

1-1-2003

## High-stakes testing : truth or consequential validity.

Mary L. Zanetti  
*University of Massachusetts Amherst*

Follow this and additional works at: [https://scholarworks.umass.edu/dissertations\\_1](https://scholarworks.umass.edu/dissertations_1)

---

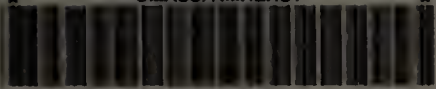
### Recommended Citation

Zanetti, Mary L., "High-stakes testing : truth or consequential validity." (2003). *Doctoral Dissertations 1896 - February 2014*. 5505.

[https://scholarworks.umass.edu/dissertations\\_1/5505](https://scholarworks.umass.edu/dissertations_1/5505)

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

UMASS/AMHERST



312066 0288 2629 0

HIGH-STAKES TESTING: TRUTH OR CONSEQUENTIAL VALIDITY

A Dissertation Presented

by

MARY L. ZANETTI

Submitted to the Graduate School of the  
University of Massachusetts Amherst in partial fulfillment  
of the requirements for the degree of

DOCTOR OF EDUCATION

May 2003

School of Education

© Copyright by Mary L. Zanetti 2003  
All Rights Reserved

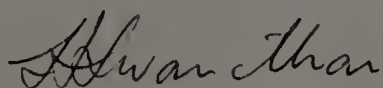
HIGH-STAKES TESTING: TRUTH OR CONSEQUENTIAL VALIDITY

A Dissertation Presented

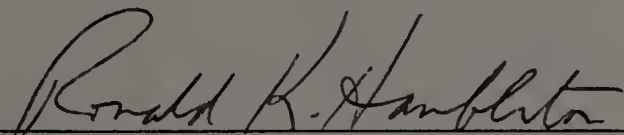
by

MARY L. ZANETTI

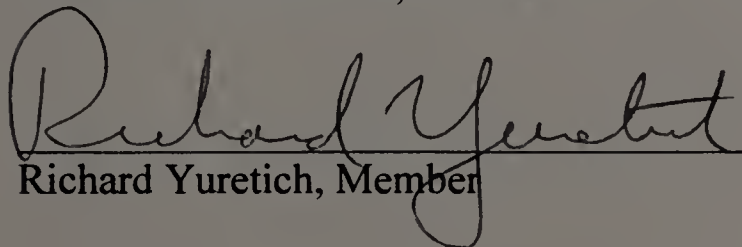
Approved as to style and content by:



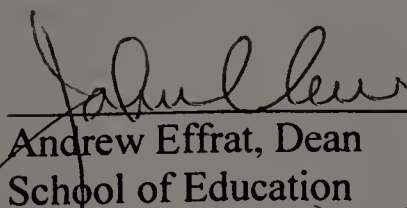
Hariharan Swaminathan, Chair



Ronald K. Hambleton, Member



Richard Yuretich, Member



Andrew Effrat, Dean  
School of Education

## DEDICATION

To my husband, Richard,  
who encouraged and supported me throughout my graduate studies.

To my sons, Anthony M. and Steven J.,  
who inspired me to continue my educational journey.

To my father, Paul V. Slattery, and my sister, Ann L. Slattery,  
who reminded me that I had the determination and ability needed to successfully  
complete this program of study.

## ACKNOWLEDGEMENTS

It is important to acknowledge the professors in the REMP program, who have been my teachers and mentors for the past four years. Most importantly, I would like to thank Stephen G. Sireci, whose guidance and support ensured my success as a doctoral student. Stephen G. Sireci, Ronald K. Hambleton, and Hariharan Swaminathan, individually and collectively, provided me with a rich measurement environment in which to grow and develop as a psychometrician. Each professor adds an important dimension to the REMP program of study and I feel fortunate to have benefited from such capable education measurement experts.

I would also like to extend my gratitude to the members of my dissertation committee, Ronald K. Hambleton, Hariharan Swaminathan, and Richard Yuretich for their participation in this important event. In particular, I would like to thank Swami, the Chair of my committee, whose advice and suggestions proved to be invaluable during the entire research process. In addition, I will be forever grateful to Ronald K. Hambleton for his timely and in-depth critique of my thoughts and written work as I progressed through the dissertation process.

During my REMP educational experience, I have been truly blessed with the blossoming of new friendships and the added bonus of a strong student-to-student learning network. While these events are not unique, the friendships developed during my studies are indeed very special and everlasting. I am truly grateful for the friendship and support of all of my classmates, in particular, Michael Jodoin, Kristen Huff, and Shameem Khaliq, each of whom offered words of encouragement and support at crucial

times, but I truly would not have completed my course of study without the unconditional friendship and acceptance of my friend, Xiaoying Ma, whose strength and resilience continues to amaze me.



## ABSTRACT

### HIGH-STAKES TESTING: TRUTH OR CONSEQUENTIAL VALIDITY

MAY 2003

MARY L. ZANETTI, B.S., NORTH ADAMS STATE COLLEGE

M.Ed., UNIVERSITY OF MASSACHUSETTS AMHERST

Ed.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Hariharan Swaminathan

This study consisted of a comprehensive review of the consequential aspects of validity of a grade 10 mathematics assessment. This test is part of a larger state-mandated assessment system in which the studied test is one of two assessments that a student must pass in order to graduate from high school in the state of Massachusetts. Validity evidence was collected using three rigorous measurement methods. Qualitative and quantitative procedures were used to ensure a more complete collection and analyses of validity evidence. A survey was developed and administered to all participating high school mathematics teachers and key education personnel. Fifty-six percent of the surveys were completed and analyzed. In addition, focus group and one-on-one interviews were conducted within each participating school district. The results indicated that the Massachusetts' education reform initiative had created significant changes in high school mathematics curriculum and instruction. In addition, many positive and negative intended and unintended consequences connected to this assessment system were identified. The results were discussed based on a classification system in which a representative sample of school districts was selected from the state population.

In this study, a comprehensive analysis of a few specific consequential validity questions was addressed using sound quantitative and qualitative research methods. This type of research, examining the consequential aspects of validity of a state-mandated test as a component of a larger assessment system, represents a huge undertaking. The social, politic, and educational implications involved in any reform effort are complex and difficult to document. As education reform affects more and more students across this nation, answers to the outlined questions may assist key administrators in the state of Massachusetts, perhaps even other states in the middle of similar reform efforts, in making important mid-course corrections, and/or merely provide needed validity evidence regarding intended and unintended consequences of the program using solid, data-driven information.

## TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGMENTS .....	v
ABSTRACT .....	vii
LIST OF TABLES .....	xiii
LIST OF FIGURES .....	xiv
CHAPTER	
1. INTRODUCTION .....	1
1.1 Background .....	1
1.2 Statement of the Problem .....	3
1.3 Purpose of the Study .....	5
1.4 Significance of the Problem .....	6
1.5 Outline of the Study .....	8
2. REVIEW OF THE LITERATURE .....	9
2.1 Introduction .....	9
2.2 Consequential Aspects of Validity .....	10
2.2.1 1985 <i>Standards</i> Offer Definition of “Validity” .....	11
2.2.2 Proponents of the Consequential Aspects of Validity .....	11
2.2.3 Opponents of the Consequential Aspects of Validity .....	15
2.3 Types of Validity Evidence that should be Gathered .....	17
2.4 Empirical Studies on the Consequential Aspects of Validity .....	22
2.5 Mixed Method Approaches .....	25
2.6 Summary .....	28
3. METHODOLOGY .....	31
3.1 Introduction .....	31
3.2 Sampling Procedure .....	32
3.3 Survey Design .....	35
3.4 Focus Group Protocol .....	37
3.5 Interview Protocol .....	38

3.6	Data Analyses . . . . .	40
3.6.1	Survey Data Analyses . . . . .	40
3.6.2	Focus Group Data Analyses. . . . .	40
3.6.3	Interview Data Analyses . . . . .	41
4.	RESULTS . . . . .	43
4.1	Survey Results . . . . .	43
4.1.1	Participant Response Rate . . . . .	43
4.1.2	Background Information . . . . .	44
4.1.3	Opinions Related to Aspects of MCAS Assessment System . . . . .	47
4.1.4	Curriculum and Instructional Practices with Respect to Time . . . . .	50
4.1.5	Content Analyses of Open-Ended Questions . . . . .	54
4.2	Focus Group Results . . . . .	57
4.2.1	Changes in Curriculum . . . . .	58
4.2.1.1	Urban Schools . . . . .	58
4.2.1.2	Suburban Schools . . . . .	62
4.2.2	Changes in Instruction . . . . .	64
4.2.2.1	Urban Schools . . . . .	64
4.2.2.2	Suburban Schools . . . . .	66
4.2.3	Educators' Opinions of a Standards-Based Curriculum . . . . .	68
4.2.3.1	Urban Schools . . . . .	68
4.2.3.2	Suburban Schools . . . . .	69
4.2.4	Morale and Motivation of Educators . . . . .	70
4.2.4.1	Urban Schools . . . . .	70
4.2.4.2	Suburban Schools . . . . .	71

4.2.5	Improvement in Student Learning .....	71
4.2.5.1	Urban Schools .....	72
4.2.5.2	Suburban Schools. ....	75
4.2.6	Suggested Changes to Assessment System .....	76
4.2.6.1	Urban Schools .....	76
4.2.6.2	Suburban Schools .....	77
4.3	Interview Results .....	78
4.3.1	Changes in Curriculum and Instruction .....	79
4.3.1.1	Mathematics Department Chairs/Heads .....	79
4.3.1.2	Principals and Assistant Principals .....	81
4.3.1.3	Superintendents and Assistant Superintendents	82
4.3.2	Opinions on Standards-Based Curriculum .....	83
4.3.2.1	Mathematics Department Chairs/Heads .....	83
4.3.2.2	Principals and Assistant Principals .....	84
4.3.2.3	Superintendents and Assistant Superintendents	85
4.3.3	Effect of MCAS on Students, Teachers, and Non-Teachers .	85
4.3.3.1	Mathematics Department Chairs/Heads .....	85
4.3.3.2	Principals and Assistant Principals .....	88
4.3.3.3	Superintendents and Assistant Superintendents	88
5.	CONCLUSION .....	90
5.1	Summary of Results .....	90
5.1.1	Summary of Survey Results .....	90
5.1.2	Summary of Focus Group Results .....	91
5.1.3	Summary of Interview Results .....	92
5.1.4	Study Summary .....	94
5.2	Delimitations .....	96
5.3	Discussion of Findings and Implications .....	97
5.4	Future Research .....	98
5.5	Conclusion .....	99

APPENDICES

A. WRITTEN CONSENT FORM ..... 101  
B. COVER LETTER ..... 103  
C. LETTER OF SUPPORT ..... 106  
D. TEACHER'S SURVEY ..... 108  
E. NON-TEACHER'S SURVEY ..... 113  
F. FOCUS GROUP PROTOCOL ..... 118  
G. INTERVIEW PROTOCOL ..... 120  
H. TELEPHONE CONTACT SUMMARY SHEET ..... 122

BIBLIOGRAPHY ..... 124

## LIST OF TABLES

Table		Page
3.1	Twelve Group Cluster with Variable Averages .....	34
4.1	Survey Participation Rates .....	44
4.2	Student Enrollment in Participating Schools .....	45
4.3	Categories of Courses Taught by Teachers .....	46
4.4	Years in Teaching Profession .....	47
4.5	Descriptive and Inferential Statistics for Influence Statements .....	53

## LIST OF FIGURES

Figure	Page
2.1 Messick's Facets of Validity Framework . . . . .	12
3.1 A Concept Map of the Research Study . . . . .	33



## CHAPTER 1

### INTRODUCTION

#### 1.1 Background

It is important that all educators and education researchers understand the role democracy plays in public education. A clear understanding of the connection between democracy and our public schools will help practitioners and researchers see a common purpose or meaning in their work. Ensuring a civil society gives significant meaning to public education.

In 1786, Thomas Jefferson wrote a letter to his old professor George Wythe in support of a bill for general education

I think by far the most important bill in our whole code is that for the diffusion of knowledge, among the people. No other sure foundation can be devised for the preservation of freedom and happiness...Preach, my dear sir, a crusade against ignorance; establish and improve the law for educating the common people. Let our countrymen know...that the tax which will be paid for this purpose is not more than the thousandth part of what will be paid to kings, priests, and nobles who will rise up among us if we leave the people in ignorance. (Alexander & Alexander, 1992, p.20)

This country's founding fathers went to great lengths to guarantee a free public education to all its citizens. In 1835, Thaddeus Stevens believed that education was a public obligation that must be nurtured to develop the entire civic intelligence to better govern through an elective republic (Alexander & Alexander, 1992). It is clear democracy has always played an important role in this country's public education system. Many will argue to what degree, but it is safe to say that it has always been of great significance to the people of the United States of America. Education is a great

equalizer among people. It can balance the social conditions among and within the people living in a society.

A democratic society relies on its citizenry to elect responsible people to maintain its government; therefore, the education of all people living in a democratic society is paramount to its success as a nation. As a country, the United States of America remains faithful to its forefathers' belief that all citizens have a right to be educated in a fair and equitable manner. The definition of a fair and equitable education is the subject of much debate in state legislatures across the United States of America.

Education will remain an important political issue in the United States because the education of its people is a guaranteed inalienable right. With that in mind, it is not surprising to see education reform efforts taking place at some level in most states across this nation. In fact, all six of the New England states currently have some form of education reform effort in effect. Many states include a mandated testing requirement in its education reform effort, while other states across the nation connect a graduation requirement to its state-mandated assessment. In other words, a test(s) must be passed in order to graduate from high school.

The trends outlined above highlight the potentially important role evaluation plays in education reform. Measurement experts must analyze and evaluate these state-mandated tests to ensure fair and equitable treatment of all students affected by an education reform initiative.

## 1.2 Statement of the Problem

It is important to note that the Massachusetts Education Reform Act (MERA) enacted in June 1993 put a three-step reform process in motion. These three steps were the creation of the (1) Common Core of Learning, (2) curriculum guides in seven academic areas, and (3) a comprehensive assessment system. In addition, section 72 of the MERA required all school districts to submit to the Board of Education a plan to eliminate the general track and to create program options that assisted all students to meet the high standards of the outlined education reform initiative. The general track was defined as any course of study where students were not being specifically prepared for postsecondary higher education, technical training, or meaningful employment with career advancement potential, and for full participation as citizens in a democratic society (Massachusetts Department of Education, 1996).

The first step set forth the broad goals for education identifying what students should know and be able to do. These broad goals emphasized that teaching and learning must be interdisciplinary. In addition, it was outlined that individuals must think and communicate, gain and apply knowledge, and work and contribute to society. The second step, the creation of curriculum guides, developed into K-12 Curriculum Frameworks in seven academic areas. The third step was the implementation of a comprehensive assessment system. This system was initially implemented in 1998 and the current version of this assessment system is described in the next paragraph.

The Spring 2001 testing schedule in the studied state included students in grades three through ten with students in each grade level taking one or more of the four core-subject area (English Language Arts (ELA), Math, Science, and Social Studies)

examinations. Beginning with the high school class of 2003, students must pass the tenth grade English Language Arts and mathematics exams in order to receive a high school diploma. Therefore, this study has attempted to determine the effect the grade 10 mathematics component of this pending graduation requirement has had on teachers, administrators, and school districts.

In September 1999, the Massachusetts Board of Education adopted the School and District Accountability System and part of that system was the implementation of the School Performance Rating Process (SPRP), which rated school performance and improvement. This rating system consisted of a multi-year cycle whereby schools were assigned overall performance ratings based on two years of assessment results and an overall improvement rating based on the comparison of those results to a baseline. Performance was measured by the percentage of students demonstrating proficient and advanced performance on Massachusetts Comprehensive Assessment System's (MCAS) ELA, mathematics, and science & technology tests and the decreasing percentage of students failing those same tests (Massachusetts Department of Education, 2001). The SPRP measured the improvement using the average scaled score of the ELA, mathematics, and science and technology assessments.

Because of the important role the MCAS mathematics assessment played in this process the appropriateness of examining and documenting changes in high school mathematics as a result of the assessment system was further bolstered. It was also appropriate to determine if school districts were connecting the assessment system's grade 10 mathematics results to high school mathematics curriculum and classroom instruction. The intended and unintended consequences of using the grade 10

mathematics test results as a means of evaluating a district's mathematics curriculum, student achievement, and educational accountability was analyzed as a result of this test's important educational role.

Overall, this study attempted to identify the intended and unintended consequences of the MCAS. More specifically, this research study attempted to gather evidence pertaining to the consequential aspects of validity connected to the MCAS grade 10 mathematics assessment.

### 1.3 Purpose of the Study

It is important to note that state testing programs are usually carefully implemented and evaluated including item reviews, item bias reviews, equating, setting standards, etc. While the MCAS has been carefully implemented and continues to be evaluated, it is unfortunate little to none consequential validity studies have been conducted in the studied state. This study was developed with that in mind. The three procedures used in this study were specifically designed to answer the following consequential validity questions: (a) Have changes in curriculum occurred within high school math departments across the commonwealth as a result of the statewide grade 10 mathematics assessment and its corresponding curriculum framework? (b) Has high school mathematics instruction changed due to the state's mathematics curriculum standards? (c) Has district high school mathematics curriculum been aligned with the state's mathematics curriculum framework?

All of these questions investigated or examined various aspects of the consequential validity of the assessment system. Messick (1989) suggested the validity

of a system as a whole needed to be evaluated in terms of its effects in improving teaching and learning. It is for this reason that the consequences of the MCAS included evaluation of the impact the state's frameworks had on high school mathematics curriculum and instructional practices. In this study, careful analyses of the goals of the comprehensive assessment system and the outlined consequential validity questions have been addressed using sound quantitative and qualitative research methods. This type of research, examining the consequential aspects of the validity of all or part of an assessment system, represents a huge under-taking. The social, political, and educational implications involved in any reform effort are complex and difficult to document. As education reform affects more and more students across this nation, answers to the previously outlined questions may assist key administrators in Massachusetts, perhaps even other states in the middle of similar reform efforts, in making important mid-course corrections, and/or merely provide needed validity evidence regarding intended and unintended consequences of the program using solid, data-driven information.

#### 1.4 Significance of the Problem

Since legislators in the state of Massachusetts passed this highly ambitious education reform act in 1993, billions of tax dollars have been spent on education reform in the 245 local and/or regional school districts across the Commonwealth. Inaccurate and negative articles relating to this important topic, in particular the state-mandated grade 10 test, which is a component of a larger comprehensive assessment system, can be found in virtually every newspaper across the state. For example, a

recent article entitled “MCAS Foes Undeterred by Test Trend” ran in The Boston Globe on December 25, 2001. Globe Staff Writer, Scott S. Greenberger interviewed Monty Neill, head of Fair Test, an anti-testing organization located in Cambridge, Massachusetts. Neill mentioned the recent gains in the studied assessment may slow opposition to the exam, but his organization will continue the battle against the state-mandated testing system. In addition, The Boston Globe recently ran an article entitled “Ranking of Schools Draws Official Ire, Activist Says Ratings Help Inform Parents.” The article outlined an education activist’s annual rankings of the state’s public schools’ average scores. Department of Education officials argued that her system was flawed and further explained “scores should be used to measure each school against itself, not against one another” (Vaishnav, 2002, p. B6). These two articles reflect the level of disagreement and misinformation that is often connected to the media’s coverage of any high-stakes testing program.

While it may have been premature to proceed with an evaluation earlier, it is appropriate and very necessary to begin an evaluation of the strengths and weaknesses of all aspects of this huge reform initiative, which is currently in its ninth year. This is especially true today because the high school graduation requirement goes into effect with the class of 2003. It is now apparent that taxpayers, legislators, students, educational researchers, and policymakers are in need of concrete, accurate information regarding the effectiveness of this state’s education reform initiative. This study focused on one component of this assessment system. That is, the consequences of this reform effort as it relates to the state-mandated grade 10 mathematics assessment and correlating curriculum standards.

