.
- Stonewater, J. K., & Opera, J. M. (1988). An analysis of in-service teachers' mathematical beliefs: A cognitive development perspective. In M. J. Behr, C. B. Lacampagne, & M. M. Wheeler (Eds.), *PME-NA: Proceedings of the tenth annual meeting* (pp. 356–363). DeKalb, IL: Northern Illinois University.
- Susuwele-Banda, W. J. (2005). Classroom assessment in Malawi: Teachers' perceptions and practices in mathematics (Doctoral dissertation, Virginia Polytechnic Institute, 2005). *Dissertation Abstracts International*, 66, 123A.
- Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, 15, 105–127.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York, NY: Macmillan.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67, 223–265.
- Yates, D. S., Moore, D. S., Starnes, D. S. (2008). *The practice of statistics: TI-83/84/89 graphing calculator enhanced* (3rd ed.). New York, NY: Freeman.
- Yates, D. S., & Starnes, D. S. (2008). *The practice of statistics platinum binder: TI-83/84/89 graphing calculator enhanced* (3rd ed.). New York, NY: Freeman.

APPENDIX A
INTERVIEW PROTOCOLS

Interview 1

A. Conceptions of statistics

1. Why did you decide to teach statistics?
2. Briefly describe the courses you took in college and state whether that helped prepare you to teach statistics?
3. How do you see your role as a statistics teacher?
4. How would you define *statistics* to anyone who has never done statistics?
5. How would you define *someone who knows statistics*?
6. How much mathematical background do you expect your students to have before they can take your statistics course?
7. Some people argue that statistics has too many formulas. Do you think students should be expected to memorize these formulas? Explain.
8. What role does technology play in the teaching of statistics?
9. What is the role of mathematics in statistical learning?

B. Conceptions of mathematics

1. Why did you decide to teach mathematics?
2. Briefly describe the courses you took in college and state whether that helped prepare you to teach mathematics?
3. How do you see your role as a mathematics teacher?

4. How would you define *mathematics* to anyone who has never done mathematics?
5. How would you define *someone who knows mathematics*?
6. Do you think students should be expected to memorize the formulas that are used in mathematics? Explain.
7. What role does technology play in the teaching of mathematics?

C. Statistics and mathematics

1. Between mathematics and statistics, which one do you like better? Why?
2. Is there any difference between mathematics and statistics? Explain.
3. How do you reconcile teaching the two subjects?

Interview 2

Assessment in mathematics and AP Statistics

1. What does *classroom assessment in statistics* mean to you? Is there any difference between the way you view assessment in the mathematics class and the statistics class? Explain.
2. How do you assess your students in statistics formally and informally?
3. What challenges do you face in implementing an assessment practice?
4. What are the factors that shape your assessment and grading practices in statistics and mathematics? (What are the challenges you face?)
5. What are your favorite assessment methods? Explain why you like them.
6. Is there any difference or similarity between the way you assess students in mathematics and the way you assess them in statistics?

7. What factors do you take into consideration when deciding which classroom activities to grade?
8. What is your grading scheme? Explain the rationale behind it?
9. What do you rely on mostly to develop your assessment items?

APPENDIX B

EXAMPLES OF MATHEMATICS AND AP STATISTICS ASSESSMENT TASKS

FOR EACH TAXONOMY

Where 1: *mathematics problem* and 2: *AP Statistics problem*.

Group A

Factual knowledge and fact systems

1. State the Intermediate Value Theorem.
2. What is an unbiased statistic?

Comprehension

1. Determine which equation represents y as a function of x ?
 - A. $y = 2x^2 + x - 8$
 - B. $2x^2 + y^2 = -8$
 - C. $x = -8y^2$
 - D. $x = 10$
2. Suppose the Atlanta Journal asks as a sample of 200 Atlanta residents their opinions on the quality of life in Atlanta. Is this study an experiment? Explain why or why not. Identify the sample and the population in this study.

Routine procedures

1. Write each of parametric equations as a single equation in x and y .

$$x(t) = 4t + 8$$

$$y(t) = 6t - 1$$
2. A random sample of 1200 teenagers (ages 12 to 17) was asked whether they played watched football on television; 885 said they did. Construct and interpret a 99% confidence interval for the population proportion p . Follow the Inference toolbox.

Group B

Information transfer

1. If $f(x) = x^2 + 1$
 - I. Find all the zeros of the function $f(x)$.
 - II. Find the x-intercepts for the graph of $f(x)$, or explain why they do exist.
2. Peter and Isaac play are playing in a golf tournament. They play repeatedly and their scores vary. Peter's score X has the $N(100, 7)$ distribution, and Isaac's score Y has the $N(95, 9)$ distribution. Is it reasonable to take the variance of the total score to be $\sigma_x^2 + \sigma_y^2 = 7^2 + 9^2 = 130$? Explain your answer.

Application in new situations

1. Solve the following equation: $x^2y^3 + 5xy - 50 = 0$ by showing that the substitution $a = y^3$ and $b = 5y$ transforms the equation to the one which is quadratic in x .
2. The general addition rule of events states that for any two events A and B ,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
 - i. Use Venn diagram to prove the result.
 - ii. Prove that the addition rule for any three events A , B , and C is given by:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Group C

Justifying and interpreting

1. Two students evaluate the expression $-(3 - x)^2$

The first student writes the following:

$$-(3 - x)^2 = -[(3 - x)(3 - x)]$$

$$\begin{aligned}
&= (-3 - x)(-3 - x) \\
&= (-3)(-3) + (-3)(-x) + (-x)(-3) + (-x)(-x) \\
&= 9 + (-3x) + (-3x) + (-x)^2 \\
&= 9 - 6x - x^2
\end{aligned}$$

The second student writes the following:

$$\begin{aligned}
-(3 - x)^2 &= -[(3 - x)(3 - x)] \\
&= -[(3)(3) + (3)(-x) + (-x)(3) + (-x)(-x)] \\
&= -[9 + (-3x) + (-3x) + (x)^2] \\
&= -[9 - 6x - x^2] \\
&= -9 + 6x + x^2
\end{aligned}$$

Who is right between the two? Find and explain the error(s)

2. A random sample of 1500 teenagers (ages 12-17) were asked whether they watched basketball online; 990 said that they did.
 - a. Construct and interpret a 95 % confidence interval for the population p .
 - b. Suppose that the results of the survey were used to construct separate 95% confidence intervals for boys and girls. Would the margins of error for those two confidence intervals be the same as, larger than, or smaller than that of the interval you constructed in part (a). Justify your answer.

Implications, conjecture, and comparison

1. Compare the elimination method with the substitution method for solving systems of equations.
2. Prove or disapprove that for any fixed confidence level C , and n observations, the margin of error is cut in one third if we take nine times as many observations.

Evaluation

1. Write a short paragraph to discuss the relative merits of solving systems of equations with two unknown variables by elimination method and by graphing method.
2. A level C confidence interval for the mean μ based on a simple random sample n can be calculated by using:

- i. $\bar{x} \pm z^* \frac{\sigma}{n}$, where z^* is the critical value for the standard normal distribution.
- ii. $\bar{x} \pm t^* \frac{s}{n}$, where t^* is the critical value for $t(n-1)$ distribution.

Describe the circumstances under which each one, (i) and (ii) would be appropriate. Are there any similarities or differences between the distributions?