More specifically, this study examined the effect the comprehensive assessment system had on district curriculum standards connected with high school mathematics, and indirectly, instructional practices. It should be kept in mind that the assessment system was created by the Commonwealth's Department of Education (DOE) to fulfill an important tenet outlined in the state's education reform initiative enacted in 1993. The newspaper articles previously cited reflect the huge debate that continues across this state and often times at the national level. So many individuals think they know what is going on and offer sharp criticism regarding this assessment system, but little hard evidence is available. The majority of criticism directed at the assessment program is based on speculation rather than empirically supported facts. Educational researchers need to step forward and offer a means of methodically analyzing various aspects of this education reform effort.

### 1.5 Outline of the Study

This section describes the organization of the remainder of this dissertation. Chapter 2 reviews literature relating to the following topics: consequential aspects of validity, types of validity evidence that should be gathered, empirical studies on consequential aspects of validity, and mixed method approaches. Chapter 3 describes the following topics that relate to methodology: mixed method research design, survey design, focus group and interview protocols, sampling procedure, and data analysis. Chapter 4 presents the results of the survey, focus group meetings, and interviews. Chapter 5 presents the conclusions for this research, limitations of the study, and discusses the implications for future research.



## CHAPTER 2

### REVIEW OF THE LITERATURE

#### 2.1 Introduction

This chapter includes a comprehensive review of the consequential aspects of validity literature. It has six sections including the introduction and summary. The second section outlines the current and past debate about the merit of the colloquial term “consequential validity” and includes researchers’ viewpoints of the pros and cons of considering the consequential aspects of validity as a separate type of validity. The third section discusses the validity evidence that should be gathered in a study examining the consequences of a test or testing program. The fourth section contains a review of empirical studies on this subject. Evaluating the intended and unintended consequences of tests, especially high-stakes assessments, are important and very necessary because of the impact these tests have on people including students, teachers, policymakers, and employers. The fourth section also mentions specific quantitative and qualitative methods that were found to be useful in studying the social consequences of tests. Finally, the fifth section provides an overview of quantitative and qualitative paradigms, also known as mixed method approaches, which were used in the majority of the cited empirical studies and subsequently adapted for use in this research study.

The debate surrounding the term “consequential validity” continues in many psychometricians’ hearts and minds. Highly regarded measurement experts, such as Samuel Messick, right down to fledgling doctoral students enrolled in various

Education Measurement Programs throughout the world have a wide range of opinions regarding the proper definition of this term. Is it a kind of validity that can stand on its own merit? Is it a type of validity evidence? Is it an element of construct validity? Is it a kind of evidence that can sometimes inform policy decisions but fall outside the technical purview of validity (American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME), 1999)? Based on a thorough review of the literature on this topic, one can respond “yes” and “no” to all of these questions.

## 2.2 Consequential Aspects of Validity

While many measurement experts casually use the term “consequential validity”, it appears to be more of a colloquial term rather than a clearly defined measurement term. The literature on the consequential aspects of validity including the debate about how this evidence should be handled is intermingled with these terms: the colloquial term “consequential validity” and the measurement phrase “consequential aspects of validity.” Nevertheless, the consequential aspects of validity, also known as “consequential validity”, are important components of test interpretations meriting further review. For example, studies investigating the consequences of testing raise research or testing experts’ awareness of the social consequences connected with testing results. Cronbach and Meehl (1955) coined the term “nomological net,” which meant construct validation involved evaluating the construct validity of so-called criterion measures in the hypothesized network as well as the target measure. Shepard (1997) explained that Cronbach and Meehl believed test developers were responsible for

comparing the theorized relationships between test results and outcomes and were accountable for the validity of both the test and the explanatory theory.

### 2.2.1 1985 Standards Offer Definition of “Validity”

The American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) outlined in the 1985 Standards for Educational and Psychological Testing “validity is the most important consideration in test evaluation” (p. 9). Validity is also clearly defined as a unitary concept. These same standards indicated there are many ways to gather validity evidence to support the inferences made from test scores. In other words, various types of evidence can be gathered to support or oppose the validity of the score-based inferences. The next two sections focus on measurement experts’ positive and negative views of validity evidence, which is deemed to evaluate the consequences of testing as well as whether this evidence should be considered in a validity study at all.

### 2.2.2 Proponents of the Consequential Aspects of Validity

Messick was clearly a proponent of the 1985 Standards; however, in 1989, he further outlined his conception of validity in a chapter in Educational Measurement (see Linn, 1989), stating that the consequential basis of test interpretation comprises the value implications of constructs and their associated measures. In this 1989 chapter on validity, Messick used a two-by-two matrix (see Figure 2.1) to explain the different facets of validity and to show that construct validity was contained in all four of the

	Test interpretation	Test use
Evidential Basis	Construct validity	Construct validity + Relevance/utility
Consequential Basis	Value implications	Social consequences

Figure 2.1. Messick's Facets of Validity Framework

cells. Messick indicated in his writing about this matrix that value implications and social consequences make up the consequential basis of evidence pertaining to validity. Then in 1995, Messick further stated "in particular, what needs to be valid is the meaning or interpretation of the scores as well as any implications for action that this meaning entails" (p. 743). Messick's writings have been accepted by many in the measurement field as the theoretical foundation of validity. Messick's statements in 1989 and 1995 clearly supported the idea that social consequences, referred to by some researchers as "consequential validity," are an important aspect of validity.

Messick (1989) wanted measurement professionals to consider the value implications connected to test names, construct labels, theories, and ideologies. It is important to consider values in score interpretations because they can directly and/or indirectly affect individuals, schools, and society. For example, state mandated high-stakes test results have obvious consequences for a student, school, school district, and society as a whole. Whenever a test score is interpreted, it is judged and connected to some broader category. If a particular minority group scores low on a specific

assessment, how will their individual scores and that specific group be viewed and judged by others?

Messick (1989) provided a great example of how a specific situation can be viewed differently. He offered the following situation: if overrepresentation of minority children in special education programs is perceived as a violation of equal protection under the law, strategies for reducing the inequality might be sought. But if seen as demographic differences in educational needs, then strategies for equitable provision of educational services would be stressed. This is an extremely useful and clear example of how the interpretation of a situation can carry out at least three major sources of value connotations: the evaluative overtones of the construct labels themselves; the value connotations of the broader theories or nomological networks in which the constructs are embedded; and the value implications of still broader ideologies about the nature of mankind, society, and the science that color our manner of perceiving and proceeding (Messick, 1980).

It is very difficult to separate value implications of test interpretations from the validity of the interpretations themselves. Messick (1989) stressed that the validation process must include the value implications of the score interpretations. The social consequences of testing are value laden. The value aspects of score meaning must be taken into account. Since test scores continue to affect the setting of social policy, it is important that the social consequences of testing continue to remain a part of the validation process. Measurement professionals know policymakers use test score interpretations during the process of setting policies; therefore, these professionals must continue to ensure the validity of their meaning and their value implications.

Shepard (1993) agreed with Messick's unified theory of validity, but she took issue with his use of a four-celled table (see Figure 2.1) describing the validity framework. Her concern centered on Messick's use of construct validity in more than one cell. It is widely understood that he believed construct validity may be construed as the whole of validity, but his table does not clearly show that. Shepard (1993) believed values are separate and distinct from test interpretations. She was also concerned about Messick's omission of practical validity questions that could assist researchers in evaluating or supporting test use.

While Messick and Shepard concur on the unified concept of validity, they differ in their views of how important validity information should be relayed to those in an applied measurement setting. Providing measurement professionals with practical validity questions would absolutely enhance their ability to provide important validity evidence in the area of test use and test interpretations. Shepard (1997) indicated that depending on test use, intended consequences of testing may or may not be a part of the relationships represented in Cronbach and Meehl's nomological net. For instance, if a teacher prepared a test as a culmination of a curriculum unit, the consequences of that test may not be part of the nomological net. She also referred to "side effects" of a test as the unintended consequences of a test used for its intended purpose. Shepard wisely cautioned psychometricians against confusing the question of who is responsible for the consequences of improper test use with the question of what to include when evaluating validity for the test's intended use.

### 2.2.3 Opponents of the Consequential Aspects of Validity

Maguire, Hattie, and Haig (1994) took issue with aspects of Messick's 1989 chapter in Linn's Educational Measurement. One specific concern related to the consequential aspects of validity. They discussed whether or not Messick's concern with "consequential validity" was in direct response to the increase in litigation surrounding tests and testing practices. They believed Messick gave "consequential validity" a more important role within construct validity because of the increasing amount of legal challenges within the testing world. They further stated consequences of test use are important, but "consequences should be moved out from the umbrella of construct validity and into the arena of informed social debate and formulated into ethical guidelines" (Maguire, Hattie, & Haig, 1994, p. 115). They added

...implicit value assumptions and social consequences of testing are better examined through processes such as those raised by Kane (1992) who presented an argument-based approach to validity that is positioned to deal with how tests should be used and interpreted. (Maguire, Hattie, & Haig, 1994, p. 114)

Does Kane's approach move the consequences of testing out of the validity arena? Kane's (1992) approach to gathering validity evidence further bolsters the need to investigate the consequences of test usage. Using this argument-based approach, intended and unintended consequences of inferences from test results must be investigated in light of the intended construct meaning. Thus, there appears to be a conflict in Maguire, Mattie, and Haig's (1994) argument.

Many highly regarded measurement experts continue to weigh in on whether or not "consequential validity" is a legitimate term. Popham (1997) stated he agreed with advocates of consequential validity who believe the social consequences of test use

should be considered when judging whether a test's use is appropriate. He further indicated that issue should be considered separately from Messick's (1989) validity framework whereas proponents of consequential validity would intertwine the two issues. Tenopyr (1996) stated

Conceptions of construct validity have changed somewhat over the years, but vagaries of measurement aside, most psychologists have agreed that constructs basically pertain to living things. To expand construct validity to cover actions by the test user or others appears to be a misinterpretation of the common consensus of measurement experts. To carry this further and speak of "consequential validity" is a perversion of the scientific underpinnings of measurement. (p. 14)

Mehrens (1997) joined the debate by stating the examination of consequences may tell us more about the adequacy of the treatment or the general wisdom or social acceptability of the action than the validity of the inference about the construct. In other words, an examination of the consequences of the grade 10 mathematics test involved in this study may only tell us about the social acceptability of the DOE's actions rather than the validity of the inference about the mathematics construct being measured. Mehrens' (1997) statement reflects a growing sentiment among many measurement experts concerning the social consequences of an assessment. This group believes social issues emanating from test inferences should be examined separate from any validity investigation. That is, a validity study should investigate inferences about the construct being evaluated at the exclusion of social consequences. The next section outlines the types of validity evidence that should be gathered when a researcher examines the consequential aspects of validity of a test or assessment program.



### 2.3 Types of Validity Evidence that should be Gathered

In the 1999 Standards for Educational and Psychological Testing, validity is referred to as “the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests” (AERA, APA, & NCME, 1999, p.9). The Standards (1999) clearly explained that the process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations. The Standards (1999) further stated, “as the validation proceeds, and new evidence about the meaning of a test’s scores becomes available, revisions may be needed in the test, in the conceptual framework that shapes it, and even in the construct underlying the test” (p. 9). If subsequent validity evidence suggests construct-related problems, such as construct under-representation or construct-irrelevant variance, then measurement experts may need to review the validity of the test program. The 1999 Standards discussed sources of validity evidence rather than distinct types of validity. Validity is again defined as a unitary concept.

It is important to note that the 1999 Standards included a section entitled “Evidence Based on Consequences of Testing.” In this section, the AERA, the APA, and the NCME “distinguished between evidence that is directly relevant to validity and evidence that may inform decisions about social policy but falls outside the realm of validity” (p. 16). Researchers were advised to consider how these consequences affect the validity or invalidity of the intended test interpretations. Even if information regarding the consequences of testing influences policies related to test use, these issues do not automatically render the test interpretations invalid. Given these reasonable

guidelines, it seems appropriate that a validity investigation of the sources of those consequences must be pursued in order to determine a test's validity or invalidity.

Employment tests offer a clear example, if a test measures a construct that mirrors the skills required to perform a specific job, then it is appropriate to require potential employees to pass the test even if certain groups perform differently on the assessment; however, if an employment test measures skills that are unrelated to the job, then the validity of the test results should be called into question. The latter is an example of construct-irrelevance, which is a source of invalidity. If the employment test failed to include important aspects of the skills required to perform a specific job, then this would represent construct under-representation, another source of invalidity. Evidence that produces important yet valid differences in performance between groups should be collected and used when making policy decisions. This distinction was clearly stated in the 1999 Standards to assist measurement experts in sifting through the increasing demands regarding the consequential aspects of validity.

Tests are used to evaluate examinees' level of proficiency, make a selection among examinees, or to determine the interests of examinees. The validation process should be used to evaluate whether or not the specified benefit is realized. Consider the grade 10 mathematics assessment mandated by the studied state. One of the purposes of this assessment is to determine a student's proficiency in mathematics based on the curriculum standards outlined by that state. School districts across the commonwealth have been asked to use the state's mathematics standards as a guide in aligning its K-12 curriculum. Therefore, an appropriate validity question would be: does high school mathematics curriculum within each district meet the outlined standards and are they

appropriate standards? A consequence of this assessment program would be changes in high school classroom instructional practices that occurred due to high school teachers' increased awareness of the state's mathematics curriculum standards. The validation process could include evidence gathered from an investigation of a representative sample of high school mathematics teachers in the studied state to determine if the assessment system, specifically the grade 10 mathematics test and its corresponding standards, assisted in deepening teachers' subject/content knowledge and in turn improved student learning.

Mehrens (1998) discussed how to evaluate whether the consequences of an assessment are good or bad. He also explained that he is not a proponent of the term "consequential validity"; however, he is interested in the consequences of assessment. After reading Mehrens (1998), the following is a sample of validity questions related to the studied assessment program: What are the consequential aspects of validity in a state-mandated standards-based assessment system? Does this type of assessment bring about improved classroom activities? Does involvement in this assessment program deepen an educator's subject knowledge? How does participation in a state-mandated assessment program affect a teacher's motivation and morale? What are the intended and unintended consequences of these assessments? Mehrens (1998) added for testing to be a good thing, the positive consequences must outweigh the negative consequences.

Baker, O'Neil, and Linn (1993) stated, "an explicit relationship between validity standards and policy uses should be conceived for performance assessments" (p. 1215). It appears that their summary of validity issues can be easily adapted for any type of assessment. They explained

In accordance with concerns for validity, including equitable inferences for members of different groups, the higher the individual stakes, the more evidence should be required to substantiate the claim of valid measurement. Validity at a minimum should be established for the policy decision and purpose for which the assessment is intended. Standards used to classify successful and unsuccessful candidates must also be validated. Furthermore, if assessments provide a wide range of topic options or assignments to examined students, validity evidence should be assembled to document the comparability of the assessments in the interest of fairness. Comparability may be established through empirical or a combination of empirical and analytical approaches. (Baker, O'Neil, & Linn, 1993, p. 1215)

It is clear that gathering validity evidence, in particular the social aspects of validity, is an important component of the work measurement professionals do, but it is easy to understand why there are few validity studies given the abstract nature of validity and the expense and time needed to gather such evidence.

Lane, Parke, and Stone (1998) outlined a comprehensive framework that can be used to investigate the consequences of an entire assessment program. These researchers suggested such a study should collect evidence of intended and unintended consequences at the following levels: state, district, school building, and classroom. A wide variety of measurement tools should be used to gather validity evidence in this type of study, such as: surveys, interviews, and focus group work. In addition, curriculum and test preparation materials should be collected and reviewed. Curriculum specialists and other key educators should be asked to discuss the impact the assessment program has on students, curriculum, and instructional practices. The study should focus on the validity of the assessment system as a whole as suggested by Messick (1995). Consequential evidence should be purposely obtained from many levels within the studied assessment program. This triangulation of evidence offers a more comprehensive view of the consequential aspects of the studied testing program.

Baker, Linn, Herman, and Koretz (2002) offered in a CRESST Policy Brief twenty-two standards representing models of practice derived from three perspectives: research knowledge, practical experience, and ethical considerations. These researchers suggested that the outlined standards should be applied to accountability systems, while tests included in an accountability system should continue to meet the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 1999). They also stated “it is likely that additional standards will be subsequently developed based on evaluations of accountability system effects” (Baker, Linn, Herman, & Koretz, 2002, p.2). These researchers’ work represents an emerging body of knowledge investigating the accountability of an education system. Due to on-going education reform initiatives throughout the United States, it is very prudent to pay close attention to this type of research, which will ultimately affect the quality of consequential validity studies in the future.

This section ends with an overview of an article written by Moss (1998). She wrote about the consequential aspects of validity within a social science framework. She argued for a larger, long-term research agenda concerning the consequences of recurrent or regularized testing on schools and society as a whole. In addition, she pointed out the importance of considering these issues because the use of these types of assessment combined with other social factors have rapidly altered our conceptions of individual identity and enhanced our ability to monitor and control people’s actions. She further argued

To the extent that these testing practices in which we engage change the social reality we study, the study of the consequences becomes an essential aspect of validity even for those who choose to limit the scope of validity to a test-based interpretation. (Moss, 1998, p.11)

## 2.4 Empirical Studies on the Consequential Aspects of Validity

This section outlines empirical studies that have investigated the consequential aspects of validity. It is more practical to appraise the consequential aspects of validity after tests are in use for a period of time. Using surveys, focus groups, and interviews, key education personnel can be asked about the intended and unintended consequences of a specific test or testing program. It is important to remember to include students at some point in the validation process, also.

Since empirical studies on this topic are sparse, the first review involves an untraditional research report. Taleporos (1998) outlined “consequential validity” from a practitioner’s perspective. As a member of the New York City Assessment staff, she compared running a testing program in a large urban city to trying to survive in the epicenter of an earthquake. She likened the consequences of testing to the earthquake’s aftershocks, happening at unpredictable times, and lingering on long after the major event. Nitko (1993) was cited by Taleporos (1998) when she stated, “tests are supposed to measure what is thought of as important, but not define it. One should build curricula and develop or select tests to determine if the desired learning outcomes have occurred” (p.20). While improved instruction is a goal, steps need to be taken to maximize both intended and unintended positive consequences and minimize the negative unintended consequences. Taleporos (1998) believed test consequences needed to be managed. The huge New York City school district found using multiple measures that were broad enough to serve more than one single interest to be a necessary component of their test management plan. Teachers, administrators, parents, and district superintendents were asked to describe what they needed and wanted in a testing program. The assessment

staff in conjunction with test publishers dealt with the unintended and intended consequences of their testing program as each issue presented itself. Taleporos (1998) emphasized that the information in her practitioner's report is just as important and meaningful as research data, but collaborative efforts between practitioners and researchers need to be endorsed, also.

Khatti, Kane, and Reeve (1995) summarized what they learned from Studies of Education Reform: Assessment of Student Performance, a 3-year national study about the impact of performance assessments on teaching and learning. They visited 16 schools across the United States. Each school was developing and implementing performance assessments due to a national, state, district, or school education reform initiative. The researchers conducted interviews, collected student work, and observed classrooms and professional development workshops. In general, they found "the effect of assessments on the *curriculum* teachers use in their classrooms has been marginal, although the impact on *instruction* and on *teacher roles* in some cases has been substantial" (Khatti, Kane, & Reeve, 1995, p. 80). They further explained "even when teachers adopt the format of performance assessments, the content and sequencing of subject matter remain largely unchanged" (Khatti, Kane, & Reeve, 1995, p. 80). This explained how curriculum could remain largely unchanged even though teachers' roles and instructional practices were changed in a more significant manner. This specific issue is connected to a major finding in this study related to the ease of adopting and implementing performance assessments. These researchers stated designers and proponents assumed teachers:

- (1) possess a clear understanding of the domains of knowledge to be assessed by the new assessments;
- (2) are knowledgeable about the most

effective approaches of teaching to these assessments; (3) have expertise in a variety of teaching modalities; and (4) know what constitutes evidence of critical thinking skills and multidisciplinary understanding. (p. 83)

These assumptions are directly related to the most important finding in this study and that is “there was no evidence that assessment and instructional changes are driven by a clear understanding of the issues [assumptions #1-4 outlined above]” (Khatti, Kane, & Reeve, 1995, p. 83).

Chudowsky and Behuniak (1998) used teacher focus groups to obtain information about teachers’ perceptions of the impact of the Connecticut Academic Performance Test. The sample included a cross-section of Connecticut schools. Each of the seven participating high schools was asked to voluntarily convene 10-12 high school teachers for the focus group meetings. The interviewers investigated the assessment program’s impact on:

- \*curriculum and instruction
- \*teachers’ expectations of students
- \*students’ behavior and attitudes
- \*parents’ behavior and attitudes
- \*professional development

The most significant finding in this small study involved teachers’ perceptions of the time spent on testing. More specifically, teachers reported the most negative impact of the test was its significant detraction from instructional time for their students. These researchers indicated that the results of this focus group study would be included in a larger, more comprehensive data collection effort in the future. The larger study



will include surveys completed by a random sample of Connecticut teachers. The researchers stated they planned on combining the focus group results with other data collection methods because this action would provide a more comprehensive collection of evidence of the consequential aspects of this testing program.

Linn, Baker, and Betebenner (2002) provided a summary of implications of requirements of the No Child Left Behind Act of 2001 as it relates to state accountability systems. They specifically discussed adequate yearly progress (AYP), which is not addressed in this study, by connecting trends in state assessments and the National Assessment of Educational Progress (NAEP) in grade 8 mathematics performance in five states (California, Maryland, Massachusetts, Oregon, and Texas). These researchers urged states

...to invest in continuing studies (as some of them have) of the impact of their accountability model and the details of its implementation in order to increase the chances of yielding the desired outcome of higher quality education and significantly improved preparation of students. (Linn, Baker, & Betenenner, 2002, p.15)

It is important to note the use of the term “impact” is closely associated with consequential aspects of validity in that these researchers are suggesting state education professionals conduct systemic validity studies, which “will provide states and districts with feedback about the utility of their [educational] systems” (Linn, Baker, & Betenenner, 2002, p. 15).

## 2.5 Mixed Method Approaches

Combining qualitative and quantitative approaches within a research study has long been accepted in the academic and education research world. Currently, more and

more researchers are viewing a mixed method study as a viable alternative for specific types of research. Denzin (1978) used the term triangulation to give merit to a study using combined methods. His argument has further bolstered by Jick (1979) when he stated that the strengths and weaknesses inherent in some quantitative methods would be offset by the strengths and weakness contained within some qualitative methods. Mathison (1988) outlined a triangulation strategy that provided better evidence for researchers by using mixed methods, which can be used to look for convergence, inconsistency, and contradictory evidence. Finally, Morse (1991) outlined two types of methodological triangulation: simultaneous and sequential. Simultaneous triangulation uses both methods at the same time, while sequential triangulation uses the results of one method to plan the next method.

Green, Caracelli, and Graham (1989) reviewed approximately 60 mixed method studies from 1980 to 1988 developing from their work five mixed method purposes for a single study:

- (1) triangulation (convergence of results)
- (2) complimentary (overlapping and different aspects of results may emerge),
- (3) developmentally (first method sequentially informs the second method),
- (4) initiation (contradictions and fresh perspectives emerge),
- (5) expansion (mixed methods add scope and depth to study).

Creswell (1994) built on Green, Caracelli, and Graham's (1989) work by advancing three models of combined designs, one of which will be described here. The dominant-less dominant design was described as a study presented in "a single, dominant paradigm with one small component of the overall study drawn from the alternate paradigm" (Creswell, 1994, p. 177). This type of study presents a consistent picture of the dominant paradigm while still gathering additional information to further

probe one aspect of the study using the less-dominant approach. That is, the introduction, literature review, method, results, and conclusion sections follow the dominant framework with the less-dominant paradigm minimally mentioned, but appropriately defined.

Smith (1994) acknowledged that rationality, rigor, and fairness are viewed as the standards of qualitative research as opposed to the standards for quantitative research: objectivity, reliability, and unbiasedness. Perhaps these standards need not be viewed as opposing forces rather they can be seen as useful determinants in creating a viable research design based on thought-provoking research questions. More and more researchers are viewing mixed method studies as a viable alternative to the all or nothing view held by members in both research camps. With that in mind, Reichardt and Rallis (1994) offered an interesting quote from St. Exupery's The Little Prince in comparing these two paradigms

Grownups love figures. When you tell them that you have made a new friend, they never ask you any questions about essential matters. They never say to you, "What does his voice sound like? What games does he love best? Does he collect butterflies?" Instead they demand, "How old is he? How many brothers has he? How much does he weigh? How much money does his father make?" Only from these figures do they think they have learned anything about him. (p. 8)

Perhaps researchers should ask both types of questions in order to obtain a clear view of the complex environments they study.

A multi-site study was selected for this dissertation study and multiple sources of information have been collected. This research project is a mixed method study conducted by a quantitatively trained doctoral candidate who is by nature a qualitative researcher. Creswell (1998) defined qualitative research as

An inquiry process of understanding based on distinct methodological traditions of inquiry that explore a social or human problem. The researcher builds a complex, holistic picture, analyzes words, reports detailed views of informants, and conducts the study in a natural setting. (p.14)

Creswell (1998) further stated that qualitative research shares good company with quantitative research, but it should not be viewed as an easy substitute for a statistical study. Before considering a qualitative research study, a researcher must be aware of the huge time commitment involved in the field and subsequent data analyses. The researcher must be willing to include multiple perspectives when writing up the results. In addition, the researcher must understand that this type of research is a bit abstract and always evolving. Creswell (1998) mentioned that a qualitative research study employs the researcher as an active learner who can tell the story from the participants' view rather than as an expert who passes judgment on participants. Richardson (1994) and Adler and Adler (1994) used the term "verisimilitude" to capture the type of persuasive writing needed. The research writer attempts to draw the reader into the subjects' worlds. With that in mind, this study was designed using a holistic approach enabling the researcher to obtain information in many ways in order to present a credible and authentic view of the participants' environment.

## 2.6 Summary

Linn (1997) does not believe artificially narrowing the concept of validity is a reasonable solution to the confusion surrounding the meaning of validity. At this point in time, it appears that many measurement professionals generally agree that the appropriateness, meaningfulness, and usefulness of score interpretations are the three

elements that comprise the conceptual validity framework (AERA, APA, & NCME, 1985). Researchers must continue to ensure that the uses and interpretations of tests contribute to enhanced student achievement and at the same time, not produce unintended negative outcomes (Linn, 1994). In addition, researchers must keep in mind that the meaningfulness of score interpretation may come into question when a validity investigation does not include the intended and unintended consequences of an assessment.

While this chapter has outlined the research and related issues involved in this important and infrequently discussed measurement topic, it is clear that a tremendous amount of research time and effort is still needed to further investigate the consequential aspects of validity in small- or large-scale assessment programs. Many studies have been completed; however, much still needs to be done. It is difficult to gather validity evidence without the use of deductive and inductive reasoning and analyses. Therefore, it is appropriate to design some validity studies using both quantitative and qualitative methods. The education profession will greatly benefit from such mixed methods studies because studying the consequences of testing requires researchers to enter into the examinees' environment, the natural setting, in some manner. Shepard (1993) pointed out that validity evidence is often simplistic and incomplete in practice. Linn (1998) stated measurement experts need to "assume the major responsibilities for taking the lead in making it clear why evaluation of plausible effects of testing policies deserves to be given high priority" (p. 29). Evaluating the consequences of high-stakes tests are important and very necessary because of the

impact these tests have on people including students, teachers, policymakers, and employers.

While the balance is currently leaning towards the inclusion of the consequential aspects of validity within the unified validity framework rather than creating a separate and distinct type of validity called “consequential validity,” the debate surrounding this measurement term is far from over. It is important that education measurement professionals foster this discussion and promote continued research in this area.

## CHAPTER 3

### METHODOLOGY

In this chapter, the methodology for the study is presented. The chapter is divided into six sections: introduction, sampling procedure, survey design, focus group protocol, interview protocol, and data analyses. The data analyses section is further broken down into three subsections: survey data analyses, focus group analyses, and interview data analyses.

#### 3.1 Introduction

It is particularly important to address both negative and positive consequences as well as intended and unintended consequences of the studied assessment tool. Linn (1994) stated

If the argument that validation should include an evaluation of the consequences of the uses and interpretations of assessment results is accepted, then it is not sufficient to provide evidence that the assessments are measuring the intended constructs. Evidence is also needed that the uses and interpretations are contributing to enhanced student achievement and, at the same time, not producing unintended negative outcomes. (p. 8)

This study examined the consequential validity of a grade 10 math test as it related to tenth grade mathematics curriculum at the district level, as outlined in the state frameworks, using a mixed method approach. The measurement tools used were a survey, interviews, focus groups, and statistical data obtained at the state, district, and building level. While both quantitative and qualitative paradigms collect information using surveys, interviews, and documents, the actual approach used in the design,

administration, and analysis of each collection device determines the paradigm in which it is most closely associated. The mixed method research design, more specifically known as the dominant-less dominant design, used in this study is reflected in Figure 3.1.

In this study, the survey design, administration, and analyses were connected to the quantitative paradigm with the focus group interviews corresponding to the qualitative paradigm. It is important to note that each focus group interview, a qualitative tool, was designed and conducted using a qualitative genre, but the actual analyses of the focus group sessions took on a mixed approach. Regarding the one-on-one interviews, each was conducted using a standardized set of semi-structured questions and a quantitatively driven numeric system was used in analyzing the information gleaned from each interview. Ethical issues must always be addressed in any project; therefore, consent forms outlining confidentiality were prepared and completed because they were vital to the success of the research study (see Appendix A).

### 3.2 Sampling Procedure

The first step in the study was to collect data, which included identifying variables enabling the classification of all public school districts in Massachusetts. Using the hierarchical cluster analysis procedure in SPSS, similar school districts were grouped by the following identifying variables: per pupil expenditure, average grade 10 mathematics test score, percent of free and reduced lunch, and district enrollment. Then a MANOVA procedure was used to cross validate the groups created in the cluster



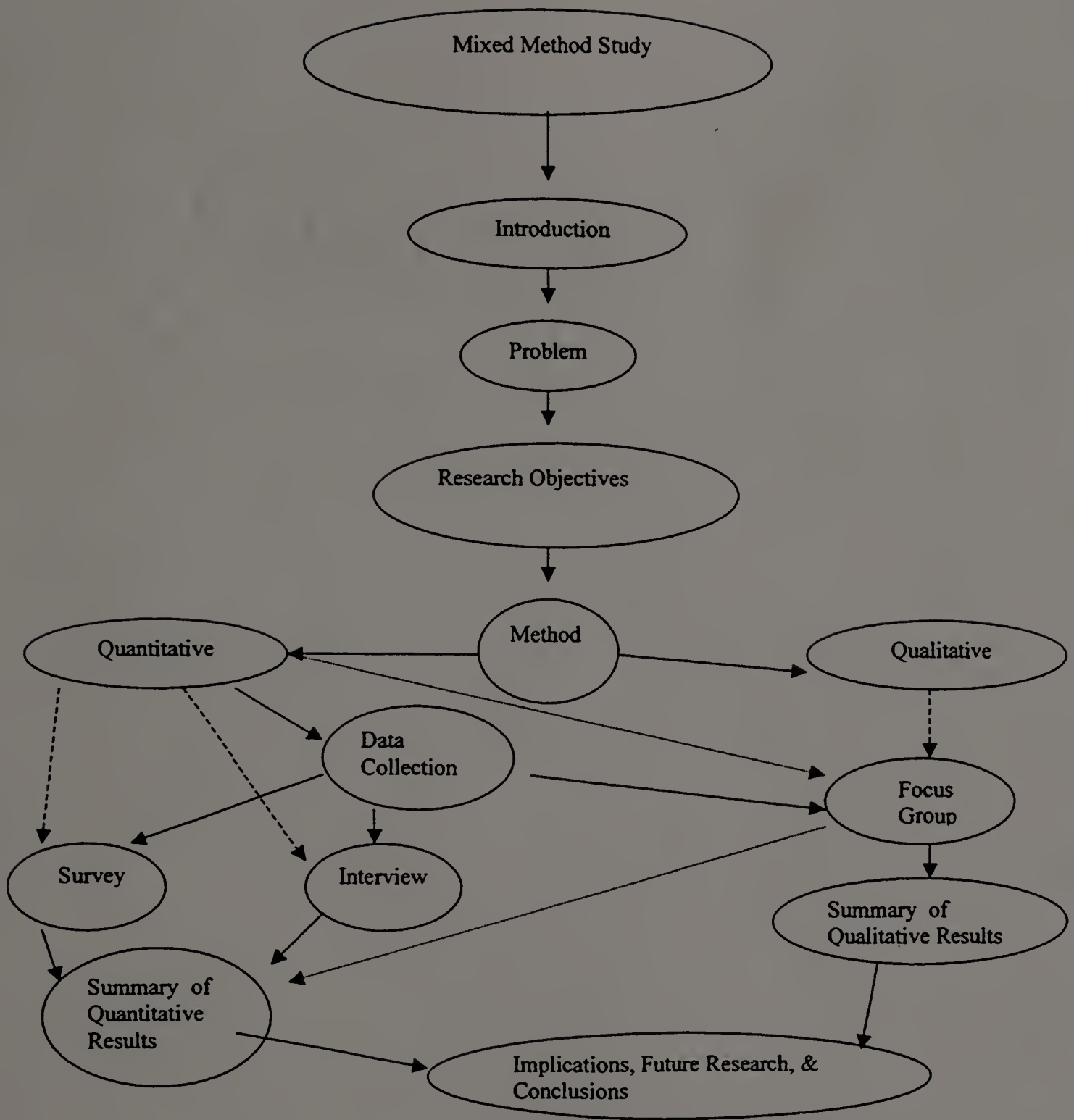


Figure 3.1 A Concept Map of the Research Study

analysis. That is, a statistically significant difference between groups was confirmed using the MANOVA procedure, which validated the 12 classification groups created using the cluster analysis procedure (see Table 3.1).

Table 3.1  
Twelve Group Cluster with Variable Averages

Cluster	Percent Free/Reduced Lunch	Average District Enrollment	Average Day Per Pupil Expenditure	Average Raw Math Score
1	9	1996.22	6394.42	32.02
2	5	4446.39	6326.84	33.09
3	4	3215.58	7558.54	37.61
4	9	7451.90	6584.93	30.40
5	50	63360.50	8880.00	26.00
6	26	15729.43	6990.50	26.64
7	26	5793.50	8601.25	30.07
8	11	7383.30	12623.00	28.37
9	47	12369.95	7060.50	21.95
10	12	774.20	12212.50	34.19
11	4	11201.70	8978.00	41.01
12	53	26538.35	7114.50	21.97

The next step involved choosing a school district from each classification group so each selected district could be analyzed using the measurement tools previously mentioned. The sample included a representative cross-section of Massachusetts' school districts using the categorization developed by the classification process.

A total of 12 school districts were requested to participate in the mixed method study. Telephone calls and direct personal requests were used to confirm each district's voluntary participation in this study. Within each participating school district, surveys were distributed to all of the mathematics teachers and key administrators in the selected high schools, seven or more teachers from each selected high school

volunteered to participate in a 2-hour focus group session, and two or three education specialists or administrators were interviewed. Each focus group member, survey participant, and interview participant signed a written consent form (see Appendix A) outlining confidentiality rules and regulations for this research study.

### 3.3 Survey Design

Once a selected school district agreed to participate in this study, a “point” person was selected within each high school to serve as the survey distribution and collection coordinator. Typically, that person was the mathematics department chair. A mailing packet was sent to each “coordinator” for subsequent distribution to each survey participant. The packet included a cover letter, letter of support, one postage paid return envelope, and a survey for each participant. Appendices B and C contain the cover letter and letter of support, respectively. The coordinator then distributed a survey to each high school mathematics teacher, mathematics curriculum specialist, and key administrator in the participating high school.

The purpose of the survey was to learn about the opinions and perceptions of the high school mathematics teachers, specialists, and administrators related to the state-mandated mathematics assessment and its corresponding curriculum standards. A two-sided two-page survey was developed. The survey was kept brief to encourage a high return rate, yet it covered as many aspects of the assessment system as possible. With that in mind, survey questions inquired about a participant’s view of the impact the state-mandated test had on his/her district’s high school mathematics curriculum including course sequencing and offerings, student achievement including the

difference between student/class mathematics grades and scores on the state-mandated test, instructional practices, and teacher motivation and morale.

After several initial changes were made to the “draft” survey, the instrument was pilot tested in February 2002. The survey was piloted tested to determine areas needing improvement. Each of the five individuals who agreed to be a reviewer was a high school mathematics teacher, education administrator, and/or curriculum specialist. The group consisted of two central office administrators, two high school mathematics teachers, and one mathematics department head.

The reviewers’ tasks were outlined as follows: (a) complete the survey under “survey-like” conditions making note of his/her completion time, (b) review all aspects of the survey with the following questions offered as assistance in their critique: are there important questions missing? should any of the questions be deleted? and is the instrument clear, concise, and easy to complete? In addition, a stamp self-addressed envelope was included to maximize the return rate. Finally, each reviewer completed the survey, provided an insightful critique, and returned the instrument in the envelope provided.

Several changes were made to the survey as a result of the pilot testing. For example, all of the non-teachers indicated that the question asking for title of position would definitely identify each of them, as there was only one “department head or assistant superintendent.” Therefore, this question was removed. In addition, two questions inquired about explaining MCAS test results to students and parents, all of the reviewers indicated that teachers do not perform this function. These two questions were removed, also. Several grammatical and/or semantical changes were made to add

clarification to many questions. A few questions were added so that MCAS remediation efforts at the district level could be explored and documented. The majority of reviewers indicated that it took them approximately 12 to 15 minutes to complete the survey, which seemed reasonable so no substantial change was made to the length of the instrument. The final survey consisted of 32 selected-response items and two open-response questions and it is included in Appendix D. The same survey was prepared for non-teachers with subtle changes made to the wording of the questions (see Appendix E). For instance, a question on the non-teachers version inquired about “the teachers in your district.” Each survey yielded 104 variables that were analyzed.

Follow-up phone calls, faxes, and/or emails with letters as attachments were sent to each “coordinator” on a regular basis to ensure a strong participation rate in the interviews and focus groups, and to ensure a high survey return rate.

### 3.4 Focus Group Protocol

Each school district was asked to convene a group of seven or more mathematics teachers for participation in a two-hour focus group discussion. Participation was voluntary. Focus groups are a type of in-depth interviewing where the researcher encourages group members to express different beliefs and points of views (Marshall & Rossman, 1999). The focus group protocol used in this study was carefully created to maintain consistency between group interviews and to foster a rigorous procedure that was precise and reproducible.

The researcher remained neutral during the focus group sessions acting as a facilitator rather than a leader. Much research and preparation went into the facilitator’s

delivery style and appropriate technical terms to use. Then a pilot focus group session was conducted using three doctoral students who were also certified teachers. While most of the feedback from this process was positive, the focus group protocol was slightly changed to ensure success.

Since the education reform initiative in the state of Massachusetts continues to stir up strong positive and negative opinions as well as a wide range of political viewpoints, the facilitator attempted to be perceived by all participants as a credible, trustworthy, good listener. The questions were purposely open-ended to encourage discussion among group members; however, a broadly stated introductory statement, which included each respective district's average mathematics test scores from the previous four years were used to set the tone and direction of the group discussion. In addition, the facilitator used subtle conversation management techniques to keep the group on topic. Each session was recorded with the group's permission as outlined in the written consent form (see Appendix A). Subsequently, a transcript of each session was prepared using each respective tape. The focus group interview protocol in its final form has been reproduced in Appendix F

### 3.5 Interview Protocol

Any interview method offers advantages and disadvantages to a research study. The results gathered using an interview method generally yield more complete information (Jackson & Rothney, 1961); however, it is difficult to standardize the interview environment. With these issues in mind, a semi-structured interview technique outlined by Gall, Borg, and Gall (1996) was used in this study. A group of

structured questions occasionally followed-up with more probing open-ended questions was used. This technique provided the researcher with standard information across respondents, while also enabling her to gain more in-depth information when needed. That is, the interviews were very structured. The order and wording of the questions were predetermined and standardized across interviews to promote consistency and reproducibility; however, the interviewer did have subsequent open-ended follow-up questions that were used to glean more detail from a participant, when necessary.

Gall, Borg, and Gall (1996) stated researchers should consider how they will handle the following interview tasks: (1) determine how to present oneself, (2) establishing rapport, (3) gaining trust, and (4) being sensitive to nonverbal information. In addition, the guidelines for conducting a research interview, which were outlined by Gall, Borg, and Gall (1996), were used in the construction of the interview protocol.

Previous research has indicated that interviews are very susceptible to bias; therefore, the interview guide and procedures were given a pilot test to determine how a high school educator would interpret the interviewer's style and questions. Using the feedback developed from the mock interview, the final version of the interview protocol used in this study was developed and it is contained in Appendix G. It is important to note that the opening statement and interview questions were carefully prepared and used consistently throughout each interview.

Key education personnel were selected to participate in an individual interview based on common names given to the researcher during telephone calls to each participating district's superintendent, high school principal, and curriculum coordinator/specialist. A telephone contact summary sheet (see Appendix H) was

developed to maintain consistent documentation and to promote meaningful contacts within each school district. Once key personnel were identified, interviews were scheduled and subsequently conducted.

### 3.6 Data Analyses

#### 3.6.1 Survey Data Analyses

The data for all survey respondents were coded and entered into the computer via the SPSS statistical software package. Descriptive statistics were computed for all selected-response survey questions. Some inferential statistical comparisons were made across school districts. Statistical comparisons were also made to test for significant differences in curriculum and instructional practices with respect to the past, present, and future as a result of the MCAS assessment system. An alpha of .01 was used in all tests of statistical significance. Effect sizes (eta-squared) less than .10 were considered negligible; effect sizes between .10 and .30 were considered small; effect sizes between .30 and .50 were considered moderate; and effect sizes greater than .50 were considered large (Cohen, 1988). A content analysis of each open-ended question was performed. Each analysis identified and coded common themes across the respondents.

#### 3.6.2 Focus Group Data Analyses

Information obtained during the focus group interviews was summarized using an adapted pen and paper technique (Babbie, 1988; Stewart & Shamdasani, 1990). This technique involved the coding of various themes using colored pencils. Content analysis was conducted on each focus group transcript by first reviewing the content of



each session, and then deriving themes from the participants' responses. Briefly, the technique involved counting the number of times a given theme (word or phrase) appeared in all of the transcripts. The analyses were guided by the questions included in the focus group protocol (see Appendix F). All key phrases and comments were color-coded for each transcript, and then the transcripts were compared to one another. This process was developed to promote consistency in coding and to add depth to the information collected from the focus group sessions. These analyses were completed at the district level and then aggregated at the state level.

### 3.6.3 Interview Data Analyses

To summarize the information obtained in the interviews, content analyses were conducted on the transcripts by first reviewing the content of each interview, and then deriving themes from those responses. All key phrases and comments were coded during the content analyses of the first transcript. Subsequent transcripts were coded by comparing them to the transcripts that were previously coded. Themes were developed by grouping similar comments made by different respondents into the same category for each respective interview question. This methodical process was followed to promote consistency within the information gleaned from the interviews.

The analyses were guided by the research questions previously outlined in Chapter 1: (a) Have changes in curriculum occurred within high school math departments across the commonwealth as a result of the statewide grade 10 mathematics assessment and its corresponding curriculum framework? (b) Has high school mathematics instruction changed due to the state's mathematics curriculum standards?

(c) Has district high school mathematics curriculum been aligned with the state's mathematics curriculum framework? More specifically, the questions included in the interview protocol (see Appendix G), which were developed from the three research questions, played a pivotal role in these analyses. These analyses were completed at the district level and then aggregated at the state level.

## CHAPTER 4

### RESULTS

In this chapter, the results of the survey, focus group sessions, and interviews are presented. The results of the survey analyses are outlined in section 4.1, and then information gleaned from the focus group sessions and interviews follows in sections 4.2 and 4.3, respectively.

#### 4.1 Survey Results

The survey results section is broken into five sub-sections: participant response rate, background information about the survey participants, opinions related to aspects of the MCAS assessment system, curriculum and instructional practices with respect to time, and content analyses of open-ended questions.

##### 4.1.1 Participant Response Rate

Of the twelve target high schools, 11 participated in one or more components of this study. The surveys were distributed to the math teachers and collected by each respective department head, which ensured a better return rate. Each individual high school response rate is listed in Table 4.1. The total participation rate for the 11 schools participating in the survey component of this study was 56.2%. That is, 116 surveys were completed and returned out of a possible 206. The individual participation rate at each high school is outlined in Table 4.1.

Table 4.1

## Survey Participation Rates

School District	# of Surveys Completed	% of H.S.	% of Total
A	11	52.4	5.3
B	9	52.9	4.4
C	18	60.0	8.7
D	12	100.0	5.8
E	12	85.7	5.8
F	11	84.6	5.3
G	11	44.0	5.3
H	10	100.0	4.9
I	9	25.0	4.4
J	9	36.0	4.4
K	4	100.0	1.9
Total	116		56.2

4.1.2 Background Information

Fifty-three percent of the survey participants were males and forty-four percent females with three participants not identifying their sex. Eighty-seven percent of the respondents were high school mathematics teachers with 54.3%, 41.3%, and 3/4% working in urban, suburban, and rural school districts, respectively. The number of students enrolled in each of the 11 school districts has been outlined in Table 4.2. Regarding the grades taught by these teachers, approximately 75 percent taught 9th,

Table 4.2

## Student Enrollment in Participating Schools

School District	Approximate Student Enrollment
A	1250
B	1100
C	3000
D	3000
E	1200
F	1050
G	900
H	2200
I	180
J	2500
K	1500

10th, and/or 11th grade with 60% teaching 12th grade.

The math teachers taught a wide range of mathematics courses. For reporting purposes, courses were grouped according to similarities in subject matter (see Table 4.3). For instance, Algebra 1, Algebra A or B, Honors Algebra 1, and Standard Algebra 1 were grouped into the "Algebra 1" category. The three largest groups were Algebra 1 (63%), Algebra 2 (42%), and Geometry (47%), with the participants teaching one or all of these courses. Thirty percent of the survey participants have taught 5.5 years or less; 14% have taught for 6 to 10.5 years; 9% from 11 to 15.5 years; 6% from 16 to 20.5

Table 4.3

## Categories of Courses Taught by Teachers

Course Category	# of teachers	% of total
Algebra 1	66	63%
Geometry	49	47%
Algebra 2	44	42%
Precalculus/Calculus	28	27%
Probability, Statistics, Analysis & Trigonometry	20	19%
Integrated Math 1 & 2	7	7%
AP Courses	7	7%
MCAS Remediation	6	6%
Computers	5	5%
General/Technical Math	4	4%
SAT Preparation	1	1%

years; 8% from 21 to 25.5 years; and 33% have taught for 26 or more years (see Table 4.4). It appeared the bulk of the teachers were at the extreme ends of experience. That is, they were very new to the profession or they were veteran teachers approaching retirement.

Table 4.4

## Years in Teaching Profession

Years	# of Teachers	Percent of total
0-5.5	32	30%
6-10.5	15	14%
11-15.5	10	9%
16-20.5	6	6%
21-25.5	9	8%
26 +	36	33%

With respect to the level of education attained by the teachers, approximately half had a master's (55.2%) and the others had received a bachelor's degree (37.1%), Certificate of Advanced Graduate Studies (CAGS) (4.8%), or doctorate (1.9%).

Regarding the administrators, there were 63.6% males and 36.4% females and a little more than half of them had attained a doctorate (54.5%). The remaining administrators had received a master's (27.3%) or CAGS (18.2%).

#### 4.1.3 Opinions Related to Aspects of the MCAS Assessment System

Survey participants were asked about their opinions regarding the impact the Massachusetts' Mathematics Curriculum Frameworks and assessment system had on their math curriculum and instructional practices. Ten opinion questions used a five-point Likert-type format ranging from "strongly agree" to "strongly disagree."

Generally, the participants had positive perceptions of the impact the education reform

initiative had on their math departments. More specifically, the majority of teachers “agreed” they had received adequate professional development at the district level for implementing the Mathematics Curriculum Frameworks and their own curriculum had changed to better reflect the frameworks; however, they “disagreed” when asked if the state had provided adequate professional development for framework implementation.

In addition, they “agreed” their instructional practices had changed due to the math frameworks and their students’ performance on the MCAS grade 10 math test. Sixty-three percent of the respondents agreed to strongly agreed course offerings had been changed due to the MCAS assessment system, while 48% agreed to strongly agreed course sequencing had changed with 26% undecided on this issue. Forty-two percent of the teachers indicated that the number of students enrolled in more difficult mathematics courses had remained the same. Approximately 75% of the participants agreed to strongly agreed their high schools were using MCAS test results to identify students who may need specialized instruction and had provided remediation instruction and/or courses for those students who had not passed the grade 10 assessment.

When asked about students’ course grades mirroring their performance on the MCAS grade 10 mathematics test, 42 % indicated “yes,” 24% responded “no,” and the remaining 29% were not sure about the correlation. If teachers indicated that their grades did not mirror MCAS test results, they were asked to comment on the reasons for the difference. Twenty-two of 116 teachers responded. Of those responding, 54% wrote about the test’s lack of fairness and/or test anxiety creating the difference. In addition, a few teachers indicated that their courses included a wide range of assessment tools that enabled them to obtain a more accurate view of their students’ knowledge. A



few more teachers indicated that MCAS test scores mirrored students' course grades at each end of the spectrum with the middle being less of a match.

It is interesting to note there was significant diversity of opinion regarding the MCAS grade 10 mathematics assessment being a useful tool to improve the long-term education of students. Thirty-three percent disagreed to strongly disagreed; 33% were undecided; and 32% agreed to strongly agreed about the long-term usefulness of this assessment. Thirty-nine percent of the teachers did not believe the education reform initiative in the state of Massachusetts had changed how they teach for the better with 36% unsure.

Teachers (78.4%) were very clear in stating that averaging all student performances on the MCAS grade 10 mathematics test was not a good indicator of their high school's overall educational success in teaching mathematics to its students. In addition, 61.2% indicated classroom time for test preparation of the MCAS grade 10 math assessment had meant other valuable topics were not being taught. Finally, teachers were asked to describe their morale as mathematics teacher based only on their district's implementation of Massachusetts' Mathematics Curriculum Frameworks. Fifty-five percent indicated their morale was poor to fair, 31% good; and the remaining 14% were not sure or did not respond to the question.

Survey participants were asked if the Massachusetts' education reform initiative had ensured that all students received the best possible education. Forty-nine percent of the respondents indicated that it had not done so with 34% unsure. Two major themes emerged from the teachers' comments on the issue of the education initiative ensuring that all students receive the best possible education. Two-thirds of the comments were

about the test being unfair to all students or specific subgroups of students. Examples of unfairness included flawed test items that were too wordy, too difficult, or not important applications of math concepts. The remaining 1/3 of the teachers' comments revolved around the disadvantage to lower and upper level students caused by the change in the number of levels within each course. Many teachers wrote about below average students struggling to keep up, while the more capable students "twiddled their thumbs."

Survey participants were asked, "Do you believe the current system of MCAS score reporting at the high school level is adequate?" Forty-six percent responded "no," with the remaining 50% split between "yes" and "not sure." Respondents were then asked to comment on how the score reporting could be improved. Of the 65 respondents, 50% indicated that the MCAS grade 10 math scores needed to be reported faster. The remaining respondents wrote about a wide range of reporting issues including the need to distribute individual student information at the classroom level and the need to include information about diversity, language spoken at home, and/or special education.

#### 4.1.4 Curriculum and Instructional Practices with Respect to Time

Respondents were asked to rate the relative influence nine different statements had on their high school mathematics curriculum and instructional practices. Non-teachers were asked to indicate their perception of the relative influence each statement had on the math teachers in their building. The scale for the influence ratings ranged from zero (no influence at all) to five (primary influence). For each of the nine

statements, each respondent was asked to rate (a) the relative influence that aspect of the state's Mathematics Curriculum Frameworks and assessment system currently has on his/her math instruction, (b) the relative influence that aspect had on his/her math instruction three years ago, and (c) his/her opinion regarding the influence that aspect will have on his/her math instruction three years from now. Table 4.4 contains the mean response for each "time" category: current, past, and future, along with additional statistical information, which will be discussed further in this section.

All nine statements were currently perceived to have a strong influence on the math teachers with mean ratings of 3.5 to 4. These statements related to teachers' perceptions of the influence various aspects of the MCAS assessment system had on their own curriculum and instructional practices. These surveys items inquired about the influence of successfully matching the state's frameworks to the curriculum taught by them and successfully implementing instructional changes to better reflect those standards. In addition, respondents were asked about the influence the following four factors had on them as educators: his/her commitment to adequately preparing students to take the MCAS grade 10 mathematics test, the past performance of students, improving students' MCAS math test scores, and the motivation of their students to do well on the assessment. They were also asked about the degree of influence their own personal motivation had on teaching mathematics as outlined in the state's Mathematics Frameworks as well as the influence administrator's pressure to obtain the best MCAS math test results had on them.

In addition, one-way repeated measures analyses of variance (ANOVAS) were computed for each statement, using time (past, present, and future) as the within-

subjects factor. First, all of the individual participant responses were aggregated for each of the 11 high schools. Second, the eleven mean influence ratings were input into a database for each of the time periods investigated. Keep in mind, “time” was the repeated measure. The results are summarized in Table 4.5. For each test, the mean rating with respective standard deviation for current, past, and future is listed in addition to the F-statistic, significance value (p), and effect size. The effect sizes for the eight statistically significant statements were large (.85 to .96), which adheres to the effect size guidelines previously described in the data analyses section of this study’s methodology chapter (Cohen, 1988).

All of the statements except one were viewed as having more influence today and/or in the future than three years ago. The one aspect that was shown to have no statistically significant difference in relation to time was the teacher’s success in implementing instructional changes to better reflect the state’s Mathematics Frameworks. The remaining eight aspects or statements, which were outlined in the previous paragraph, were weighted as having more influence in the future than in the past. This suggested that many aspects inquiring about the influence of the reform initiative were viewed as increasing in influence in the future compared to the past.

Four statements/aspects were shown to have more influence today and would continue to have more influence in the future than in the past. The four aspects shown to have more influence today with increasing influence in three years follow: commitment to adequately preparing students to take the MCAS grade 10 math test, matching the state’s Mathematics Frameworks to the curriculum taught, motivation of

Table 4.5

## Descriptive and Inferential Statistics for Influence Statements

Statement	Current Influence	Influence 3 Yrs. Ago	Future Influence	F (2, 11)	p	$\eta^2$
Commitment to adequately preparing students to take the MCAS grade 10 mathematics test*	3.63 (.47)	2.10 (.60)	3.89 (.53)	98.61	.000	.96
Matching Massachusetts' Mathematics Frameworks to the curriculum that I teach*	3.80 (.52)	2.65 (.49)	3.99 (.58)	25.20	.000	.85
Success in implementing instructional changes to better reflect the state's Mathematics Frameworks	3.42 (.45)	2.72 (1.91)	3.68 (.60)	3.65	.069	.45
Preparing students to take the MCAS grade mathematics test**	3.75 (.49)	2.18 (.30)	3.92 (.59)	70.46	.000	.94
Student performance on the MCAS grade 10 math test**	3.51 (.61)	2.05 (.54)	4.11 (1.13)	57.19	.000	.93
Improving students' mathematics test scores on the MCAS grade 10 math assessment**	3.71 (.58)	2.21 (.60)	4.28 (1.25)	73.55	.000	.94
Motivation of my students to do well on the MCAS grade 10 mathematics test*	3.63 (.42)	1.89 (.60)	3.92 (.50)	104.50	.000	.96
Pressure from administrators to obtain the best MCAS math test results from my students**	3.88 (1.79)	2.30 (.84)	3.77 (.90)	71.75	.000	.94
Motivation to teach mathematics as outlined in the Massachusetts Mathematics Frameworks*	3.60 (.67)	2.37 (.43)	3.71 (.66)	25.22	.000	.85

Notes: The influence scale ranged from 0 (no influence at all) to 5 (primary influence). Each mean rating is followed by its respective standard deviation in parentheses.

\*Indicates future > current > past.

\*\*Indicates future > past.

students to do well on the MCAS grade 10 mathematics test, and motivation to teach mathematics as outlined in the state's Mathematics Frameworks.

#### 4.1.5 Content Analyses of Open-Ended Questions

The content analyses of the two open-ended questions were conducted by analyzing, identifying, and then coding common themes across respondents. A total of 99 participants (85%) responded to one or both of the questions. During the content analyses, it became apparent that the same themes were emerging for each question. As a result, the themes were merged and have been outlined in the next few paragraphs.

The two open-ended questions inquired about respondents' opinion/concerns related to the consequences of the MCAS grade 10 mathematics assessment and the impact of the state's Curriculum Frameworks on high school math instruction. Both positive and negative themes emerged related to these issues. The negative themes included three major and two minor concerns, while two major positive issues were indicated and have been discussed later in this section. The major negative themes consisted of the (1) MCAS grade 10 mathematics test being unfair, (2) dissatisfaction with the make-up of the test and concerns about teaching to the test, and (3) one test not being a fair measure of ability.

Twenty-one percent of the respondents believed the test and corresponding assessment system were unfair to certain subgroups of students, such as: English as second language, special education, socially and economically disadvantaged. Many described the raising of the bar to average college preparation as an unrealistic goal for many students. That is, the one size fits all curriculum had not worked for many

students. They also indicated the math test was too dependent on reading comprehension. They further explained that too much classroom time was spent on non-achievers or non-performers. Specifically, their behavior disrupted or interfered with student learning. This situation was exacerbated by an assumption all students were college bound and were interested in taking Algebra 1 and Geometry. When often times, students' interests were reflected in their effort, grades, and ability. Many teachers observed students thrown into math classes that were above their ability level, which translated into low-level students not receiving necessary instruction on basic mathematics concepts. The self-esteem of these low-end students was a concern due to this phenomenon.

In addition, an interesting twist to the test being unfair to certain subgroups was the perceived effect it had on average to above-average students. Teachers indicated high ability math students were at a disadvantage because less and less material was covered due to heterogeneous classes, which meant they were at a disadvantage when they eventually enrolled in precalculus or advanced placement math classes. This was stated as a watering down of math instruction or a de-emphasize on mainstream and high-end students.

Twenty-nine percent of the respondents were dissatisfied with the make-up of the test. They did not like the test items. They viewed them as too wordy, too complicated, and even "tricky." They believed important math concepts were being overlooked and more practical, realistic math situations should be examined. They further stated too much attention was given to past test questions and answers rather than the curriculum, which resulted in teaching to the test or test items. Because they

did not like the test items this was seen as an especially negative situation. They had focused more and more energy on test-taking strategies, rather than developing critical math skills and helping motivate students in math.

Twenty-two percent of the respondents believed one test, as a measure of math ability, was unfair. That is, one assessment cannot tell the whole story. Multiple assessment measures must be considered especially with complex mathematics concepts. They mentioned this one test highlighted previously documented demographic differences in educational performance among students, which served no educational purpose.

The remaining minor negative issues were the loss of teaching time due to the actual test administration (12%) and the grade 10 math assessment placing too much pressure on students and teachers (15%). Respondents outlined the difficulty involved in covering more material in less time with each class containing a wider range of student ability. Many respondents viewed the pressure placed on students as detrimental to their education. It hampered students who had already developed “math anxiety” even more. Teachers also experienced increased pressure to improve student test scores from many groups, such as: media, administrators, and parents.

The two major positive themes that emerged from the content analyses were the implementation of statewide mathematics standards and increased student accountability. Twenty-eight percent of those responding believed the implementation of the statewide K-12 Mathematics Curriculum Frameworks, which was also defined as “math standards”, improved their district’s math curriculum and indirectly improved instruction. Improvement was described as increased communication among math



educators, professional development, regular curriculum reviews, and textbook alignment with frameworks. Twenty percent of the respondents viewed increased student accountability as a positive consequence of the education reform initiative. Many students were exerting more effort and motivation, and it was hoped that overall student performance would increase as a result. Respondents indicated that students now knew when they needed remediation. The increased student accountability also created an increased awareness of the consequences of the education reform initiative among parents and teachers.

#### 4.2 Focus Group Results

The focus group results have been summarized using the content analyses procedure outlined in Chapter 3 (see Section 3.6.2). After reviewing the content of each focus group transcript, themes were identified, coded, and counted. Six high schools participated in the focus group component of this research study. Each of the six high schools was assigned a letter of the alphabet (e.g., School district “A”, School District “B”) to ensure anonymity. Three of the six school districts were in urban settings, while the remaining three were in suburban settings. The average number of focus group participants in the urban and suburban schools was 17 and 8 teachers, respectively. It is interesting to note that each respective department head participated in each of the suburban focus group sessions, while none of the department heads attended the urban focus group sessions.

Each district’s focus group transcript was analyzed and six major themes emerged. These issues evolved from the implementation of the MCAS grade 10

mathematics test and its corresponding K-12 frameworks: changes in curriculum, changes in instruction, educators' opinion of a standards-based curriculum, morale and motivation of educators, improvement in student learning, and suggested changes to the assessment system. These six major themes can be clearly connected to the questions outlined in the focus group protocol outlined in Chapter 3 (see Appendix F). Emergent themes emanating from the focus group protocol are discussed in terms of urban and suburban school districts in the following six sub-sections.

#### 4.2.1 Changes in Curriculum

##### 4.2.1.1 Urban Schools

Curriculum changes had occurred in varying degrees at all three urban high schools. MCAS preparation courses had been developed in two of the three urban high schools. Students who had failed the MCAS grade 10 math test had been enrolled in the MCAS prep course in each of the two urban high schools. This is in addition to their regular mathematics courses.

A wide variety of MCAS remediation help had been implemented at all three of the urban schools. Computer software used for self-paced instruction, one-on-one tutorial help, and small class size were some of the methods used to implement remediation help. The majority of urban teachers stated the after school and Saturday sessions had not appeared to be as effective in helping students. They explained this was due to students' extracurricular activities, their employment schedule, or their general lack of interest. For example, urban school district "A" had volunteers help out on Saturdays, yet as these volunteers found few students showing up, they slowly

started showing up less and less. The teachers participating in these programs indicated that students who did show up were not always interested in learning. Often times they wanted their school records to reflect participation, but their interest in truly learning was not there. It appeared that most of these programs were implemented the previous school year. It was not made clear whether these programs continued today, but these examples outlined the reasons teachers found these programs to be less successful than originally planned.

Due to the problems with the previously outlined programs, teachers from two of the three urban high schools indicated that their high schools had implemented “pull-out” programs during the school day. These programs were set up after the Massachusetts Department of Education amended wording in their grants to include remediation programs that occurred during the school day. The majority of teachers indicated that “school day” remediation efforts appeared to be more successful because the students were being reached during the normal school day.

Regarding specific changes in curriculum within each of the urban mathematics departments, all of the urban teachers indicated they had changed the order in which topics were presented to better prepare their students for the MCAS grade 10 mathematics test. That is, grade 10 math teachers covered certain chapters before the administration of the assessment. For many, this presented a problem in that teachers found many students confused because they had not retained critical concepts from a previous math course or the act of skipping around within the textbook meant critical concepts were not presented in a thoughtful manner. More specifically, many teachers did not have the luxury of stopping and reviewing an old concept or a more difficult

concept because they felt pressured to cover the material at an inappropriate pace. This was especially true because the majority of teachers at all three urban schools stated there was an increase in poorly prepared students in all of their classes. While they understood one goal of the education reform initiative was to increase the number of students enrolled in Algebra 1 and Geometry, the actual result was that course levels within two of the three urban high schools were condensed which resulted in more heterogeneous classes. Stronger students were made to work at a slower pace and weaker students fell further and further behind.

One urban teacher from school district “C” stated, “The notion that the top brings the bottom up is ridiculous. And you have people in high places pushing that idea and people who have not been in the classroom to watch what really happens [in that situation.]” Another teacher in the same district added

I think maybe some of the lower level kids in some situations can be helped but in general I think everybody’s hurt. I think the kids at the lowest levels are hurt because they’re put into classes and they can’t keep up. How does that help their self-esteem? It just makes them feel more dumb....The smart kids are bored to tears and basically hate school....They’re not being challenged, so we’re basically wasting their brain power and the kids in the middle are just sort of bopping along. And the teachers are going insane because how can we possibly deal with all of these kids at all of these levels and put more of them in a class and expect us to be able to meet all their needs in some creative way....and cram the curriculum in....and have them do projects....and have them pass MCAS.

These two statements reflect major concerns mentioned by a majority of the urban teachers who participated in a focus group.

In school district “A”, teachers discussed confusion around their understanding of two integrated math courses, Integrated Mathematics Program and Connected Mathematics Program, which were implemented due to the education reform initiative.

Then one teacher who had previously taught the middle school version of these courses clearly explained that these courses were not created for the low-ability student rather these courses were intended to present math concepts in a more holistic or realistic manner. The instructional methods used were supposed to help students who needed a more hands-on approach to learning math concepts. Much discussion followed regarding students and teachers' confusion around how these courses fit into their curriculum and their general success. In general, these two courses received mixed reviews from the teachers in school district "A."

Teachers in school district "C" indicated general/technical course offerings had been dropped from their schedules, while no changes in course offerings had occurred at school district "B." One teacher in school district "C" exclaimed, "Now it's almost stripped of everything." Another teacher in the same district added, "We used to have consumer mathematics courses, which were a type of "life skills" math."

As previously mentioned, all three urban math departments indicated that the order in which concepts were presented within each course had been changed to better align with the content of the MCAS grade 10 math test and frameworks. These changes had occurred after each district reviewed the K-12 mathematics curriculum framework. The actual review process varied from district to district and the specifics of that process were not investigated. Each math department head appeared to play a pivotal role in dispersing MCAS information to teachers. Frequent memos had been prepared and distributed to all teachers and MCAS information and instructional techniques were discussed at most department meetings.

#### 4.2.1.2 Suburban Schools

Results emanating from focus group sessions at each of the three suburban high schools were similar in that course offerings had been changed and course sequencing had not. Each discussion then turned to curriculum alignment within each suburban high school. Teachers from school district “D” indicated that their curriculum and textbooks were strongly aligned with the K-12 mathematics curriculum frameworks due to the two district-wide curriculum reviews completed in the recent past. In addition, the teachers in the high school math department in district “D” examined every question over the past three test administrations and found a total of approximately five questions that weren’t addressed in their curriculum. They immediately made changes to include these overlooked concepts.

Math teachers in the remaining two school districts, “E” and “F”, stated some progress in curriculum alignment had been made, but more work needed to be done. That is, some of their textbooks were aligned with the frameworks; however, more attention needed to be given to the actual curriculum being taught and the match between course textbooks and the state’s mathematics frameworks. For example, a math teacher from school district “F” stated

We have not worked on curriculum for several years. So we have not aligned formally, while we have completed some work informally.”  
Another teacher from the same district further explained, “We have very little time where we have common planning time...it’s kind of hard to align things in our own subject area. Fortunately, there are four of us who are teaching Algebra II this year who have a common time that we can sometimes use...from what I understand there’s not going to be any time this summer for curriculum planning either. So again—there we go again. You do it on your own.

All three high schools had experienced a decrease in the number of course levels offered. The teachers explained that this action stemmed from the education reform initiative requiring the deletion of general/technical math courses, which resulted in more heterogeneous classes within the Algebra 1 and Geometry courses. There was not much discussion about the changes in the number of levels. It is important to note that school district "E" had a Special Education teacher go into certain math classes to assist with instruction for the first time during the 2001-2002 school year. This action seemed to be supported by the majority of math teachers present.

Regarding specific remediation efforts, school district "D" implemented a tutoring program during the 2001-2002 school year whereby students who had failed an MCAS math test were taken out of extended learning classes, also known as "study halls", to receive additional instruction via a computer software program. Math teachers were disappointed their input was not sought before this computer program was purchased. They believed tutoring via a computer software program would not be very effective due to the fact that most students who struggled with the MCAS math test needed one-on-one instruction or at least small group instruction. They stated self-paced learning would probably not work for the majority of students failing the MCAS math test in their district.

School district "F" had also implemented a "pull out" program for those students who had failed the MCAS math test. One teacher reflected upon one shortcoming involved with this program, "I have a student who had a dilemma. He told me that he had a chemistry test today first period but he was supposed to attend an MCAS help

session. He asked me what he should do.” This district’s implementation of the MCAS remediation program sometimes created serious scheduling issues for some students.

Placement of 9th grade students into appropriate math courses was mentioned as a concern during the focus group session at school district “E.” Teachers indicated scheduling decisions were made without the benefit of the grade 8 MCAS math results. In addition, focus group participants in school districts “D” and “E” discussed weak lines of communication between their high school math department and each respective middle school math department. There continued to be an on-going concern about ninth graders arriving at high school ill prepared for Algebra 1. Teachers in school district “E” discussed the fact that the lower level courses were no longer available due to the MCAS initiative. Teachers in school district “F” discussed the same issue and they agreed that a lot of the students coming through their classes were not finding Algebra 1 and Geometry to be an attainable goal. While it was agreed this particular issue was not new to education, it was exacerbated by the MCAS graduation requirement. Many of the high school math teachers indicated that student success on the MCAS grade 10 math test was viewed as their responsibility. In other words, they felt the bulk of the burden on their shoulders rather than dispersed throughout the K-12 system.

#### 4.2.2 Changes in Instruction

##### 4.2.2.1 Urban Schools

Mathematics instructional techniques had definitely changed at all three urban high schools. Teachers stated that more open-ended questions were presented on tests, during classroom discussions, and homework assignments. A wide range of positive



and negative opinions was expressed regarding these types of questions. Many teachers believed open-ended questions enabled them to present more realistic problems to their students. Others indicated that the process of obtaining an answer was stressed as most important, but there would come a time in their students' lives when they would be required to obtain the correct solution or answer. That is, students needed to know and understand fundamental multiplication and division facts, etc. Always offering partial credit on quizzes and tests because a student had a portion of the process correct may send the wrong message to some students.

In school district "C", teachers explained that students were being exposed to a wide range of concepts in the elementary and middle school years, yet not enough about any one concept to really be able to go on. One teacher in school district "C" stated, "They're weak in everything --Jack of all trades and master of none." Teachers in the remaining two urban high schools indicated similar concerns surrounding the knowledge and ability of incoming students. In addition, it was consistently indicated that strong students will always be motivated to do well, but these students should be allowed to move at their own pace rather than being forced to move at a slower pace.

The more heterogeneous classes presented a real challenge to teachers who were trying to explain more complex mathematics concepts. Often times, they found themselves moving forward in the curriculum and instruction even though all students had not understood or absorbed what was being taught. This is reflected in a comment from a teacher in school district "B", " We do see an increase in students enrolled in algebra/geometry courses regardless of ability, knowledge, and preparation. These

students move through these courses without truly understanding many of the concepts.” Another mathematics teacher in school district “B” stated

In the past, a teacher would stop, review, and re-teach a concept that presented a problem to the kids. Today, we need to move forward, regardless of understanding, so that all of the information is at least presented before the MCAS test is administered.

School district “C” implemented a block schedule during the 2001-2002 school year. While the first year of any new program was viewed as a transition period, many of the math teachers in this district indicated their concern about covering a whole year of algebra in a half-year course. Regardless of lengthier classes, many of these teachers did not feel they were able to cover the same amount of material in a half year. In addition, teachers from this district indicated that they used one textbook for all algebra classes and one for all geometry classes. A teacher in school district “C” explained, “We used to have three textbooks for algebra, three textbooks for geometry. Now one size fits all.” A few more teachers further explained that many students struggled with comprehending the text as written due to the book’s reading level. That is, reading comprehension was a real issue for some students regarding the “one size fits all” textbook. In addition, high ability students were using books that were considered too easy for them.

#### 4.2.2.2 Suburban Schools

Instructional changes had occurred as a result of the MCAS grade 10 math test and the MCAS assessment system as a whole in all three suburban high schools participating in this study. Analysis of each of the three focus group sessions gleaned similar changes in math instruction at these suburban high schools when compared to

the urban results. Generally, every teacher had included more open-ended questions in his/her classroom instruction. Many teachers indicated that open-ended questions had always played a role in their teaching, yet they now consciously included a wide range of open-ended questions in everything they do. That is, homework assignments, quizzes, projects, and regular tests now contained many open-ended questions. This action had seemed to help students become more comfortable with, and capable of, successfully answering these types of questions.

Teachers in school district “D” agreed that you were always going to get an improvement in student performance when you spend more time on a concept or technique (e.g., open-ended questions). One teacher in that district indicated that the use of “Buckle Down”, a MCAS booklet, was very helpful. Another teacher added, “And we counted it as a chapter. So it was part of the curriculum. This wasn’t just for their enjoyment. This was also a grade for them as well. We gave quizzes and worksheets on it.” These teachers arrived at a consensus that including MCAS preparation as part of the course grade was a win-win situation for the students and school district. Additionally, teachers in this district further explained that many projects and classroom activities had been dropped. If it was not going to contribute to the MCAS scores then it was eliminated. In addition, there had been a concerted effort to coordinate their grading system at the department level.

Teachers from school district “E” expressed their concern about Hispanic students understanding concepts in class with a lot of help, but not being able to successfully complete MCAS test items because they didn’t understand the words. A teacher stated, “I mean they may know what to do somewhere in the back of their

minds, but reading it, they don't know that that's what they're supposed to do on that question." Another teacher added

In class they can ask you a question. You can help them key into the concept. And once you give them a hint they're off and running. But in the MCAS [grade 10 math test]...they get to a paragraph and sometimes the paragraph is enough to turn them off to not even try and these kids know what they're doing.

#### 4.2.3 Educators' Opinions of a Standards-Based Curriculum

##### 4.2.3.1 Urban Schools

Participants in one of the three urban high schools (school district "B") indicated they were comfortable with the implementation of a standards-based math curriculum as outlined in the Massachusetts' K-12 curriculum frameworks. This group reached consensus on this important point. The teachers who participated in this focus group session mentioned several benefits to having a statewide standards-based curriculum. Many of these teachers thought a uniform curriculum provided transient students with a better shot at succeeding throughout the commonwealth. While a few teachers stated that the mid-course changes made to the mathematics frameworks by the Massachusetts' Department of Education created some discussion in their department, they for the most part believed these changes were manageable.

Teachers in the remaining two urban school districts had two very different views on a standards-based curriculum. In school district "A", teachers were united in their concern that a standards-based curriculum meant all teachers were to be on the "same page" on any given school day. The states of New York and Connecticut were cited as examples of the regimentation that is involved with a standards-based curriculum. One teacher in school district "A" stated, "Either you're going to make up

your mind you're going to just teach the subject or you're going to teach the kids." This seemed to sum up these participants' sentiment regarding a standards-based curriculum. Regarding school district "C", when the teachers were asked to discuss their opinion of a standards-based curriculum, they immediately went into a heated discussion about the elimination of tracking in their middle schools and high school. One teacher from school district "C" further explained, "It would only work if you had standardized leveling in all the schools. Like all the top-level kids in Springfield, Boston, Waltham, or New Bedford. They'd all be basically around the same academic levels."

#### 4.2.3.2 Suburban Schools

Mathematics teachers in all three suburban high schools indicated they were in support of a standards-based curriculum; however, they outlined two problems inherent in implementing a standards-base curriculum. First, a standards-based math curriculum must be uniformly implemented in the K-12 grades, which had not fully occurred in any of the school systems according to the participating high school math teachers. A teacher in school district "E" explained

I personally think it's still too early to tell because I don't think the younger grades have fully implemented the standards- [until] a group of students makes it through twelve years under the Education Reform Act, I don't think you're going to see it [the impact of a standards-based curriculum]. Currently, everybody is still trying to pick up the pieces.

Second, professional development must be offered specific to the implementation of the K-12 math standards. That is, participants in all 3 suburban high schools indicated that professional development workshops were often chosen based on "more bang for your buck!", which translated into a workshop presenter outlining how

to integrate standards-based instruction across subject areas. This presented a problem in that almost all of the presenters were not knowledgeable or even conversant in mathematics. Another teacher in school district “E” stated, “We are working on standards-based instruction now, but it takes a long time. It doesn’t happen overnight.” A math teacher in school district “D” outlined his concern regarding the standards-based curriculum limiting his ability to teach students to analyze a math problem in three or four different ways. He mentioned that he used to teach in a more in-depth manner, now “we teach to the test.”

One other issue came up in one suburban school district. In school district “E”, teachers indicated they still needed more money in their budget to fully implement the math standards in their high school. One teacher stated, “We should have more computers available....There were a lot more things we should have....He [their math department head] still works within his budget...he doesn’t always have enough in the budget to buy everything needed.”

#### 4.2.4 Morale and Motivation of Educators

##### 4.2.4.1 Urban Schools

When the urban mathematics teachers in each participating high school were asked how the test and the assessment system affected their morale and motivation, they were all united in stating they felt defeated. The majority of the high school teachers agreed they were eagerly waiting for retirement. In addition, many of the teachers indicated they were just plain sick of the entire initiative. One teacher in urban school district “A” stated, “I’m just sick of it. I’m sorry- that’s a gut reaction.” Another

teacher in school district “C” asked, “How many of us are retiring as soon as possible? [The majority of teachers raised their hands.] There you go.” Another teacher in school district “B” explained

In addition to being blamed for student failure on the MCAS test, we were also being hit with the threat of teacher testing. While the majority of us feel confident about passing such a test, this test will be used against us just as the MCAS test and its results have been used to blame us.

#### 4.2.4.2 Suburban Schools

While the morale and motivation among the suburban math teachers was low, the words used to describe their feelings were not as strong or negative as the urban teachers. A consensus among each group of participants indicated that morale and motivation was at an all time low; however, they did not direct their negativity at the reform initiative. For instance, one teacher from school district “F” stated “As for testing- from what I understand the math portion of the MCAS is corrected by out-of-state teachers and that annoys me because they use in-state teachers to correct portions of other MCAS tests.” A teacher in school district “E” outlined his discouragement as follows

It’s discouraging when you start reading in the paper about how incapable math teachers were...Right now we’re getting a lot of kids who were not prepared at all. A lot of it comes from home... You’ve got kids coming into school who don’t know letters, colors, or even the alphabet..

#### 4.2.5 Improvement in Student Learning

Each group of teachers participating in a focus group was asked if the MCAS assessment system, specifically the MCAS grade 10 mathematics test, had brought about improvement in student learning. Depending on the responses of the teachers, a

follow-up question was asked: Do the math grades received by students in your district mirror their scores on the MCAS grade 10 math test? If teachers' responses to the follow-up question were negative, they were asked to discuss the reasons for the difference.

Generally, the majority of the participating teachers indicated that student learning had improved; however, each group then proceeded to qualify what that meant to them. The following two sections shed some light on their discussion regarding improved student learning in urban and suburban settings, respectively.

#### 4.2.5.1 Urban Schools

The majority of teachers in the three urban high schools indicated that some improvement in student learning had occurred as a result of the MCAS mathematics assessment initiative. One teacher from school district "B" summed it up for her peers, "In some cases, sure it has made a few students try a little harder. Still many others do not see the usefulness of such a test." Another teacher in the same district stated, "The graduation requirement does put some pressure on the students to succeed, but it can't change their view of the importance or the role education plays in their world, or their future." Still another teacher from that district added

A few international exchange students commented to me about how much teachers do for their students in the U.S.A. compared to their own countries. In each respective country, more responsibility is placed on the student's shoulders in the area of education/learning.

This series of comments created quite a bit of discussion related to American students not taking their education seriously and not taking responsibility for their own success.

The discussion quickly turned to the role parents/families play in education. A



consensus was reached that students need support at home in order to succeed in education.

In school district “C,” teachers spoke about getting what was desired out of the education reform initiative. That is, one teacher stated

In almost a crazy sense, they’ve actually received what they wanted. We’ve taught these kids how to study for a test. Not to study to learn, not to study to know, but for instance, open-ended questions, if it does work, now they will be good at open-ended questions, but life isn’t always an open-ended question. Sometimes you have to have accuracy. You’re not going to get partial credit if you get near the right answer. Your boss is going to want the right answer....You know what, the world out there isn’t really a multiple choice test. It’s knowing what you know and being able to use what you know to help someone else do something. That’s what we should be doing. Before, we gave them these tools. We gave them skills, the tools necessary to do that and along the way we actually taught them creative learning so they could actually think up how they’re going to use their tools.

Another teacher in the same district added

We used to have consumer mathematics courses, which was a type of “life skills” math, and that’s what some of these lower ability kids need. They’re never going to head to a job where they’re going to use algebra, but they should know how to balance a checkbook and do their income taxes.

Many teachers in school district “A” had similar thoughts and those sentiments can be summarized in this statement, “They really, truly, just can’t do it. I mean they literally, especially with the math, can’t do it!” Another teacher in school district “A” further explained

The whole basis of the MCAS system is that in order to be a productive member of society you have to pass MCAS math and English because other than that you’re worthless. I think our system in the United States is built on strengths and weaknesses. I hire someone to do work on my house because I’m not good at that. Some people were good at math and some weren’t and to say that the kids were not worthy of a diploma is just wrong.

Two of the three urban school districts discussed their students' math grades mirroring MCAS scores. Teachers in school districts "A" and "B" indicated that there was a match between the two; however, teachers in district "B" added there were a few surprises in the match up of average students or students in the middle. For example, one teacher stated, "Wow—that student did not complete one homework assignment all year, but he managed to pass the MCAS!" Another teacher in the same district agreed by adding, "I had a few students not pass who were conscientious students. I can't help but wonder about their test taking skills or anxiety. Perhaps they were not able to indicate what they knew on that test."

One final issue emanating from the "improved student learning" discussion was brought up in two of the three focus group sessions. The issue revolved around "retention." Teachers in two of the three urban school districts indicated that lack of retention was a real problem in that students were routinely promoted year after year regardless of academic success. For instance, a school teacher in district "A" stated

I have a student in school who is a junior this year who was in tears when he found out he flunked the MCAS math test for the second time and he told me that I didn't understand because he hadn't passed a math class since fifth grade....How did he get so far? It's the system that has failed him.

Another teacher in the same district added, "I have kids sitting in my classes doing nothing all day for the second year in a row. Something has to change!" Generally, teachers in these school districts indicated that social promotion in K-8 grades was a problem and more stringent promotion requirements should be considered at the high school level.

#### 4.2.5.2 Suburban Schools

The suburban teachers had a more positive take on the degree of improvement in student learning caused by the MCAS system, specifically the MCAS grade 10 math test. Each of the three groups reached consensus in stating “yes” – student learning had improved but each group qualified their positive responses. For instance, a math teacher in school district “D” explained

I think it had to...I think it almost had to improve student learning because we changed our teaching styles...We have tried to coordinate our grading system. I know I have taken a lot of stuff out of my teaching style. I used to do a lot more than I do now....count a project as a test grade or something like that. We don't have the luxury of doing that anymore. We have to make sure that they know the material. So the focus is different which means the kids can't receive B's for a nice fancy project...now they really have to do well on quizzes and tests. In other words, if it doesn't contribute to MCAS scores we basically don't include it.

One teacher in that same district added, “Then maybe we've changed the focus rather than improved student learning.” Another teacher countered

I think it has elevated the lower level students because they're now doing the algebra. They're doing the geometry. I don't think they're happy about it. I don't think that they ultimately really understand it either because most of the lower level students were not abstract thinkers. Nor do I think that they see it tying into their future.

Teachers in school district “E” discussed the fact that the general math courses were no longer available due to the MCAS initiative, which created a catch-22 effect on student learning. Teachers in school district “F” discussed the same issue, but agreed that a lot of the students coming through their classes were not finding Algebra 1 and Geometry to be an attainable goal. One teacher yelled, “It's a nightmare!” Another added

The whole philosophy is everyone must take Algebra 1 and Geometry now. Does everyone need Algebra 1 and Geometry to be a successful

person after high school? Not everyone is going on to college after high school; therefore, is it necessary for everyone to know how to do Algebra 1 and Geometry?...Is that necessary to be a productive human being in society?

Finally, a teacher in school district “E” wondered out loud, “I’ll be interested to see over time if more students drop out because there’s not a whole lot of hope out there for some of these kids to stay in school.”

Regarding math grades mirroring MCAS scores, one of the three suburban groups addressed that follow-up question. In school district “E”, teachers indicated that their high- and low-ability students matched up well. The middle-ability or average students were not as clear-cut. That is, teachers mentioned that perhaps some of their students were not be able to show what they knew on the test and/or they suffered from test anxiety, which caused the difference in the match up.

#### 4.2.6 Suggested Changes to Assessment System

##### 4.2.6.1 Urban Schools

Teachers in the urban high schools who participated in a focus group session indicated that the MCAS grade 10 math test should be changed so that it assesses a range of skills that everyone is going to need in order to live. That is, make the questions more relevant to real life. One teacher in school district “A” explained, “For example if a kid makes \$300 a week and his rent is \$xxx, does he have enough money to live on?” Another teacher in school district “C” stated

Do I want accountability? Absolutely. Based on the social climate in the United States and watching how kids were reacting to it, I want something that makes them accountable. However, to make a judgment that everyone must know algebra before exiting high school, I think is wrong. The assessment itself has to change and it should represent a core of skills

that everyone is going to need. I sat one evening taking one version of the MCAS math exam and pondered how many questions my very intelligent, capable wife could answer on that test. I questioned whether or not one needs algebra to exit high school. My answer to that is no, you do not.

Many teachers in school district “C” indicated that the MCAS grade 10 math test should be dropped as a graduation requirement, but it should continue to be used as an accountability tool. Still a few more in that same district added that it should be moved into a higher grade. One teacher in school district “A” mentioned

There were a few people that we work with that probably shouldn't be working anymore, but in all my twenty-plus years of teaching I have never seen anyone fired. Perhaps instead of using the MCAS system maybe we should be moving towards eliminating people that were not doing their job. I think we would have been a lot better off if that was the focus.

Many heads nodded after this statement was made, but consensus was not reached on that specific issue.

#### 4.2.6.2 Suburban Schools

Teachers in the suburban high schools considered the algebra and geometry knowledge needed to pass the MCAS grade 10 math test to be unfair for some students. According to these teachers, the students who had not passed and/or taken Algebra 1 and Geometry, found this test impossible to pass. In addition, students in vocational schools and special education programs were considered to be at a disadvantage. Labeling high school students who had not passed the MCAS test as “failures” was also an issue for many suburban math teachers. Many teachers mentioned there was still confusion over what kind of test it was. Many educators and non-educators compared the MCAS tests to basic skills tests and that was obviously a problem.

Finally, one teacher from school district “E” indicated a friend of his served on the committee responsible for creating the grade 10 math frameworks and he told him that he was the only representative from Western Massachusetts. In addition, he stated many of those serving on that board were from very affluent communities. He found himself consistently stating that many of the selected objectives/questions were unfair to the average student living in his geographic area. Students in higher socio-economic communities typically have more self-motivation and more parental interest/control, which translated into a greater probability of successfully passing the test. He found himself consistently asking the frameworks committee members about the remaining student population. He added that many of the objectives/questions were not reasonable for many tenth grade students living in the western part of the state. These specific issues concerned the majority of teachers participating in that focus group.

#### 4.3 Interview Results

The one-on-one interview results have been summarized using the content analyses procedure outlined in Chapter 3 (see Section 3.6.3). After reviewing the content of the first interview transcript whereby all key phrases and comments were coded, subsequent transcripts were then compared to the first and themes were identified. Major themes were developed by grouping similar comments made by different respondents into the same category.

Twenty-three education professionals from ten different high schools participated in the interview component of this research study. Each of the ten high schools was assigned a letter of the alphabet (e.g., School district “A”, School District

“B”) to ensure anonymity. The letters assigned to the six high schools in the focus group analyses remained and the 4 additional high schools received assigned letters in the same manner.

Three major themes emerged from the content analyses of the 23 transcripts: changes in curriculum and instruction, opinions on standards-based curriculum, and the effect of MCAS on students, teachers, and non-teachers. These three major themes can be clearly connected to the questions outlined in the interview protocol in Chapter 3 (see Appendix G). The themes have been discussed based on the responses of individuals holding the same positions. That is, the types of position held by the interview participants: (1) mathematics department chairs/heads, (2) principals and assistant principals, and (3) superintendents and assistant superintendents. Eleven of the 23 participants (48%) were principals or assistant principals, seven were superintendents or assistant superintendents (30%), and five were mathematics department heads (22%). The three themes are discussed in terms of position held by the participants in the following three sub-sections.

#### 4.3.1 Changes in Curriculum and Instruction

##### 4.3.1.1 Mathematics Department Chairs/Heads

All five (22%) mathematics department heads indicated that levels within each course had been collapsed with the number of levels falling between three and four among the participating high schools. The department head from school district “I” stated they had four curriculum levels with most of the students in the bottom two

levels. At school district “J”, the department head indicated they had three levels: honor, regular, and lower level.

The collapsing of course levels seemed to be a huge issue at the remaining three high schools (“A, E, and G”) because the change increased the range of abilities within each class making it more difficult to teach the mathematics concepts. This problem was exacerbated by the deletion of a whole sequence of course offerings in the area of general/technical mathematics. In addition, preparation of incoming ninth grade students remained an issue for these three high schools. Two of the three high schools were in urban areas with the remaining one in a suburban setting. It is important to note that the remaining high school, which was in suburban school district “E”, was grappling with a changing student population in that the community was looking more and more like an urban city every day.

From an instructional standpoint, all of the mathematics department heads clearly stated that their teachers had increased the use of open-ended questions in all aspects of their teaching. That is, classroom and homework assignments, quizzes, and tests routinely included open-ended questions. In addition, the importance of successfully passing the MCAS grade 10 mathematics test was stressed and released test questions were used by all of their teachers. Mathematics department heads from school districts “I” and “H” indicated that their course content was still being tweaked and revised to better align with the state’s MCAS math frameworks. The remaining department heads (“A, E, and G”) stated that more attention needed to be given to framework and textbook alignment.



Remediation help for those failing or in danger of failing the MCAS grade 10 mathematics test had been implemented in all five high schools. This help had been offered in many forms: after school tutorials, additional in-school instructional blocks, independent computer tutorials, smaller class sizes, and team teaching where the additional teacher was often a special education instructor. The math department head from school district "I" mentioned that their mathematics laboratory, which was created before the Massachusetts Education Reform Act of 1993 (MERA), had seen an increase in usage. This laboratory was a drop-in center that was staffed by various math teachers throughout the day to field student questions about math instruction and/or homework.

#### 4.3.1.2 Principals and Assistant Principals

All 11 (48%) of the principals/assistant principals agreed that their building's mathematics curriculum and instruction had changed as a result of the MERA of 1993. While the bulk of the changes occurred in instruction, the degree of curriculum change and/or alignment to the Massachusetts mathematics frameworks varied from school to school.

Three principals indicated that their math curriculum had been changed but more work needed to occur. That is, they felt there was room for improvement. The principal from school district "F" indicated that teacher turnover had negatively impacted his math department's ability to offer consistent and up-to-date instruction. In addition, the same principal mentioned that his school district went without a K-12 curriculum specialist for many years. Two principals from school district "C" and "E" indicated that the new frameworks had been adopted and they had subsequently

changed their math curriculum but more time needed to be given to the actual implementation of the adopted frameworks. One area needing attention included the need to review whether or not their textbooks were aligned with the frameworks. In addition, three principals/assistant principals from “A, E, and G” school districts clearly stated that the deletion of all general/technical math courses had a mixed effect on the individual math teachers and their instruction. That is, requiring all students to take Algebra 1 and Geometry impacted how their teachers taught and at what pace they covered the material.

#### 4.3.1.3 Superintendents and Assistant Superintendents

Four of the seven superintendents/assistant superintendents (57%) stated their district’s (C, E, F, & G) high school math curriculum and instruction had changed but more time and effort needed to be expended to fully implement the adopted frameworks with the remaining three administrators having varied opinions on this issue. The superintendent from school district “D” indicated his high school math curriculum had been rearranged so instructional sections met the needs of the students taking the MCAS grade 10 math test; however, he added that no substantial changes in course offerings had occurred. The same superintendent indicated that the high school’s math curriculum was aligned with the frameworks and had recently gone through a second curriculum review as outlined in the district’s policy and procedures manual. In addition, the assistant superintendent in that same district echoed his sentiment, but he further explained that instructional practices were constantly being tweaked based on past student performance on the MCAS grade 10 math test.

The assistant superintendent from school district “I” stated their curriculum was not necessarily influenced by the MCAS assessment system. That is, MCAS was viewed from a systemic perspective in that they wanted to know if they were reaching their students; however, their math curriculum continued to be reviewed using a previously established curriculum revision cycle. The assistant superintendent from school district “E” stated “specialization is killing us.” She further explained that her district needed a more integrated approach to learning especially in mathematics. That is, conceptual learning rather than repetitive or rote learning was needed. She also stated that many teacher preparation programs were not in line with this notion. Therefore, many new teachers were not properly prepared to teach the math frameworks.

#### 4.3.2 Opinions on Standards-Based Curriculum

##### 4.3.2.1 Mathematics Department Chairs/Heads

The five department heads supported a standards-based curriculum. The mathematics department head from school district “I” indicated a few particular skills/concepts were pushed into the frameworks by college professors and perhaps those should be removed, but otherwise the MCAS K-12 Mathematics Framework was “O.K.” The department head in school district “G” believed in a standards-based curriculum but thinks “the state of Massachusetts is clueless as to what it should look like.” Another department head from school district “H” strongly supported the frameworks, but was against “the club” used against students, which came in the form of the graduation requirement. She further added that data driven decision-making was

a necessity in education; however, the graduation requirement placed an unfair amount of pressure on students. She stated some kids meet all of her district's requirements, but just couldn't pass the test. Obviously, these students were taking retake after retake and that process labeled them as "failures," which concerned her to a great extent.

#### 4.3.2.2 Principals and Assistant Principals

All of the principals/assistant principals believed in the merits of a standards-based curriculum, while several qualified their support with a comment on the actual implementation of the Massachusetts K-12 mathematics frameworks in their school buildings. The principal from school district "F" stated that once it was clear what the standards were going to be the math teachers in his building went to work on them; however, he believed the implementation of the MCAS math tests should have followed the children through the school system. In other words, the district's curriculum alignment with the K-12 math frameworks had not happened over night. Budgets needed to be adjusted and professional development needed to occur. He further explained the MCAS tests should have been implemented in the earlier years and should have slowly moved up through the grades with those students who were learning the revised K-12 math curriculum.

The principal from school district "H" indicated that the implementation of the math standards had created an atmosphere of improved K-12 communication. More specifically, math teachers were now discussing the standards, their district's curriculum, and how they might tweak their instructional practices to better align classroom activities with the adopted frameworks. The principal from school district

“I” explained the standards should definitely remain because they made teachers focus on the under-served student population in the area of mathematics.

#### 4.3.2.3 Superintendents and Assistant Superintendents

All seven of the superintendents/assistant superintendents believed a standards-based curriculum was very important to education. They indicated that mathematics teachers needed to be familiar with what students should know and be able to do. This issue was viewed as a very important component in the Massachusetts’ education reform initiative. A few of the administrators elaborated by outlining the importance of defining the standards first, and then moving forward. Common or uniform standards were viewed as a common goal or objective. A standards-based curriculum was viewed as a logical business-minded decision according to the superintendent from school district “D.” He also believed that it was a better idea than local control, which created a lack of curriculum uniformity across the state.

#### 4.3.3 Effect of MCAS on Students, Teachers, and Non-Teachers

##### 4.3.3.1 Mathematics Department Chairs/Heads

The department heads clearly believed in the merits of accountability for teachers, non-teachers, and students. What that accountability looked like was perceived differently by each participant. In addition, the accountability tool was viewed to have a different effect on each group of stakeholders.

Regarding the effect on students, the department chairs believed students should be held accountable because today’s society often protected students from taking

responsibility for themselves. The department head from school district “I” stated, “It has helped to underscore that this time in students’ lives when they need to learn. Adults expect them to learn rather than work full-time.” The department head from school district “C” stated, “This school has improved student learning, and it’s done so on the backs of teachers. Teaching in an urban district is tough.” In urban school district “A”, the department head explained, “If this test reflects what society needs in terms of students’ math knowledge, then students were all set. If not, students will suffer. It appears to me that they were not examining broader learning experiences, such as: career exploration.” The same math chair indicated that field trips were not approved unless they were directly linked to the math standards. This brought up another issue, that is, the arts and sports often sparked an interest in the minds of many urban high school students and that connection was what kept those children in school. In other words, this particular math chair believed that everything else has been deemphasized in order to focus on English and mathematics due to the reform initiative. He believed broader experiences had been taken away, which was particularly dangerous in an urban setting.

Regarding teachers and non-teachers, professional development and morale and motivation had been affected by the MCAS assessment system. In addition, one district had experienced an increased workload for high school math teachers only.

Professional development had played a key role in helping teachers implement the math standards adopted by their school districts. It was clear that professional development was an on-going process containing positive and negative elements. The positives included K-12 collaboration among all math teachers, visits to other schools, common

planning time, and general intellectual growth for math teachers as a result of these learning experiences. The negatives included initial training completed with subsequent follow-up cancelled due to budget cuts, lack of interest on teachers' part after sufficient professional development points (PDP's) had been acquired for recertification, and professional development trainer(s) or presenter(s) not conversant or knowledgeable in mathematics.

According to 3 of the 5 math department heads, morale and motivation of math teachers in their high schools were quite low. An increase in cynicism, a dislike for the Massachusetts Department of Education, and general frustration at the increasing demands placed on math teachers had contributed to low morale and motivation. The remaining two department chairs indicated there were pockets of low morale and motivation. The math department head from school district "C" stated that two-thirds of his teachers were excited and interested and the remaining one-third were "dinosaurs" and more difficult to change. The department chair from school district "I" believed the tenth grade teachers felt most of the pressure. Consequently, they were overly sensitive about using all of their available instructional time without any outside interruptions (e.g., field trips).

An increased workload had occurred in district "A's" high school. The increase was exclusively in the mathematics department, which had caused negative repercussions throughout the department reverberating throughout the rest of the high school. The math department chair in this district explained that the additional instructional blocks were added so MCAS remediation classes could be added to their students' schedule.

#### 4.3.3.2 Principals and Assistant Principals

Four of the 11 principals/assistant principals (36%) indicated the morale and motivation of the math teachers in their building (“D, E, F, and H”) was good or high. One principal from school district “E” stated, “They were handling it well.” Two assistant principals from school district “B” and one principal from district “J” (27%) were not comfortable answering the question pertaining to the morale and motivation of the math teachers in their building. Each of the three administrators did not believe they could gauge their math teachers’ morale and motivation. Still two other principals from districts “A” and “G” clearly stated the morale and motivation of their math teachers was “lousy.” The principal from school district “C” indicated that morale was both up and down among teachers in the math department. Finally, all of the administrators indicated their own personal morale and motivation was high when considering the implementation of the MCAS assessment system.

#### 4.3.3.3 Superintendents and Assistant Superintendents

The superintendents and assistant superintendents who participated in an interview indicated in the early years of the MERA teachers’ morale and motivation was low. As the education reform initiative continued, the movement brought out mixed emotions among math teachers. Many wondered why the state had meddled in local school districts’ business, while some veteran math teachers went into denial believing that “this too shall pass.” Some chose early retirement with the newer teachers embracing the MCAS initiative. More recently, a lot of energy had been placed on analyzing student test results and making appropriate changes to instruction



based on weaknesses indicated from the analyses. The superintendents and assistant superintendents stated many teachers were still concerned about the amount of anxiety and pressure placed on students as a result of the graduation requirement.

As a group, the superintendents and assistant superintendents indicated their morale and motivation had not changed, but they were very sure the majority of students' motivation had increased after the graduation requirement became a reality. Five of the seven administrators (71%) believed student learning had improved with the remaining two stating that it was too soon to tell.

## CHAPTER 5

### CONCLUSIONS

The final chapter of this study includes five sections: summary of results, delimitations, discussion of findings and implications, future research, and conclusion. The summary section provides a synopsis of the results of the study including the intended and unintended consequences of the grade 10 MCAS. The delimitations section discusses issues concerning the generalizability of the findings to the state population. The third section provides a discussion of findings and implications based on the results of this validity study. The fourth section presents suggestions for future research with the conclusion offering closing thoughts.

#### 5.1 Summary of Results

##### 5.1.1 Summary of Survey Results

The strong survey participant response rate (56.2%) added credence to the findings of this study. It is important to note the majority of educators taught Algebra 1, Algebra 2, and/or Geometry with 30% teaching 5.5 years or less and another 33% teaching 26 or more years. About half of the teachers had earned a master's degree and approximately half of the administrators had attained a doctoral degree. Generally, the participants had positive perceptions of the impact the education reform initiative had on their math curriculum. Approximately half to two-thirds of the respondents agreed to strongly agreed that course offerings and course sequencing had changed in their high school math departments. In addition, about 75% of the respondents strongly

believed that MCAS test results were used to identify those students needing math remediation. However, three-fourths of the teachers did not consider averaging student MCAS scores as a good indicator of their high school's overall educational success in teaching math. About two-thirds of the teachers indicated that grade 10 MCAS test preparation meant other valuable topics were NOT being taught and the test was unfair to many students for many different reasons.

Regarding the ratings of the nine influence statements, all except one were viewed as having more influence today and/or in the future than in the past suggesting many aspects of the reform movement had been and continue to be implemented. Four specific statements were viewed as growing in influence in the future, which indicated the reform initiative would have a continuing impact on high school math curriculum and instruction.

The three negative themes emanating from the two open-ended questions suggested that the actual MCAS grade 10 math test was viewed as an unfair assessment for a wide range of reasons. The two positive themes indicated that the survey participants welcomed increased student accountability and the implementation of the statewide math standards.

### 5.1.2 Summary of Focus Group Results

A wide range of significant changes had occurred in mathematics in the participating high schools as a result of the MCAS. General mathematics courses had been deleted from the program of studies at five of the six high schools and more students had been enrolled in Algebra 1 and Geometry with a strong remediation effort

for those failing the MCAS grade 10 math test. It is important to note the remaining high school did not need to delete any general math courses because the school itself was created after the education reform initiative had been implemented; therefore, its program of studies was built on the tenets of the initiative, which advocated algebra and geometry courses rather than general or technical math courses. In addition, participants in five of the six high schools indicated individual levels within each course had been reduced producing more heterogeneous math classes. An increase in the usage of open-ended questions in homework assignments, quizzes, classroom instruction, and tests had also occurred in the six participating high schools.

The high school math curriculum had been aligned with the Massachusetts' K-12 Frameworks in all of the participating high schools with two of those needing more time and attention to fully implement the standards as outlined. Teachers in four of the six focus groups supported a standards-based math curriculum. The morale and motivation of teachers was low with retirement viewed as an enticing solution to the pressures of the reform initiative. Student learning had improved according to the participants; however, the quality of improvement was debated. Finally, numerous changes to the MCAS were discussed with one general issue emerging. That is, a request to review the body of math knowledge needed to pass the MCAS grade 10 math test, specifically the algebra and geometry concepts.

### 5.1.3 Summary of Interview Results

The results of the 23 interviews clearly indicated that the MCAS had impacted students, teachers, and non-teachers in the ten different high schools participating in the

interview component of this study. Increased accountability for students, teachers, and non-teachers was supported. The impact of the increased accountability had positive and negative affects on the participants. Morale and motivation of teachers represented a mixed bag. That is, some had been positively affected by the reform initiative, while many veteran teachers had seen their morale and motivation drop, which had led many of them to consider early retirement. The administrators indicated their own personal morale and motivation was good. Student motivation had been viewed as increasing due to the graduation requirement attached to the MCAS grade 10 math test.

According to the mathematics department heads, principals and assistant principals, and superintendents and assistant superintendents, open-ended questions were now routinely and regularly used in math classes. These questions permeated throughout classroom instruction, homework assignments, quizzes, and tests. Consequently, students were definitely more competent and capable of answering this type of question.

The education reform initiative had also resulted in fewer academic levels for each math course taught, while it was indicated that course sequencing had not changed. Regarding course offerings, all general and technical math courses had been deleted. In addition, a wide range of remediation help had been implemented over the last few years in each of the participating high schools. Some remediation programs had been discontinued due to their lack of success or low student attendance, while programs that occurred during the school day had been perceived to be more effective.

A standards-based mathematics curriculum was fully supported by all participants; however, the superintendents and assistant superintendents indicated that

more time was needed to fully implement all of the components of the frameworks or standards as outlined in the education reform initiative. On-going and effective professional development was viewed as an important component of the successful implementation of a standards-based mathematics curriculum.

#### 5.1.4 Study Summary

The survey response rate (56.2%) and parallel results from each of the three components of the study reinforce the general findings discussed in this section. The results from each component of the study were very similar with only minor differences in the degree of participants' opinions. Approximately half to two-thirds of the survey respondents agreed course offerings and sequencing had changed in their high schools. In the focus group and interview sessions, the participants referred to changes in course levels and the deletion of the general track, but the survey was more limited in that area so perceived change showed up in the area of course offerings and sequencing. Specific changes in the number of academic levels within each course, including the deletion of the general track, points out an intended consequence of the MCAS as previously highlighted in Chapter 1. This issue also highlights an unintended consequence of the MCAS as it relates to high school mathematics in that teachers were concerned that the deletion of the general track AND the collapsing of levels within each math course may be adversely affecting low- and high-level students.

Participants consistently agreed MCAS results were successfully used to identify students needing math remediation and that remediation was offered in many forms. Some remediation programs had been more successful than others with in-

school help viewed as the most effective. Identifying students in need and offering assistance via math remediation is also an intended consequence of the reform initiative.

The MCAS grade 10 math test and the assessment system as a whole had greatly influenced all of the participants. That is, the Massachusetts' education reform initiative had influenced and continues to influence the participating high school's math curriculum and instruction. The most prominent aspect of change involved the increased use of open-ended questions by teachers. In addition, the alignment of high school math curriculum to the state's frameworks had occurred in varying degrees within the 11 participating high schools. These findings illuminated the intended consequence of curriculum reform outlined in Section 1.2.

Participants in all three components of the study indicated that the MCAS grade 10 math test was an unfair assessment for many reasons. Most of these reasons revolved around student needs. Responses to the questions inquiring about the morale and motivation of the participants were mixed. The "fairness" and "morale and motivation" issues represented unintended consequences of the MCAS. This concern will be further discussed in Section 5.3, Discussions of Findings and Implications.

Throughout the interview and focus group sessions, participants consistently stated increased accountability and uniform statewide math standards were a positive addition to high school math education in Massachusetts. These issues were also reflected in the content analyses of the two open-ended question contained in the survey. As previously discussed, the two major positive themes were the implementation of statewide mathematics standards and increased student accountability. That is, 28% of those responding believed the implementation of the

statewide K-12 Mathematics Curriculum Frameworks, which was also defined as “math standards,” improved their district’s math curriculum, and indirectly, improved instruction. Improvement was described as increased communication among math educators, professional development, regular curriculum reviews, and textbook alignment with K-12 math frameworks, which should be viewed as a positive intended consequence. Twenty percent of the respondents viewed increased student accountability as a positive consequence of the education reform initiative. Many students were exerting more effort and motivation, and it was hoped that overall student performance would increase as a result. Respondents indicated that students now knew when they needed remediation. The increased student accountability also created an increased awareness of the consequences of the education reform initiative among parents and teachers.

## 5.2 Delimitations

Generalizability of the findings of most research studies involving a representative sample of a population must always be examined. While care was taken in the selection of the representative sample in this study, consideration of the success of that process must be pondered. The cluster analysis and subsequent MANOVA procedure, which cross-validated the 12 clusters of school districts, strengthened the generalizability of the findings. The fact that 11 of the 12 selected high schools participated in one or all three components of the study further strengthened the study’s generalizability. The educators from the participating high schools who completed one or more component of this study reflected a cross-section of school districts across



Massachusetts, which translated into a sample that is representative of the population of high school mathematics teachers and administrators in the state of Massachusetts.

### 5.3 Discussion of Findings and Implications

The findings summarized in this chapter define and illuminate implications of the reform initiative in the state of Massachusetts. States under-going education reform efforts may find these results interesting and helpful in planning similar validity studies. As reflected in this dissertation, it is very common to have the findings in an educational research study answer as many questions as it creates. That is, the results indicated many positive changes in mathematics education had occurred throughout Massachusetts, but more research needs to be conducted in order to clarify the concerns and perceptions of the participants uncovered in this study.

It is also important to note that every educational reform initiative is held accountable by local, state, and/or federal statutes. Education is a fluid-like process. While teachers work “in the trenches”, administrators must clearly outline a school’s mission and maintain a leadership role in communicating a common purpose within a school community. The mixed levels of participants’ morale and motivation found in this study reflect the varied perceptions of the challenges presented every day in the classroom, such as: students with low motivation or one type of algebra textbook for all math students. Are teachers’ perceptions of MCAS entangled with the day-to-day difficulties of the profession?

A teacher’s perception of the fairness of an assessment is very different than that of an education measurement expert. Frequent and productive communication at the

school, district, state, and federal levels must occur in order to alleviate concerns and correct any element of “unfairness” in relation to a high-stakes test. It is relatively easy to conduct item bias reviews, a psychometrically-driven fairness issue; however, a different kind of fairness issue, highlighted by teachers in this study, involved a low-level student who is not successfully passing the MCAS test. This student, based on local, state, and/or federal statutes, must receive comprehensive support. While that student is given five opportunities to pass this graduation requirement, a safety net must embrace that student so he/she will succeed before, during, and after those test administrations. Heterogeneous classrooms that don’t address individual student needs and textbooks with inappropriate reading levels are just two issues highlighting holes in that student’s safety net. Incomplete or inappropriate professional development for mathematics educators is another example.

Another implication emanating from this study involves the participants’ desire to revisit the specific algebra and geometry concepts included in the grade 10 MCAS test. Perhaps an inclusive mid-course review process would alleviate, clarify, or correct some elements of this test that seemed to create a lot of concern and anxiety on the part of the participants.

#### 5.4 Future Research

There is a wide range of potential future research topics emanating from this study. Sound validity evidence should be gathered to further examine the education reform initiative in Massachusetts and to assist policy makers in making necessary mid-course corrections to the movement. One specific research topic is a validity study of

current instructional practices involving actual classroom observations, which would further validate educational change occurring in classrooms due to MERA. Larger replication studies should be considered to further validate the findings outlined in this study. Students, parents, and community members' perceptions and opinions should be gathered to add depth to current and future education reform research in the state of Massachusetts. It is important to note that gathering validity evidence is an on-going, always evolving process. Validity studies are expensive and time intensive but nevertheless crucial to the future of education.

## 5.5 Conclusion

All education reform initiatives involve the creation of educational goals and objectives related to the desired outcome. These goals and objectives can come in various forms (e.g., policies, rules, regulations). For the state of Massachusetts, the education reform act of 1993 was outlined in a three-step process (see section 1.2). In order to ascertain the success/failure or strength/weakness of the reform program, a regular review of these goals is deemed appropriate. The review can be done informally by policy makers and/or more formally by measurement experts but research has proven that validation should occur.

This study examined a very specific aspect of the MERA. An investigation of the consequential aspects of validity of the MCAS grade 10 mathematics test and corresponding assessment system was conducted. Three research questions were outlined: a) Have changes in curriculum occurred within high school math departments across the commonwealth as a result of the statewide grade 10 mathematics assessment

and its corresponding curriculum framework? (b) Has high school mathematics instruction changed due to the state's mathematics curriculum standards? (c) Has district high school mathematics curriculum been aligned with the state's mathematics curriculum framework? The findings of this study summarized in section 5.1 indicated that math curriculum and instructional changes in the participating high schools had occurred. Therefore, these findings suggested that aspects of the education reform goals as previously outlined have been realized with further educational change predicted.

APPENDIX A  
WRITTEN CONSENT FORM

Study of the Consequential Aspects of Validity of a  
State-Mandated Grade 10 Assessment  
**Consent for Voluntary Participation**

I volunteer to participate in this mixed method study and understand that:

1. I will complete a 2-page survey and be interviewed by Mary L. Zanetti using a focus group or one-on-one interview format each consisting of 7 questions.
2. The questions I will be answering address my views on issues related to the grade 10 state-mandated mathematics assessment used in the studied state. I understand that the primary purpose of this research is to identify the intended and unintended consequences of this assessment tool.
3. The focus group session will be tape recorded to facilitate analysis of the data, while the one-on-one interview will not be recorded.
4. My name will not be used, nor will I be identified personally in any way or at any time. I understand it will be necessary to identify participants in the dissertation by position and/or classification system using variables that created 12 school district categories (e.g., school district "L" or mathematics curriculum specialist(s) in school district "L").
5. I may withdraw from part or all of this study at any time.
6. I have the right to review material prior to the final oral exam or other publication.
7. I understand that results from this survey, focus group session, and interview will be included in Mary L. Zanetti's doctoral dissertation and may also be included in manuscripts submitted to professional journals for publication.
8. I am free to participate or not to participate without prejudice.

\_\_\_\_\_  
Participant's Signature

\_\_\_\_\_  
Date

\_\_\_\_\_  
Researcher's Signature

\_\_\_\_\_  
Date

**APPENDIX B**  
**COVER LETTER**

UNIVERSITY OF MASSACHUSETTS  
Amherst, MA 01003  
William M. Bulger, President

School of Education  
Bailey G. Jackson, Dean

February 7, 2002

Mr. XXX XXXX  
Principal of XX High School  
63 Chestnut Street  
Ludlow, MA 01056

Dear Mr. XXXX:

Thanks so much for agreeing to participate in Mary L. Zanetti's dissertation study. As you may already know, this research project is comprised of 3 components: a survey, a focus group session, and selected one-on-one interviews.

The attached survey instrument, interview protocol, and focus group protocol inquires about the intended and unintended consequences of the state-mandated grade 10 mathematics assessment currently used in this state. Each measurement tool is a component of Ms. Zanetti's dissertation, this project is one of the final requirements in the Research and Evaluation Methods Doctoral Program at the University of Massachusetts Amherst. This study is concerned specifically with investigating the positive and negative consequences of the grade 10 mathematics test administered in your high school. The results of this study will help to provide much needed documentation regarding the strengths and weaknesses of this statewide assessment system.

Ms. Zanetti is particularly interested in obtaining your responses and responses from key educators in your building because they will contribute significantly toward illuminating some of the issues related to this important assessment tool. The enclosed instruments have been tested in three individual pilot testing procedures, and revisions were made to each tool to ensure that information could be obtained using a minimum of your attention and your staff's time. The average time required to complete the survey was 15 minutes, while the focus group session will take approximately 1-2 hours, and the one-on-one interview will take about 20 minutes.

We would appreciate it if you would ask all of the high school mathematics teachers, mathematics curriculum specialists, and administrators in your building to



complete and return the enclosed survey to you prior to March 20, 2002. At that point in time, the completed surveys can be returned in the stamped, self-addressed envelope that has been enclosed. All responses will be held in the strictest confidence.

Regarding the focus group session and interviews to be conducted in your district, Ms. Zanetti will speak with you by telephone to schedule a convenient date and time to conduct those aspects of her study. Written consent forms, which outline confidentiality rules, will be made available to the focus group and interview participants during each scheduled event.

We will be pleased to send you a summary of the results if you desire. Thank you for your cooperation and participation in this important study.

Sincerely,

Bailey G. Jackson, Dean

cc: Enclosures

APPENDIX C  
LETTER OF SUPPORT

[Connecticut Valley Superintendents' Letterhead used]

February 8, 2002

Mr. XXX XXXX  
Principal of XX High School  
63 Chestnut Street  
Ludlow, MA 01056

Dear Mr. XXXX:

On behalf of the Connecticut Valley Superintendents' Roundtable (CVSR), I am writing to inform you of my support, the Executive Board's support, and the general membership's support of Mary L. Zanetti's research study. Your participation in this study will help document the important issues related to the state-mandated grade 10 mathematics test and the state's assessment system as a whole.

The education reform initiative in our state needs to be studied so the strengths and weaknesses of this program can be evaluated. Your input will be critically important to the success of this quantitative and qualitative study. Please keep in mind that all survey responses, focus group discussions, and interview responses will be completely anonymous and confidential. CVSR is always interested and eager to help a graduate student complete her doctoral dissertation. With this in mind, please take the time to meet with Ms. Zanetti and encourage the educators in your building to participate in her research study.

Given the laudable purposes of this study, we hope you can see why your input is so vitally important. Please take a few minutes to distribute the enclosed surveys to the mathematics teachers, mathematics curriculum specialists, and administrators in your building and work with Mary when she calls to schedule a focus group session and individual interviews. A postage paid return envelope has been provided. If you have any questions, please call or email Mary Zanetti at (413) 583-6143 or [mzanetti@educ.umass.edu](mailto:mzanetti@educ.umass.edu). Thank you for your cooperation.

Sincerely,

John Cullinan, Supt.  
President

APPENDIX D  
TEACHER'S SURVEY

## TEACHER'S SURVEY<sup>1</sup>

### PLEASE READ:

This survey is designed to determine how the MCAS grade 10 mathematics test and corresponding assessment system have affected your high school's math curriculum and you as a high school mathematics teacher or administrator. All information provided is confidential; please refer to the attached written consent form for more specific confidentiality information.

### INSTRUCTIONS:

This survey is divided into four sections. The first section provides us with some personal information about you. The second section is a list of statements that ask for your opinions based on an agreement scale. The third section includes questions about the impact your state's education reform initiative has had on your high school's math curriculum in the past, present, and future. The fourth and final section is comprised of a few selected response questions and two open-ended questions. Feel free to attach additional paper, if you would like to add comments. **The entire survey will take you approximately 15 minutes to complete.**

#### Section 1 – Personal Information

*Please complete the information below. Do not sign your name.*

1. Sex:  Female  Male
2. Are you a high school mathematics teacher at your school?  Yes  No
3. Please circle one of the following words that best describes your school district:  

Urban	Suburban	Rural
-------	----------	-------
4. What is the approximate number of students enrolled in your high school? \_\_\_\_\_
5. Please indicate the grade(s) you are currently teaching: \_\_\_\_\_
6. Please indicate the subjects you are currently teaching: \_\_\_\_\_

In addition, indicate the total number of years you have taught: \_\_\_\_\_

7. Are you a Massachusetts certified mathematics teacher?  Yes  No
8. **Education:** *Circle the highest degree you have completed.*  
Bachelor's degree      Masters degree      CAGS      Doctorate

---

<sup>1</sup> The author is grateful to Susan Bowles, Caryn McCrohon, Timothy O'Neil, Zhenhong Hu, and Stephen G. Sireci for the use of their two surveys in the creation of this survey.

**Section 2** – We are interested in your opinions regarding the impact the state’s Mathematics Curriculum Frameworks and assessment system has had on your high school’s math curriculum and your instructional practices. Please focus exclusively on the implementation of the curriculum rather than any other important issue that may be related (e.g., contract negotiations).

*Please circle the response that best indicates your opinion.*

Statement	Strongly agree	Agree	Undecided	Disagree	Strongly Disagree
9. I have received adequate professional development <b>at the state level</b> for implementing the Mathematics Curriculum Frameworks.	SA	A	U	D	SD
10. I have received adequate professional development <b>at the district level</b> for implementing the Mathematics Curriculum Frameworks.	SA	A	U	D	SD
11. I have changed my mathematics curriculum to better reflect the state’s Mathematics Curriculum Frameworks.	SA	A	U	D	SD
12. The mathematics department in my high school has changed course offerings due to the MCAS mathematics assessment system.	SA	A	U	D	SD
13. The mathematics department in my high school has changed course sequencing due to the MCAS mathematics assessment system.	SA	A	U	D	SD
14. My instructional practices have changed as a result of the state’s Mathematics Curriculum Frameworks.	SA	A	U	D	SD
15. I have made changes to my instructional practices due to my students’ performance on the MCAS grade 10 mathematics assessment.	SA	A	U	D	SD
16. The MCAS grade 10 mathematics assessment is a useful tool to improve the long-term education of students.	SA	A	U	D	SD
17. The MCAS grade 10 mathematics test results are being used in my high school to identify students who may need specialized instruction.	SA	A	U	D	SD
18. The school system has adequately provided remedial instruction and/or courses for those students who were unsuccessful on the MCAS grade 10 mathematics test.	SA	A	U	D	SD

19. Please circle the word that best describes your **morale** as a mathematics teacher based only on your district’s implementation of the state’s Mathematics Curriculum Frameworks? Excellent      Good      Fair      Poor      Not Sure

20. Over the past three years, the education reform initiatives in my state have changed how I teach for the better: \_\_\_\_\_ Yes      \_\_\_\_\_ No      \_\_\_\_\_ Not Sure

**Section 3 -**

21. Statements concerning your high school’s mathematics curriculum and your instructional practices appear below. For each statement, please indicate: (a) the relative influence that aspect of the state’s Mathematics Curriculum Frameworks and assessment system currently has on your math instruction, (b) the relative influence that aspect had on your math instruction three years ago, and (c) your opinion regarding the influence you think that specific aspect will have on your math instruction three years from now. The scale for the influence ratings range from zero (no influence at all) to five (primary influence). *Please circle your responses.*

Statement	Current Influence 0=Not at all 5=Primary	Influence 3 Years Ago	Future Influence
a. Commitment to adequately preparing students to take the MCAS grade 10 mathematics test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
b. Matching the state’s Mathematics Frameworks to the curriculum that I teach	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
c. Success in implementing instructional changes to better reflect the state’s Mathematics Frameworks	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
d. Preparing students to take the MCAS grade 10 mathematics test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
e. Student performance on the MCAS grade 10 math test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
f. Improving students mathematics test scores on the MCAS grade 10 math assessment	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
g. Motivation of my students to do well on the MCAS grade 10 mathematics test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
h. Pressure from administrators to obtain the best MCAS math test results from my students	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
i. Motivation to teach mathematics as outlined in the state’s Mathematics Frameworks	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5

22. Do you think classroom time for test preparation of the MCAS grade 10 mathematics assessment has meant other valuable topics are **NOT** being taught?  
 Yes                       No                       Not Sure

23. Do you believe averaging all student performances on the MCAS grade 10 mathematics test is a good indicator of a high school’s overall educational success in teaching mathematics to its students?  
 Yes                       No                       Not Sure

24. Over the past three years, the number of students enrolled in more difficult mathematics courses in my school has:  
 increased                       decreased                      or                       remained roughly the same

**Section 4 –**

25. Do you believe the current system of MCAS score reporting at the high school level is adequate?  Yes  No  Not Sure

If not, how could the reporting of scores be improved? \_\_\_\_\_  
\_\_\_\_\_

26. In general, my students' course grades mirror their performance on the MCAS grade 10 mathematics test?  Yes  No  Not Sure

If not, please provide reasons for the difference. \_\_\_\_\_  
\_\_\_\_\_

27. The education reform initiative in my state has ensured that all students receive the best possible education.  Yes  No  Not Sure

Comment: \_\_\_\_\_  
\_\_\_\_\_

**Open-ended Questions –**

28. What are the positive and/or negative consequences of the MCAS grade 10 mathematics assessment?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

29. What concerns do you have regarding the impact the state's Curriculum Frameworks have had on mathematics instruction in your high school?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**THANK YOU FOR TAKING THE TIME TO COMPLETE THIS SURVEY!**  
*Please return survey to your department head* so he/she can mail to: Mary L. Zanetti,  
School of Education, UMASS, 179 Hills South, Amherst, MA 01003-4140.



APPENDIX E  
NON-TEACHER'S SURVEY

## NON-TEACHER'S SURVEY<sup>2</sup>

### PLEASE READ:

This survey is designed to determine how the MCAS grade 10 mathematics test and corresponding assessment system have affected your high school's math curriculum and you as a high school mathematics administrator. All information provided is confidential; please refer to the attached written consent form for more specific confidentiality information.

### INSTRUCTIONS:

This survey is divided into four sections. The first section provides us with some personal information about you. The second section is a list of statements that ask for your opinions based on an agreement scale. The third section includes questions about the impact your state's education reform initiative has had on your high school's math curriculum in the past, present, and future. The fourth and final section is comprised of a few selected response questions and two open-ended questions. Feel free to attach additional paper, if you would like to add comments. **The entire survey will take you approximately 15 minutes to complete.**

#### Section 1 – Personal Information

*Please complete the information below. Do not sign your name.*

1. Sex:  Female  Male
2. Are you a high school mathematics teacher at your school?  Yes  No
3. Please circle one of the following words that best describes your school district:  

Urban	Suburban	Rural
-------	----------	-------
4. What is the approximate number of students enrolled in your high school? \_\_\_\_\_
5. Please indicate the grade(s) you are currently teaching: \_\_\_\_\_
6. Please indicate the subjects you are currently teaching: \_\_\_\_\_

In addition, indicate the total number of years you have taught: \_\_\_\_\_

7. Are you a Massachusetts certified mathematics teacher?  Yes  No
8. **Education:** *Circle the highest degree you have completed.*  
Bachelor's degree      Masters degree      CAGS      Doctorate

---

<sup>2</sup> The author is grateful to Susan Bowles, Caryn McCrohon, Timothy O'Neil, Zhenhong Hu, and Stephen G. Sireci for the use of their two surveys in the creation of this survey.

**Section 2** – We are interested in your opinions regarding the impact the state’s Mathematics Curriculum Frameworks and assessment system has had on your high school’s math curriculum and your instructional practices of math teachers in your district’s high school. Please focus exclusively on the implementation of the curriculum rather than any other important issue that may be related (e.g., contract negotiations). *Please circle the response that best indicates your opinion.*

Statement	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
9. Teachers have received adequate professional development <b>at the state level</b> for implementing the Mathematics Curriculum Frameworks.	SA	A	U	D	SD
10. Teachers have received adequate professional development <b>at the district level</b> for implementing the Mathematics Curriculum Frameworks.	SA	A	U	D	SD
11. Teachers have changed my mathematics curriculum to better reflect the state’s Mathematics Curriculum Frameworks.	SA	A	U	D	SD
12. The mathematics department in my high school has changed course offerings due to the MCAS mathematics assessment system.	SA	A	U	D	SD
13. The mathematics department in my high school has changed course sequencing due to the MCAS mathematics assessment system.	SA	A	U	D	SD
14. Teachers’ instructional practices have changed as a result of the state’s Mathematics Curriculum Frameworks.	SA	A	U	D	SD
15. Teachers have made changes to my instructional practices due to my students’ performance on the MCAS grade 10 mathematics assessment.	SA	A	U	D	SD
16. The MCAS grade 10 mathematics assessment is a useful tool to improve the long-term education of students.	SA	A	U	D	SD
17. The MCAS grade 10 mathematics test results are being used in my high school to identify students who may need specialized instruction.	SA	A	U	D	SD
18. The school system has adequately provided remedial instruction and/or courses for those students who were unsuccessful on the MCAS grade 10 mathematics test.	SA	A	U	D	SD

19. Please circle the word that best describes your perception of the **morale** of your district’s high school mathematics teacher based only on your district’s implementation of the state’s Mathematics Curriculum Frameworks?

Excellent      Good      Fair      Poor      Not Sure

20. Over the past three years, the education reform initiatives in my state have changed how high school mathematics teachers in my district teach for the better:

\_\_\_\_\_ Yes                      \_\_\_\_\_ No                      \_\_\_\_\_ Not Sure

**Section 3 -**

21. Statements concerning your high school’s mathematics curriculum and teachers’ instructional practices appear below. For each statement, please indicate your opinion regarding: (a) the relative influence that aspect of the state’s Mathematics Curriculum Frameworks and assessment system currently has on the math instruction in your district’s high school, (b) the relative influence that aspect had on the math instruction in your district’s high school three years ago, and (c) your opinion regarding the influence you think that specific aspect will have on the math instruction in your district’s high school three years from now. The scale for the influence ratings range from zero (no influence at all) to five (primary influence). *Please circle your responses.*

Statement	Current Influence 0=Not at all 5=Primary	Influence 3 Years Ago	Future Influence
a. Commitment to adequately preparing students to take the MCAS grade 10 mathematics test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
b. Matching the state’s Mathematics Frameworks to the curriculum that educators teach	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
c. Success in implementing instructional changes to better reflect the state’s Mathematics Frameworks	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
d. Preparing students to take the MCAS grade 10 mathematics test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
e. Student performance on the MCAS grade 10 math test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
f. Improving students mathematics test scores on the MCAS grade 10 math assessment	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
g. Motivation of your district’s high school students to do well on the MCAS grade 10 mathematics test	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
h. Pressure from administrators to obtain the best MCAS math test results from your district’s high school students	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5
i. Motivation of educators to teach mathematics as outlined in the state’s Mathematics Frameworks	0 1 2 3 4 5	0 1 2 3 4 5	0 1 2 3 4 5

22. Do you think classroom time for test preparation of the MCAS grade 10 mathematics assessment has meant other valuable topics are NOT being taught?

\_\_\_\_\_ Yes                      \_\_\_\_\_ No                      \_\_\_\_\_ Not Sure

23. Do you believe averaging all student performances on the MCAS grade 10 mathematics test is a good indicator of a high school's overall educational success in teaching mathematics to its students?  Yes  No  Not Sure

24. Over the past three years, the number of students enrolled in more difficult mathematics courses in your district's high school has:  
 increased  decreased or  remained roughly the same

**Section 4 –**

25. Do you believe the current system of MCAS score reporting at the high school level is adequate?  Yes  No  Not Sure

If not, how could the reporting of scores be improved? \_\_\_\_\_  
\_\_\_\_\_

26. In general, my district's high school students' course grades mirror their performance on the MCAS grade 10 mathematics test?

Yes  No  Not Sure

If not, please provide reasons for the difference. \_\_\_\_\_  
\_\_\_\_\_

27. The education reform initiative in my state has ensured that all students receive the best possible education.  Yes  No  Not Sure

Comment: \_\_\_\_\_  
\_\_\_\_\_

**Open-ended Questions –**

28. What are the positive and/or negative consequences of the MCAS grade 10 mathematics assessment?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

29. What concerns do you have regarding the impact the state's Curriculum Frameworks have had on mathematics instruction in your high school?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**THANK YOU FOR TAKING THE TIME TO COMPLETE THIS SURVEY!**  
*Please return survey to your department head* so he/she can mail to: Mary L. Zanetti,  
School of Education, UMASS, 179 Hills South, Amherst, MA 01003-4140.

APPENDIX F  
FOCUS GROUP PROTOCOL

Good afternoon,

Thanks so much for agreeing to participate in today's focus group session. My name is Mary Zanetti and I'm a doctoral candidate at UMass Amherst. I am currently completing my dissertation study, which investigates the positive and negative consequences of this state's grade 10 state-mandated mathematics test. As a future education measurement professional, I believe it is very important to take the time to obtain educators' perceptions about the strengths and weaknesses and the intended and unintended consequences of this test and the state assessment system as a whole. More specifically, I would like to hear how this testing program has affected your high school's math curriculum, course offerings, and instructional practices. So that's why I'm here today.

I have approximately 7 questions/topics that I would like to discuss, but I'd like to use these issues as a guide rather than turning my visit into a question and answer session. As you all know, I have distributed and collected the completed written consent form, which outlined confidentiality issues pertaining to this study and indicates that today's session will be recorded. I will be the only person using today's transcript and it will be used in the data analyses portion of my study.

I thought we could start the conversation with a review of your high school's average student math test scores over the past 3 years (visual aid used here).

What are your immediate thoughts when you view these test scores?

How has this test and your students' performance on it affected the math curriculum in your building? course offerings? course sequencing?

Follow-up questions:

What's your opinion of a standards-based curriculum?

What's your opinion of the Frameworks?

How has this test and your students' performance on it affected your own instructional practices?

How has this test and the whole assessment system affected your morale and motivation?

Has it improved student learning?

Follow-up question:

Do the math grades received by your students generally mirror their scores on the mandated test? If not, why not?

What changes would you make to this test or the assessment system as a whole?

APPENDIX G  
INTERVIEW PROTOCOL



Good afternoon,

Thanks so much for agreeing to participate in this interview. My name is Mary Zanetti and I'm a doctoral candidate at UMass Amherst. I am currently completing my dissertation study, which investigates the positive and negative consequences of this state's grade 10 state-mandated mathematics test. As a future education measurement professional, I believe it is very important to take the time to obtain key educator's perceptions of the strengths and weaknesses and the intended and unintended consequences of this test and the state assessment system as a whole. More specifically, I would like to hear how this testing program has affected your high school's math curriculum, course offerings, and the instructional practices of your district's high school teachers. So that's why I'm here today.

I have approximately 7 questions/topics that I would like to discuss, but I'd like to use these issues as a guide during our conversation rather than turning my visit into a question and answer session. As you know, I have asked you to sign a written consent form, which outlined confidentiality issues pertaining to this study. I hope you don't mind me taking notes. Keep in mind, I will be the only person using these notes in the data analyses portion of my study.

I thought we could start the conversation with a review of your district's grade 10 math scores over the past 3 years (visual aid used here).

What are your immediate thoughts when you view these test scores?

How has this test and your district's grade 10 students' performance affected the math curriculum in the high school? course offerings? course sequencing?

Follow-up questions:

What's your opinion of a standards-based curriculum?

What's your opinion of the Frameworks?

How has this test and the grade 10 students' performance in your district affected the instructional practices in that building?

How has this test and the whole assessment system affected your morale and motivation? morale and motivation of the math teachers at the high school?

Has it improved student learning?

Follow-up question:

Do the math grades received by the grade 10 mathematics students in your district's high school generally mirror their scores on the mandated test? If not, why not?

What changes would you make to this test or the assessment system as a whole?

APPENDIX H  
TELEPHONE CONTACT SUMMARY SHEET

**Name:** \_\_\_\_\_ **Date:** \_\_\_\_\_

**Address:** \_\_\_\_\_

**Phone number:** \_\_\_\_\_ **School:** \_\_\_\_\_

**Location, Date, and Time of Focus Group Session** [*include # of participants*]:

\_\_\_\_\_

**Recommended names and their telephone numbers for interviews:**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

**Name, Location, Date, and Time of each Interview:**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

**Notes:**

## BIBLIOGRAPHY

- Adler, P. A., & Adler, P. (1994). Observational techniques. In N.K. Denzin & Y.S. Lincoln (Eds.), Handbook of qualitative research (pp. 377-392). Thousand Oaks, CA: Sage.
- Alexander, K., & Alexander, M. D. (1992). American public school law (3<sup>rd</sup> ed.). St. Paul, MN: West Publishing.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Psychological Association.
- Babbie, E. (1998). The practice of social research (4<sup>th</sup> ed.). Belmont, CA: Wiley.
- Baker, E. L., Linn, R. L., Herman, J. L., & Koretz, D. (2002). Standards for educational accountability systems (National Center for Research on Evaluation, Standards, and Student Testing Policy Brief 5). Los Angeles, California: UCLA, Graduate School of Education and Information Studies.
- Baker, E. L., O'Neil, H. F., & Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. American Psychologist, 48(12), 1210-1218.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Chudowsky, N., & Behuniak, P. (1998). Using focus groups to examine the consequential aspect of validity. Educational Measurement: Issues and Practice, 17(4), 28-38.
- Creswell, J. W. (1994). Research design: Qualitative & quantitative Approaches. Thousand Oaks, CA: Sage Publications.
- Creswell, J. W. (1998). Qualitative inquiry and research: Choosing among five traditions. Thousand Oaks, CA: Sage Publications.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.
- Denzin, N. K. (1978). The research act: A theoretical introduction to sociological

methods (2<sup>nd</sup> ed.). New York: McGraw-Hill.

- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). Educational research: An introduction. White Plains, NY: Longman.
- Green, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. Educational Evaluation and Policy Analysis, 11(3), 255-274.
- Greenberger, S. S. (2001, December 25). MCAS foes undeterred by test trend. The Boston Globe, p. B4.
- Jackson, R. M., & Rothney, J. W. M. (1961). A comparative study of the mailed questionnaire and the interview in follow-up studies. Personnel and Guidance Journal, 39, 569-571.
- Jarrell, M. G. (2000). Focusing on focus group use in educational research. Paper presented at the Mid-South Educational Research Association Annual Meeting, Bowling Green, KY (ERIC Document Reproduction Service No. ED 448 167).
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. Administrative Science Quarterly, 24, 602-611.
- Kane, M. T. (1992). An argument-based approach to validity. Psychological Bulletin, 112, 527-535.
- Khatti, N., Kane, M. B., & Reeve, A. L. (1995). How performance assessments affect teaching and learning [Research Report]. Educational Leadership, 53(3), 80-83.
- Lane, S., Parke, C. S., & Stone, C. A. (1998). A framework for evaluating the consequences of assessment programs. Educational Measurement: Issues and Practice, 17(2), 24-28.
- LeCompte, M. D., & Preissle, J. (1993). Ethnography and qualitative design in educational research (2<sup>nd</sup> ed.). San Diego, CA: Academic Press.
- Linn, R. L. (1989). Educational measurement (3<sup>rd</sup> ed., pp. 13-103). New York: Macmillan.
- Linn, R. L. (1994). Performance assessment: Policy promises and technical measurement standards. Educational Researcher, 23(9), 4-14.
- Linn, R. L. (1997). Evaluating the validity of assessments: The consequences of use. Educational Measurement: Issues and Practice, 16(2), 14-16.
- Linn, R. L. (1998). Partitioning responsibility for the evaluation of the consequences

of assessment programs. Educational Measurement: Issues and Practice, 17(2), 28-30.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. Educational Researcher, 31(6), 3-16.

Maguire, T., Hattie, J., & Haig, B. (1994). Construct validity and achievement assessment. The Alberta Journal of Educational Research, 40, 109-126.

Marshall, C., & Rossman, G. B. (1999). Designing qualitative research (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage Publications.

Massachusetts Department of Education. (1996, February). State plan to eliminate the general track. Retrieved from Massachusetts Department of Education Website on the World Wide Web:  
<http://www.doe.mass.edu/edreform/ELIM/gentrac96.html>

Massachusetts Department of Education. (2001, January). School performance rating process implementation guidelines and methods. Retrieved from Massachusetts Department of Education Website on the World Wide Web:  
<http://www.doe.mass.edu/edreform/accountability.html>

Mathison, S. (1988). Why triangulate? Educational Researcher, 17(2), 13-17.

Mehrens, W. A. (1997). The consequences of consequential validity. Educational Measurement: Issues and Practice, 16(2), 16-18.

Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? Education Policy Analysis Archives, 6(13).

Messick, S. (1980). Test validity and the ethics of assessment. American Psychologist, 35, 1012-1027.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3<sup>rd</sup> ed., pp. 13-103). New York: Macmillan.

Messick, S. (1995). Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50, 741-749.

Morse, J. .M. (1991). Approaches to qualitative-quantitative methodological triangulation. Nursing Research, 40(1), 120-123.

Moss, P. .A. (1998). The role of consequences in validity theory. Educational Measurement: Issues and Practice, 17(2), 6-12.

- National Board of Professional Teaching Standards (2001). NBPTS Home webpage: <http://www.nbpts.org>
- Ponsford, B. J., & Masters, L. A. (1998). How to use focus groups in higher education research. College & University, 73(3), 2-9.
- Popham, W. J. (1997). Consequential validity: Right concern--- wrong concept. Educational Measurement: Issues and Practice, 16(2), 9-13.
- Reichardt, C. S. & Rallis, S. F. (Eds.). (1994). The qualitative-quantitative debate: New perspectives. San Francisco: Jossey-Bass.
- Richardson, L. (1994). Writing: A method of inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 516-529). Thousand Oaks, CA: Sage.
- Shepard, L. A. (1993). Evaluating test validity. Review of Research in Education, (19), 405-450.
- Shepard, L. A. (1997). The centrality of test use and consequences for test validity. Educational Measurement: Issues and Practice, 16(2), 5-8, 13, 24.
- Smith, M. L. (1994). Qualitative plus/versus quantitative: The last word. In C. S. Reichardt & S. F. Rallis (Eds.), The qualitative-quantitative debate: New perspectives (pp37-44). San Francisco: Jossey-Bass.
- Stewart, D. W., & Shamdasani, P. N. (1990). Focus groups: Theory and practice. Newbury Park, CA: Sage Publications.
- Taleporos, E. (1998). Consequential validity: A practitioner's perspective. Educational Measurement: Issues and Practice, 17(2), 20-24.
- Tenopyr, M. L. (1996). Construct-consequences confusion. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Vaishnav, X. (2002, January 27). Ranking of schools draws official ire: Activist says ratings help inform parents. The Boston Globe, p. B6.
- Wilson, V. A. (1998). Qualitative research: An introduction. Purposes, methodology, criteria for judgment, and a rationale for mixed methodology, (ERIC Document Reproduction Service No. ED 423 285).





